Social Learning and Parameter Uncertainty in Irreversible Investment

----Evidence from Greenhouse Adoption in Northern China

Honglin Wang
Michigan State University
Email: wanghon7@msu.edu

Thomas Reardon
Michigan State University
Email:reardon@msu.edu

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Orlando, FL, July 27-29, 2008.

Social Learning and Parameter Uncertainty in Irreversible Investment

----Evidence from Greenhouse Adoption in Northern China

Honglin Wang[1]
Thomas Reardon[2]

Abstract: This paper introduces social learning into irreversible investment theory through parameter uncertainty, and shows that social learning could reduce parameter uncertainty to facilitate irreversible investment technology adoption. The theoretic model is tested by using household level data from energy saving greenhouse adoption in northern China, and empirical evidences are consistent with the theory: social learning has significant positive impacts on greenhouse adoption, while market volatility discourages the adoption.

Key words: Social Learning, Technology Adoption, Irreversible Investment, Parameter Uncertainty, Energy Saving Greenhouse.

JEL codes: O12, O31, C61, D83, G12.

[1] Honglin Wang is a Ph.D Candidate in both Economics and Ag Economics, Michigan State University.
[2] Thomas Reason is a Professor in Ag Economics Department, Michigan State University.

Social Learning and Parameter Uncertainty in Irreversible Investment

----Evidence from Greenhouse Adoption in Northern China

1. Introduction

A high rate of technological change is a primary feature of modern agriculture (Schultz) and new technology adoption is a key to rising out of poverty for many poor farmers in developing countries. However, agricultural innovations are often adopted slowly or at times not adopted at all. Many previous studies have focused on removing constraints of adoptions such as lack of credit, limited access to information and insufficient human capital etc[3].

On the other hand, a lot of technology adoption is lumpy investment such as greenhouse, tractors and irrigation because of embedded lumpy investment feature of these technologies. To these kinds of technology adoption, it is not only a change from old technologies to new ones but also an investment decision. Farmers might not adopt a technology when they could adopt it because they feel it is not optimal timing to make the investment. Unfortunately, few literatures paid attention to this aspect of technology adoption, and this paper is designed to address the gap.

Moreover, lumpy investment such as greenhouse or machinery usually entails significant irreversible investment and incomplete information with respect to the performance of the invested goods, its reliability, and appropriateness of their operation (Sunding and Zilberman, 2000). The irreversibility of investment causes more risk concerns. In developed countries, the prevailing approach to address such risk is to form

[3] Feder, Just and Zilberman (1985) wrote a comprehensive review about constraints to technology adoptions.

a product-backup system such as warranties and established dealerships equipped to repair breakdowns. However, very few studies have been done on how small farmers in developing countries deal with such risk where the product-backup system is either absent or imperfect.

Therefore, how do small farmers deal with risk from incomplete information and irreversible investment in developing countries? This paper tries to provide evidences that learning from others (social learning) could be one of solutions. Economists recognized that the timing of adoption in irreversible investment varies across adopters. The difference in the timing suggests that some agents choose to wait while others adopt new technology quickly. The delaying of adoption may enable agents to obtain more information, reduce overall uncertainty and increase expected discounted benefits by avoiding irreversible investment when it is not worthwhile (Sunding and Zilberman, 2000). This suggests that the orthodox theory of investment evaluation (Net Present Value Rule) which ignores the value of delaying could be misleading. The ability to delay an irreversible investment is like holding an "option" because an agent or a firm has the right but not the obligation to buy an asset at some future time (Dixit and Pindyck, 1994, D-P model hereafter).

Real option approach has been applied to study how uncertainty and irreversibility affect agents' adoption behavior (Hasset and Metcalf ,1992; Zilberman et al, 1994; Olmstead and Rhode,1998; Nelson and Amegbeto, 1998). However, like D-P model, most of these studies assume all parameters in the dynamic process are known to agents, the only uncertainty in these models comes from stochastic process itself (volatility), all parameters (both drift and volatility) are assumed to be known. However, economists recognize that it is unrealistic to assume agents know all parameters or agents can estimate all parameters precisely using past data (Merton, 1980). In other words, there exists another layer of uncertainty: parameter uncertainty.

Some studies have been done on how parameters uncertainty affects investment behavior in finance literature (Merton, 1980; Gennotte, 1986; Brennan, 1998; Xia, 2001; Abasov, 2005; Huang and Liu, 2007). Merton (1980) proved that: when the observation period is finite and trading is continuous, it is possible to agents to estimate the variance of return precisely because the estimator of variance (volatility) would converge as time interval become smaller and smaller (more and more observations). However, the estimator of expected return, nothing is gained in terms of accuracy of the expected return estimate by choosing finer observation intervals. This well-known result suggests it is reasonable to assume agent know the parameter of volatility but not expected return.

Gennotte (1986) derived optimal estimators for the unobservable expected instantaneous returns using past realized returns and establish the separation theorem which allows the estimation to be solved in two separate steps. Tools of non-linear filtering theory (Liptser and Shiryayev, 1978) was introduced to derive the optimal estimators of drift when agents continuously update their beliefs as time evolves, and the result shows there exists a new Brownian motion process determining the conditional moments, but contains no additional information on future realizations of returns.

Brennan (1998) discusses learning effects based on Gennotte (1986) work and shows that future learning about drift on risky asset induces the agents to take larger or smaller position in risky asset, the direction depends on his risk tolerance. Xia (2001) generalized Brennan (1998) work by introducing return predictability and shows that uncertainty about the predictive relation leads to a state-dependent relationship between the optimal portfolio choice and the investment horizon. Huang and Liu (2007) introduced periodic filtering (periodic new updates) combined with continuous filtering (continuous new updates), and endogenized the news frequency and news accuracy to show that rational inattention to important news may make investors over or underinvest.

In this paper we focus on how social learning affects irreversible technology adoption under parameter uncertainty. We assume value of project follows a geometric Brownian motion, and the parameter of volatility is known while the drift is not known to agents. Agents want to find the optimal stopping point to make irreversible investments (adoption). Agents learn about the parameter (drift) in two ways: first, agent can extract information on the drift from their observations of realized past returns by continuous filtering. Second, agent can obtain direct information about the drift periodically from early adopters in his social network, but the information comes with observational errors. Since chatting with people costs little in a village, we assume there is no information cost. We derive a theoretic model including both continuous filtering and periodical filtering (social learning), and obtain differential equations to fully characterize the optimal stopping problem under the assumption that agent's uncertainty about the drift follows Gaussian distribution. Since it is impossible to get analytical solutions in this case, numerical approach is employed to get solutions, which shows that the higher uncertainty and higher expected return of the drift would induce the higher trigger value, hence more difficult to adopt. Social learning affects adoptions through its impact on agents' belief of the drift. The more social learning, the less uncertainty of the drift, hence the easier the agent adopts the technology. On the other hand, social learning also affects the expected returns, and in this case social learning could have positive or negative impacts on adoption, which depends on agent's initial belief of the drift and average beliefs in his social network.

Moreover, we test the theory using household level data about greenhouse adoption in northern China. Our sample includes 700 randomly selected households in 70 villages in Shandong Province. Linear probability model (LPM) and Probit model are used to estimate coefficients. The empirical results are consistent with theoretical model: social learning has significant positive impacts on greenhouse adoption, the more adopters in a farmers' social network (social learning), the more likely the farmer would adopt greenhouse. The results suggest that in developing countries, when small farmers face irreversible technology adoption under incomplete information, social learning could be an effective way to deal with such uncertainty and facilitate adoption.

It is worth to mention social learning in this paper is different to previous social learning studies in technology adoption (Besley and Case 1997; Foster and Rosenzweig 1995; Conley and Udry 2005; Muchshi 2004 and Bandiera and Rasul, 2006, Yamauchi, 2007). Firstly, in most above studies, new innovation means new variety, which is not an irreversible investment. Therefore, there is no option value involved in these studies. Secondly, Social learning affects expected return in those models, but social learning is kind of exogenous to the dynamic optimization problem. In other words, social learning is not a choice variable in the dynamic programming. On the contrary, in this paper agents choose the optimal timing to stop waiting, which is equivalent to choose the amount of social learning.

In brief, this paper is aimed at contributing to the technology adoption literature in three ways: first, this paper tries to fill the gap between irreversible investment and technology adoption literature in developing countries. Second, this paper introduces social learning into irreversible investment literature through parameter uncertainty[aec1]. Third, this paper is the first paper to empirically test how social learning affects irreversible investment technology adoption with parameter uncertainty.

2. Energy Saving Greenhouse in Northern China

From the 1950s to 1970s, China was known for its focus on grains self-sufficiency policy– the "iron rice bowl" Policy. At the beginning of economic liberalization in early 1980s, China introduced a policy focusing on agricultural diversification to add vegetable products to the grain foundation. Rapid economic growth created increasing demand for high value vegetables. However, poor infrastructure and high energy costs prevented shipping perishable products from south to north, and affordable fresh vegetables were still not available in winters in northern China.

It is not surprising that this huge demand for cheap fresh vegetables induced energy saving greenhouse technology (ESGT). ESGT not only changed food consumption pattern of hundreds and millions consumers, but also the pattern of agriculture production in northern China. Today, China is the biggest economy in horticultural production and contributes to one third of global horticultural output in 2003. In terms of vegetable production, China produced 520 million tons vegetables and accounts for 40% of global vegetable production. The total area of greenhouse vegetable production reached 150,000 ha in 2004 (Chinese Agriculture Yearbook, 2006). At least half a million farmers adopted ESGT.

There are mainly two types of protected facility used by Chinese farmers for vegetable production[4]. The first type is the modern greenhouse which was initially introduced into China by the government in the late 1970s from the US, Europe and Japan. This type of greenhouse has a fixed structure made of steel and glass with active climate control and hydroponics. But modern greenhouses were expensive and not modified to fit in the local conditions. Even though the government put a lot of efforts to promote it, the expensive modern greenhouse failed to be widely adopted by Chinese farmers. Till 1999, this kind of greenhouse contributed less than 0.2% of vegetable production area in China (Wan, 2000).

However, Chinese farmers somehow got ideas from these "alien boxes" and modified the expensive greenhouse to cheap and energy saving greenhouses using cheap local materials and by-products. This kind of energy saving greenhouse has a fixed structure made of bamboo and pounded clay with or without some active environmental control. Since heating costs inside greenhouse is expensive to farmers, this kind of greenhouse rarely has active environmental control, which is only used in coldest weeks in cold areas. For most energy saving greenhouses in Shandong province[aec2], the only energy source of greenhouse is sunshine through whole winter. To do this successfully, the orientation of greenhouses is carefully chosen to make sure greenhouse absorb sunshine as much as possible. Moreover, covering materials are used to keep temperature high enough to

[4] The third type is all kinds of small shading houses used by Chinese farmers to extend the production season a little bit compared with open field production. We don't include these shading houses into greenhouses. The easiest way to distinguish the shading houses with greenhouses is the back-wall. Greenhouses have back-wall and the shading houses don't. We use the back-wall criterion in our survey.

allow vegetables to survive at night. By this way, Shandong farmers could produce many kinds of vegetables such as tomato, cucumber and peppers etc in northern China's cold winter.

The yield of greenhouse crops is much higher than open-field ones. For example, average annual tomato yield in greenhouse is 200 tons/ha compared to 40 tons/ha in an open field. Several factors, including labor intensive production, contribute to the high yield. The popular greenhouse size in Shandong province is 60 meters long and 10 meters wide, two full time workers are needed for a greenhouse. Secondly[aec3], all year long greenhouse production usually lasts more than 8 months because the temperature inside greenhouse is high enough in the winter. Thirdly, many kinds of high technologies are applied in greenhouse production such as high quality varieties, organic fertilizers[5] etc. The construction cost of ESGT is roughly $\$4/m^2$, much cheaper compared to modern greenhouse ($\$80/m^2$) because of most greenhouse is made of cheap by-products and local materials. However, it is still a big investment to small farmers. For example, if a greenhouse is 60 meter long and 10 meter wide, the construction cost would be about $2400, while the average Chinese farmer annual net income per capita is less than $500 in 2005.

3. Irreversible Investment with Social Learning

Investment is often irreversible: once installed, capital has little or no value unless used in Production (Bertola and Caballero 1994). To build a greenhouse, farmers usually spend months to build the main part of greenhouse—rear-wall of greenhouse, which is

[5] It is worthy to mention that the organic fertilizer is applied much more in greenhouse than open field because farmers need grow vegetables in the same small piece of land intensively year after year. Chemical fertilizers decrease the quality of land rapidly, but organic fertilizers don't.

usually made of clay bricks. The clay bricks would go back dirt if farmers demolish greenhouses. Other materials of greenhouse are agricultural by-products such as straw mattress and bamboo beam, and salvage value of these materials are very low. Therefore, greenhouse adoption can be viewed largely irreversible investment[6].

Moreover, greenhouse adoption is an investment with uncertainty. There are two layers of uncertainties: the first layer uncertainty comes from the stochastic process of project value. In the stochastic process the current state determines only the probability distribution of future states, not actual value. So even one knows all the parameters of the stochastic process, she never knows actual values in future. The second layer of uncertainty comes from incomplete information of parameters of the stochastic process. Parameter uncertainty could be affected by estimation or learning. It is well know that the volatility of stochastic process can be estimated pretty well using past realized returns if the trading is continuous, however, estimating drift precisely using past realized returns is quite difficult (Merton, 1980).

Since it is unrealistic to assume agents would know the drift, agents have to optimize their investments under incomplete information. However, agent can extract (or filter) information on future expected instantaneous return from their observation of past returns (Gennotte, 1986). In other words, filtering information (learning) would affect the agents' perception of the parameter. That is why social learning could affect agent's decisions.

In the following subsection, we firstly briefly review irreversible investment under known parameters (D-P model), then we introduce parameter uncertainty with continuous learning and periodical learning.

---

[6] In the survey, less than 10% farmers abandoned the greenhouse after they adopted and most of cases are due to natural disasters or other non-economic reasons.

3.1 Irreversible Investment with Known Parameters

In this subsection, we use D-P model to discuss what factors affect agents' irreversible investments under the first layer of uncertainty (with known parameters).

Observers of business find that firms usually invest in projects that are expected to yield a return typically three or four times the cost of capital (Summer, 1987). On the other hand, firms stay in business for long periods while absorbing operating losses, and price can fall substantially below average variable cost without inducing disinvestment and exit (Dixit and Pindyck, 1994). This puzzle was the motivation of the paper "the value of waiting to invest" (McDonald and Siegel, 1986). The key question in that paper was at what point it is optimal to pay a sunk cost (I) in return for an investment whose value is V, which evolves according to a geometric Brownian motion over time.

$$dV = \alpha V dt + \sigma_v V dz \qquad (1)$$

The equation (1) is the Brownian motion with drift where dz is the increment of a wiener process and α is the drift parameter, σ_v is the variance parameter. As we mentioned before, this model assumes agents have complete information of parameters of the geometric Brownian motion, in other words, both α and σ_v are assumed to be known. However, because it is stochastic process, even though new information arrives over time, the future value of the project is always uncertain (Dixit and Pindyck, 1994).

The ability to delay investment creates the value of the option because of irreversibility and uncertainty. We denote the value of the option to invest as $F(V)$. We want to find a rule to maximize this value. Since the payoff from investing at time t is $V_t - I$, we want to maximize its expected present value:

$$F(V) = MaxE[e^{-\rho T}(V_t - I)] \qquad (2)$$

Where t is the unknown future time that the investment is made, ρ is the discount rate.

Write the Bellman equation for this optimal stopping problem using dynamic programming theory:

$$\rho F(V)dt = E[dF(V)] \qquad (3)$$

Then we expand $dF(V)$ using ITO's lemma in continuous time.

$$dF(V) = F'(V)dV + 0.5F''(V)d(V)^2 \qquad (4)$$

Substituting equation (1) for dV into equation (4) and noting that $E(dz) = 0$, hence the Bellman equation (3) becomes:

$$0.5\sigma_v^2 V^2 F''(V) + (\rho - \delta)VF'(V) - \rho F(V) = 0 \qquad (5)$$

Where $\delta \equiv \rho - \alpha$, we assume $\delta > 0$.

In addition, $F(V)$ must satisfy the following three boundary conditions:

$$F(V) = 0 \qquad F(V^*) = V^* - I \qquad F'(V^*) = 1$$

To find $F(V)$, we must solve equation (5) subject to the three boundary conditions. It is easy to see that, to satisfy the first boundary condition given equation (5), the solution must take the form

$$F(V) = AV^{\beta} \qquad (6)$$

Where A is an unknown constant to be solved and β is a known constant whose value depends on the parameters σ_v, ρ and δ. Then we use the other two boundary conditions to solve the A and trigger value V^*. We can get:

$$V^* = \frac{\beta}{\beta - 1} I \qquad (7)$$

$$A = \frac{V^* - I}{V^{*\beta}} \quad (8)$$

The equation (6)-(8) gives the value of investment opportunity and the optimal investment rule. The critical value V^* at which it is optimal to invest. From the equation (7), we can see that V^*>I, which implies that the conventional NPV rule is wrong, uncertainty and irreversibility drive a wedge between the critical value V^* and I.

Then solve for β from equation (6) and (5):

$$0.5\sigma_v^2\beta(\beta-1)+(\rho-\delta)\beta-\rho=0 \quad (9)$$

$$\beta_1 = \frac{1}{2} - \frac{\rho-\delta}{\sigma_v{}^2} + \sqrt{[\frac{\rho-\delta}{\sigma_v{}^2} - \frac{1}{2}]^2 + \frac{2\rho}{\sigma_v{}^2}} > 1^{7} \quad (10)$$

Differentiate the equation (9) totally, denote the left hand side of equation (9) as Q, we can get

$$\frac{\partial Q}{\partial \beta}\frac{\partial \beta}{\partial \sigma_v} + \frac{\partial Q}{\partial \sigma_v} = 0 \quad (11)$$

Since $\frac{\partial Q}{\partial \beta} > 0$, $\frac{\partial Q}{\partial \sigma_v} > 0$, we can easily see $\frac{\partial \beta}{\partial \sigma_v} < 0$ from equation (11), hence $\frac{\partial V^*}{\partial \sigma_v} > 0$.

Which means the greater is the amount of uncertainty over future of V, the larger is the trigger value V^*. Similarly, it can be shown that $\frac{\partial V^*}{\partial \delta} < 0$, which means the more difference between interest rate(ρ) and drift (α), the lower trigger value V^* given $\delta \equiv \rho - \alpha > 0$.

[7] The other root of β is negative, so we omit it here.

The D-P model shows that the trigger value is known to agents because all parameters are known. In other words, there is no uncertainty about the trigger value itself. Therefore, there is no role for learning in the D-P model.

3.2 Irreversible Investment with Parameter Uncertainty

However, it is unrealistic to assume all parameters known, particularly the drift (α). Therefore, parameter uncertainty adds the second layer of uncertainty in the model, and that is why learning could affect agents' investment decision because learning could reduce parameter uncertainty and affect agent's perception of the drift.

3.2.1 Continuous Learning without Periodical News[8]

Similar to the D-P model, we assume project value V_t evolves according to a geometric Brownian motion:

$$dV_t = \alpha_t V_t dt + \sigma_v V_t dZ_t \qquad (12)$$

Note that α_t is time dependent variable now. As we mentioned before, σ_v can be assumed to be known to agents if trading is continuous. However, the drift(α_t) is not known and realized project value (V_t) is assumed to be observable. Like in the D-P model, the agent wants a rule to maximize the following value:

$$Max \ E[e^{-\rho T}(V_t - I)] \qquad (13)$$

Since the project value V_t is observable, this filtration $\{F_t^A\}_{t\geq 0}$, which define the agent's information set, in this case, is same with $\{F_t^V\}_{t\geq 0}$, which generated by process $\{V_t\}_{t\geq 0}$.

[8] In this subsection, I follow Abasov (2005) framework about Gaussian distribution.

Hence the agent's maximization problem becomes solving an optimal stopping problem based on the information filtration $\{F_t^V\}_{t\geq 0}$ available to them:

$$\begin{aligned} &\underset{T\in\Theta}{Max}\, E[e^{-\rho T}(V_t - I) \mid F_t^V] \\ &dV_t = V_t[\alpha dt + \sigma_v dz_t] \end{aligned} \tag{14}$$

where Θ is the set of all stopping times adapted to $\{F_t^V\}_{t\geq 0}$. Since the drift is not known, so the Brownian motion process Z_t is not observable as well. Agents need to make an inference at each time t based on information filtration $\{F_t^V\}_{t\geq 0}$ to solve the maximizing problem. The conditional expectation of the drift at time t is:

$$m_t = E(\alpha \mid F_t^V) \tag{15}$$

According to Liptser and Shiryaev (1978), the problem can be solved based on the conditional distribution (m_t), which define a new observable Brownian motion Z_t'

$$dZ'_t = dZ_t + \frac{1}{\sigma_v}(\alpha - m_t)dt \tag{16}$$

Meanwhile, m_t evolves dynamically as well. Before we discuss the evolution of m_t, we need define agent' initial belief of the drift, which is normal distribution $\alpha \sim N(m_0, \gamma_0)$, then we are ready to discuss the dynamics of m_t: also according to Liptser and Shiryaev (1978),

$$dm_t = \frac{\gamma_t}{\sigma_v} dZ_t^{'} \tag{17}$$

where γ_t is the conditional variance of α at time t, which can be defined as

$$\gamma_t = E[(\alpha - m_t)^2 \mid F_t^V] \tag{18}$$

and conditional variance also has its dynamic evolution:

$$d\gamma_t = -\frac{\gamma_t^2}{\sigma_v^2}dt \qquad (19)$$

According to separation theorem (Gennotte, 1986), the path Z’$_s$ $(0 \le s \le t)$ determines the conditional moments (m_t, γ_t), however, they contains no additional information on future realizations of V_t. Agents can make their decision only based on (m_t, γ_t and V_t). In other words, (m_t, γ_t, V_t) describe sufficient knowledge of the state at time t, and agents don’t have to remember past information to make current decisions because (m_t, γ_t, V_t) determine the probability distribution of V and m over the next infinitesimal interval [t, t+dt].

So if we combine the above equations together, we can fully characterize the investment problem using the observable processes:

$$\begin{cases} Max\, E[e^{-\rho T}(V_t - I)] \\ dV_t = V_t[m_t dt + \sigma_v dz'_t] \\ dm_t = \dfrac{\gamma_t}{\sigma_v} dZ_t' \\ d\gamma_t = -\dfrac{\gamma_t^2}{\sigma_v^2} dt \end{cases} \qquad (20)$$

Since $d\gamma_t = -\frac{\gamma_t^2}{\sigma_v^2}dt$ is the Riccati type of differential equation, so we can simplify equation (20) as:

$$\gamma_t = \frac{\gamma_0 \sigma_v^2}{\gamma_0 t + \sigma_v^2} \qquad (21)$$

From equation (21), we can see that the conditional variance of drift decrease in t, which means the longer an agent observe V_t, the less uncertainty about the parameter, which is consistent with Merton (1980) results: the uncertainty of expected return is not related to the number of observations, but related to the length of observation period. However, the conditional mean about draft can fluctuate up and down, which depends on the new observable Brownian motion $Z_t^{'}$.

Substitute equation (21) into equation (20), we can get:

$$\begin{cases} Max\, E[e^{-\rho T}(V_t - I)] \\ dV_t = V_t[m_t dt + \sigma_v dz'_t] \\ dm_t = \dfrac{\gamma_0 \sigma_v}{\gamma_0 t + \sigma_v^2} dZ'_t \end{cases} \qquad (22)$$

Therefore, this optimal stopping problem is defined by three equations, and the original problem with unobservable Brownian motion Z_t is reduced to the three equations with observable Brownian motion $Z_t^{'}$. However, it is still impossible to get an analytical solution to this kind of differential equations in general. So we have to depend on numerical approaches to get solutions.

Abasov (2005) uses numerical approach to solve the equation (22), and show that option value increase as uncertainty goes up, which makes perfect sense since extra layer of uncertainty from parameter (drift) should entail more valuable option and higher trigger value. He also shows that the higher m_t induce the higher trigger value, hence the longer the agent is willing to wait. This result is completely consistent with foregoing D-P model.

A brief comparison between parameter uncertainty model and D-P model helps us understand the model better and facilitate next step study. Comparing this model to D-P model, the parameter uncertainty model has several different features. First, there are two layers of uncertainties in the model: the intrinsic uncertainty of stochastic process (same with D-P model) and the uncertainty of the drift. It can be shown that, if we set $\gamma = 0$, this model will collapse to the D-P model. In other words, the D-P model is nested in this general model. Second, unobserved Brownian motion can be transferred into observable Brownian motion, and separation theorem allows us to solve this problem in two stages: derivation of conditional expected returns and make decisions based on conditional distributions. Third, analytical solutions are usually impossible in these kinds of cases, while D-P model has its nice analytical solution because all parameters are known.

3.2.2 Continuous Learning with Periodic News (Social Learning) Updates

In developed counties, there are public economic forecasts and newsletters available to people, therefore, agents can make inference (learning) based on realized past return data. However, in rural China the information are more likely from private information sources—early adopters in farmers' social network. Similar to Huang and Liu (2007) model, we allow the agent to obtain direct information about drift (α) periodically from early adopters in his social network, but the information comes with errors because no one really knows the drift even though they have adopted the technology. Different to Huang and Liu (2007), we assume there is no information cost because chatting with friends or neighbors costs little in villages.

Assume farmers can get some information about the drift from his friends with errors:

$$\alpha_j = \alpha + \varepsilon_j \tag{23}$$

where α_j is the person j's opinion about α, and ε_j represents the noise in the information, which is assumed distributed as $\varepsilon_j \sim N(0, \sigma_\varepsilon^2)$.

According to foregoing section, the agent has to solve the problem under new observable Brownian motion Z'_t:

$$\begin{cases} \underset{\tau \in \Theta}{Max}\, E\{\int_0^\tau e^{-\rho T}(V_t - I)dt\} \\ dV_t = V_t[m_t dt + \sigma_v dz'_t] \end{cases} \tag{24}$$

The periodic news would affect his conditional expectation of the drift. Between periodical news updates, the agent infers the conditional distribution of the drift from the observation of realized past project value V_t, the conditional mean and variance follow the equations:

$$\begin{aligned} dm_t &= \frac{\gamma_t}{\sigma_v} dZ'_t \\ d\gamma_t &= -\frac{\gamma_t^2}{\sigma_v^2} dt \end{aligned} \tag{25}$$

Immediately before news α_j is received at time t_j, the conditional distribution of α is normal with mean m_{tj}^- and variance γ_{tj}^-, upon observing α_j, this conditional distribution of α is updated according to the following equations:

$$\begin{cases} m_{tj} = m_{tj}^- + \dfrac{\gamma_{tj}^-}{\gamma^-_{tj} + \sigma_\varepsilon^2}(\alpha_j - m_{tj}^-) \\ \gamma_{tj} = \dfrac{\gamma_{tj}^- \sigma_\varepsilon^2}{\gamma_{tj}^- + \sigma_\varepsilon^2} \end{cases} \tag{26}$$

From above discussion, we can consider periodic news updating as little jumps when they got the direct information from his social networks. This updating is attached to the continuous updating from continuous information sources (realized past returns).

If we don't consider the continuous filtering between two periodical updating, after receiving N times periodic news updating, the conditional expectation and conditional variance of α would be:

$$
\begin{aligned}
m_{tN} &= \frac{\sigma_\varepsilon^2}{N\gamma_{tj}^- + \sigma_\varepsilon^2} m_0 + \frac{N\gamma_{tj}^-}{N\gamma_{tj}^- + \sigma_\varepsilon^2} \overline{\alpha}_j \\
\gamma_{tN} &= \frac{\gamma_0 \sigma_\varepsilon^2}{\gamma_0 N + \sigma_\varepsilon^2}
\end{aligned}
\tag{27}
$$

where m_{tN} and γ_{tN} are the conditional mean and conditional variance of drift after receiving N times periodic news. $\overline{\alpha}_j = \frac{1}{N}\sum_{j=1}^{N}\alpha_j$.

First, look at conditional variance equation, it is easy to see that conditional variance of the parameter decreases in social learning, which means social learning helps reduce uncertainty of the parameter. It makes perfect sense because social learning allows agents get more information about the parameter, which is similar to get more information by observing longer in continuous learning. Therefore, according to the numerical solutions from foregoing part, more social learning means lower trigger value, hence easier to adopt the technology.

Second, look at the conditional mean equation, we can consider $\frac{\sigma_\varepsilon^2}{N\gamma_{tj}^- + \sigma_\varepsilon^2}$ and $\frac{N\gamma_{tj}^-}{N\gamma_{tj}^- + \sigma_\varepsilon^2}$ as two weights for m_0 and $\overline{\alpha}_j$ respectively. With N increases, the m_{tN} will be away from m_0 and converge to $\overline{\alpha}_j$ because $\frac{N\gamma_{tj}^{-}}{N\gamma_{tj}^{-}+\sigma_\varepsilon^2}$ become greater and $\frac{\sigma_\varepsilon^2}{N\gamma_{tj}^{-}+\sigma_\varepsilon^2}$ become smaller. This result indicates that the more information the agent get from social learning, his conditional expectation about the drift will converge to average level of early adopters in his social network. So the periodic news (social learning) helps agents to change his mind from his naïve prior belief to average level of beliefs in his social network. His conditional mean of the parameter goes up or down depends on the relation between m_0 and $\overline{\alpha}_j$. If m_0 is greater than $\overline{\alpha}_j$ (he is too optimistic at the beginning), the social learning would allow him low down his expected return of project, hence means lower trigger value, which facilities his adoption, vice versa.

In conclusion, both continuous filtering and periodical filtering (social learning) affect the agent's perception of the parameter. More information always reduces the uncertainty of the parameter no matter where the information comes from, hence facilitate the adoption. In rural China, since the public information is not usually available to small farmers, information from social learning could play important roles in people's adoption. On the other hand, more information could affect adoption through agents' perception about mean of the drift, and updated perception about mean could make the trigger value higher or lower, which depends on the relationship between an agent's initial belief and average belief in his social network.

4. Empirical Analysis

4.1 Survey and Greenhouse Technology Diffusion

The survey area is Shandong province which accounts for about 7 percent of China's cropping land, but accounted for nearly 12 percent of its horticulture area in 2004. This percentage has been rising over time. Moreover, since the number of greenhouses is higher than average and since the level of commercialization is typically thought to be higher than the rest of China (and so almost certainly yields are higher), it is safe to assume that in fact the share of the Shandong's total production is higher than it's area share.

In Shandong, we conducted two coordinated, community and household level surveys in 2005 and 2006 respectively. The first one, the Shandong village survey, is a provincial representative sample of tomato and cucumber growing villages in China's main horticulture-producing province. The first step in conducting the survey involved creating two sampling frames of county-level tomato production and county-level cucumber production in order to choose the five sample counties per crop. With a knowledge of the total production environment in Shandong for each crop, we ranked counties by the level of output per capita. We then divided the counties in Shandong into 3 groups: high production; medium production and low production counties. 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 decile. The two medium production counties were randomly chosen from the third and fourth quintiles. There was only one low production county chosen. After eliminating the five percent of the counties with the lowest production, the low production county was randomly chosen from the lowest quintile. In the end for each crop there were 2 counties in the high production set of counties; 2 counties in the medium production set of countries and 1 county in the low production set of counties. The total level of production in each set of countries provided data for our weighting system (which is used to create point estimates for provincial averages of each of our variables, see figure 1 and 2).

After the sample counties were chosen, a relatively similar process was used to select sample townships and villages. In total for each crop, the survey teams visited 10 townships. Moreover, for each crop (for the five counties and 10 townships), we interviewed village 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 village, townships and counties in the sample we were able to construct area-based weights in order to be able to create point estimates of our variable that are provincial representative.

After choosing the villages the enumeration team then visited each community and ran data collection activities. In each village, enumerator conducted a two hour, sit-down survey with the 3 village leaders for village survey. In each village, we divided all households into two groups: non-cucumber/tomato households and cucumber/tomato households. We randomly sample 7 cucumber/tomato farmers and 3 non-cucumber/tomato farmers. As a result, we got 350 each tomato households and cucumber households[9]. With a knowledge of distribution of two different households plus distribution of greenhouse households in each village, we could calculate the weights to adjust the selection bias problem.

After data cleaning, we finally got the 638 valid observations, and among them 362 households adopted greenhouse technology. Like other new technologies, farmers

[9] The reason we did not directly randomly stratified sample on greenhouse is the greenhouse survey is a part of big horticulture production survey which require the stratified sample on cucumber/tomato or non-cucumber/tomato households.

adopted the greenhouse at different years which can be described as a technology diffusion process. The greenhouse diffusion process can be roughly divided into three stages: early stage, take-off stage and slow-down stage. From figure 3, we can see that the diffusion process is relatively slow before 1990, only a few of farmers adopted the technology. Between 1990 and 1995, many farmers started to adopt greenhouse, and the diffusion process reaches its peak between 1996 and 2000, then the trend slowed down. This diffusion curve is very similar to the standard HYV maize S diffusion curve (Griliches, 1957).

4.2 Data Description

As we mentioned before, we got total 638 valid household observations in 70 villages. 204 (64%) households out of 317 households in tomato area adopted greenhouse, while 158 (49%) adopters out of 321 households in cucumber area. The tomato growers more likely adopt the greenhouse because a big shading house is competitive substitute for greenhouse in cucumber production.

In this study, we are interested in how social learning affects farmers' adoption. The theoretic framework predicts that social learning helps reduce parameter uncertainty, hence facilitate the adoptions. We are going to test the hypothesis using empirical data, but we need to be aware of that social learning is just one of many reasons to affect adoption in practice. For example, people have other options such as off-farm jobs, or people face other constraints such as credit constraints since greenhouse adoption is a lumpy investment as well. To identify social learning effect and test theoretic hypothesis using empirical data from real life, we need carefully define those theoretic concepts empirically and control other factors properly.

4.2.1 Social Learning

Adoption is a binary and largely irreversible decision, and it is easily observable. However, farmers adopted greenhouse in different years, which implies that we need to put year dummies to control for heterogeneities over years if possible.

Social learning is the key variable for this study, the way we measure social learning is similar to Bandiera and Rasul (2006) work: "how many people do you know adopt the greenhouse before you in your village and nearby villages respectively?", "How many of these people belong to your relatives and friends respectively?"[10] In our questions, we emphasize social learning before a farmer's own adoption, which allows us to identify causality relations instead of just correlations between social learning and adoption in some previous studies. It is reasonable to use the number of adopters in his social network to measure the size of social learning because of several reasons: first, we can not directly measure social learning per se. Second, it relates to the number of different sources of information on greenhouse adoption the farmer has access to from within the set of all people the farmer knows, corresponding the variable in the theoretic model. Third, it related to close contacts from whom information can be more easily obtained related to those outside of this reference group (Bandiera and Rasul, 2006). Four, in survey we found that farmers can remember the number of adopters before him clearly because the greenhouses are big objectives and easily observable, and greenhouse adoption is a big deal to small farmers, they usually pay a lot of attentions to other people's adoption before they adopted.

[10]We didn't ask neighbors because friends usually include neighbors in Chinese.

The first two rows of table 1 provide the means, standard errors by adoption status. In the last column of table 1, the results of test of equality are provided to examine whether the differences between non-adopters and adopters are significant. The first row indicates that adopters know about 6.9 adopters in his village on average, while non-adopters only know about 4.7 earlier adopters. The result of t-test tells us this difference is significant, which implies that the social learning of adopters is significant more than non-adopters. In other words, it implies that non-adoptions could be due to insufficient social learning. When we extend the social network to nearby villages (the second row), the trend is consistent.

4.2.2 Other Characteristics of Household

Table 1 also gives other characteristics of households. The third, fourth and fifth row of table 1 give us information about family size, family labor and off-farm labor for both groups. It is interesting that family size of adopters is significant larger than non-adopters, while the number of family labor of adopters is significant less than non-adopters, this seemingly contradiction means adopters have more dependent family members (either young kids or old parents) than non-adopters, which implies that greenhouse adoption could be a good choice if someone have to stay at home to take care of their dependent family members. The off-farm labor from non-adopters (0.8) is significant more than adopters (0.24), which indicates again that growing greenhouse and off-farm jobs are a substitute for each other. The fact that age of adopters (35) is younger than non-adopters (46.6) is also consistent with the story because people in 30s have young kids and old parents to take care of, and greenhouse growing is a heavy labor duty, which is difficult to old people.

There is no significant difference in education between adopters and non-adopters, which implies that education maybe is not very important when social learning is the main information source about the technology. It is not surprising that off-farm income from non-adopters is much more than adopters provided foregoing discussions. Farm size of adopters is larger than non-adopters, which implies that farm size could be an important factor to greenhouse adoption. Irrigation is important to greenhouse growing, but even 80% of non-adopters have irrigation implies that irrigation might not be an important prerequisite for adoption.

Land tenure security might be another concern to farmers' adoption even though the Chinese government banned major land reallocations in 1984. However, the village leaders often impel different scale land reallocations every several years to make each person in villages has relatively equal share of land. So we collected the times of major and minor land reallocations before adoption to control for land tenure security. The data shows that non-adopter face more land reallocations than adopters in both major (1.44 vs. 0.79) and minor land reallocations (4.29 vs. 3.19) significantly, which implies that land tenure security could be an important factor to greenhouse adoption.

Credit constraint is another big concern to lumpy investment technology adoption. If a farmer has no access to credit and has no enough savings, it would directly result in non-adoption. However, it is difficult to measure whether a farmer is credit constrained, which is equivalent to examine whether a farmer can borrow as much as he would like to borrow at the going market interest rate (Banerjee and Duflo, 2002). Since we focus on initial greenhouse adoption instead of optimal greenhouse scale, we just need to know whether a farmer has the ability to build a greenhouse by borrowing money or their own savings. Therefore, we collected house value of each household to indicate the household's wealth, and his credit history (maximum borrowing and maximum lending) before adoption to indicate his credit ability.

The data shows that non-adopters are significantly wealthier than adopters (mean of house value: 8,773 yuan vs. 4,294 yuan). The credit ability from social network presents the same trend: non-adopters have significant more credit ability than adopters: The maximum lending (862 yuan vs.368 yuan) and Maximum borrowing (1,352 yuan vs.925 yuan). These two evidences imply that credit constraint might not be an important constraint in greenhouse adoption given non-adopters have both more wealth and credit ability than adopters.

4.3 Methodology

Even though we use expected present value and trigger value as tools to analyze farmers' adoption behavior, this does not mean we can observe or ask farmers these two variables directly in the survey. We only can observe farmers' adoption status. To this kind of binary response dependent variable problem, LPM and Probit Model are appropriate to get good estimates, and we choose Probit model to illustrate the connections between theoretic and empirical framework since there are no much difference between Probit and LPM in this part.

When the expected present value (EPV) is greater than trigger value, it would lead to adoption:

$$\text{Y=1 if } Y^* = EPV - V^* > 0$$

$$Y=0 \text{ if } Y^* = EPV - V^* \leq 0 \quad (28)$$

Y=1 indicates adoptions and Y=0 indicates non-adoptions.

EPV is a function of current Profit (π_t) and discount rate given the geometric Brown motion, and assume the farmers are profit Maximizer,

$$\pi_t = f(p_t, w_t) \quad (29)$$

p_t is price of output and w_t are prices of inputs, so

$$EPV_t = f(p_t, w_t, \rho) \quad (30)$$

Also from the theoretic framework, we know if the drift (α) is known, the trigger value is a function of

$$V^* = f(\rho, \alpha, I, \sigma_v) \quad (31)$$

But we don't know the drift, so we have another layer of uncertainty and the function becomes

$$V^* = f(\rho, I, \sigma_v, m_t, \gamma_t, t) \quad (32)$$

Where m_t, γ_t are conditional mean and conditional variance of α, and t is the total length of observation period. Then we substitute m_t, γ_t out, we can get

$$V^* = f(\rho, I, \sigma_v, m_0, \sigma_\varepsilon, \overline{\alpha_j}, \gamma_0, N, t) \quad (33)$$

Combine equation (33) with equation (30), we get

$$Y^* = f(p_t, w_t, \rho, I, \sigma_v, m_0, \sigma_\varepsilon, \overline{\alpha_j}, \gamma_0, N, t) \quad (34)$$

p_t and w_t are supposed to be prices of output and inputs of greenhouse production, however, the history vegetable prices data in Shandong are not available, we have to use ratio of vegetable price index and inputs price index at national level as a proxy to measure the profitability of greenhouse production over years. The discount rate ρ is assumed to be constant to farmers so that its impact goes to the constant terms in regression.

For the investment cost (I), we use greenhouse construction costs (in real value) to represent the investment cost. To non-adopters, we use average costs in the neighborhood as the proxy.

σ_V is the variance of project return , which is assumed to be known in theoretic model. In empirical part, we use output and inputs price variations during 3 years before adoption to be the proxy for the volatility of return. m_0 and γ_0 are initial conditional mean and conditional variance of α, which are determined by personality of a farmer, whether he is optimistic and pessimistic, we use household characteristics such as age, family size and education of household head to capture those factors. $\overline{\alpha_j}$ is the average of expectation of α in a farmer's social network, which can be approximately represented by economic growth rate if we assume the return of project is close to average return of economy, and the average belief of many farmers is close to average return of economy in last 3 year before adoption. σ_ε is unobservable noise about information collection, which goes to error term in regression.

N is the key variable for this study, as we mention before, we use the number of early adopters in a farmer's social network to measure the amount of social learning. t is the total length of observation period, which we calculate the year he adopted minus the year he was aware of the technology to represent the length of observation period.

Beside foregoing factors, we need to be aware of that there are many factors to affect greenhouse adoption in practice such as land tenure security, off-farm jobs and household wealth which we discussed in the foregoing sector. Therefore, we also need control these factors in regression.

In brief, the empirical model is,

$$Y_i^* = f(X_i, Z_i, N_i, D_1, D_2) \qquad (35)$$

X_i are characteristics of the household before adoption (period t-1), which include age, education of household head, family size, farm size, off-farm employments, family labor, irrigation conditions and family wealth etc.

Z_i are institutional variables and other determinants at period t-1, which include land reallocations variables, the ratio of prices index, market volatility, years of awareness of technology (the length of observation period), greenhouse construction costs and conditional mean of market return.

N_i is the social learning variable at period t-1 which include number of adopters in social network.

D_1 are year dummies to control for heterogeneities over different years.

D_2 are county dummies to control for heterogeneities over different counties.

4.4 Estimation

Since the dependent variable is a binary variable and our primary interest is to see how explanatory variables affect the response probability, LPM and Probit are two popular models to get good estimates. However, both of them have their strengths and weaknesses.

4.4.1 Linear Probability Model (LPM)

The LPM model for binary response dependent variable is specified as

$$P(Y=1\mid X)=\beta_0+\beta_1X_1+\beta_2X_2+\cdots+\beta_kX_k \qquad (36)$$

Where Y is adoption status (1 or 0) and X_k are the explanatory variables which are specified according to economic model in foregoing sections.

LPM has its strengths and weaknesses: First, LPM is a linear model which brings us a lot of conveniences in model estimation, OLS can give us consistent and even unbiased estimators, and it is also easy to deal with heteroskedasticity problem by using heteroskedasticity-robust standard errors and t statistics. Second, in this model the β_j now measure the effects of the explanatory variables on a particular probability, in other words, the dependent variable is a conditional probability. Therefore, unless the range of explanatory variables is severely restricted, the LPM can not be a good description of the population response probability. The model should be seen as a convenient approximation to the underlying response probability. What we hope is that the linear probability approximates the response probability for common values of the covariates. Fortunately this often turns out to be the case (Wooldridge, 2002). Third, since OLS does not require the correctly specified conditional density functions, it is more robust than Probit with MLE. Therefore, even with some weaknesses, LPM often provides good estimates of the partial effects on the response probability near the center of the distribution of X.

4.4.2 Probit Model

Probit model is a non-linear model, and it is can be derived from underlying latent variable model:

$$Y^*=\beta_0+\beta_1X_1+\beta_2X_2+\cdots+\beta_kX_k+e \qquad (37)$$

If we assume the error term (e) follows the standard normal distribution, we can write the equation (37) as:

$$P(Y=1 \mid X) = P(Y^{*} > 0 \mid X) = G(X\beta) \qquad (38)$$

where the G is cumulative distribution function (CDF) of standard normal distribution. To this kind of non-linear function, it can be proved that maximum likelihood estimation can get consistent and efficient estimators if we specify the conditional density function correctly. However, it is almost impossible to test the distribution of error terms in empirical study. In addition, heterogeneity in Probit models could bring us troubles to estimate parameters consistently even heterogeneity is independent of X. However, if we want to estimate partial effects, ignoring heterogeneity is not a problem because we still can consistently estimate the average partial effects given heterogeneity is independent of X.

4.4.3 Endogeneity

The endogeneity is one of the most formidable problems in empirical studies. Before we discuss how to deal with endogenous problem using econometric tools, let us firstly discuss why we could face endogenous problem when we try to identify the social learning effect.

A. Reflection Problem and Endogeneity

A person usually has similar personalities with people in his social network (an old Chinese saying). That is one of reasons why we observe that persons who belong to the same group tend to behave similarly. Manski (1993) define this phenomenon as "reflection problem" which is driven by two effects: (1) endogenous effects, wherein the propensity of an individual to behave in some way varies with the prevalence of the behavior in the group. (2)Correlated effects, wherein individuals in the same group tend to behave similarly because they face similar environments and have similar personal characteristics.

In this paper, we try to prove that a farmer adopt greenhouse because he learn something from early adopters in his social network (social learning effects), instead of endogenous effects or correlated effects. Therefore, we need identify the social learning effects from endogenous effect or correlated effects.

Endogenous effect is essentially the social pressure problem, and psychologists use social pressure to interpret why people tend to behave the same way with most people in the group. In greenhouse adoption case, in most villages, greenhouse adopters are minority group, therefore it would be rare that farmer adopt the greenhouse because of the social pressure.

However, the correlated effect is main identification obstacle to us. For example, if farmers within a social network in which all members are curious to new things, a farmer's adoption could just come from his curiosity to new things, instead of social learning from others in his social network. Since there could be many such kinds of unobservable factors in error terms, endogenous problems could bias our estimates.

Endogenous problem mainly comes from two sources: simultaneous determinations and unobservable heterogeneity. Since we focus on initial adoption decisions, and we collected the information at the year before the adoption[11] so that we are able to avoid endogeneity from simultaneous determinations.

B. Endogeneity Test

[11] To non-adopters, we collected the information in the year before the survey occurred. (2005)

To test endogeneity in linear model, Hausman (1978) suggested comparing OLS and 2SLS estimators as a formal test of endogeneity. For sake of computational brevity, there is a regression-based two steps method:

(1) Run the OLS regression N_i on X, Z, Z_{iv}, D_1 and D_2, and save the residuals $\hat{v}_2$;

(2) Run the OLS on X, Z, N_i , D_1 ,D_2 and $\hat{v}_2$ to get t statistics to test endogeneity.

Where Z_{iv} is the extra instruments to identify the model. A nice feature of this procedure is that the usual t statistic on $\hat{v}_2$ is a valid test of the null hypothesis that N_i is exogenous. If we can not reject the null hypothesis, we have more confidence to believe social learning is exogenous, and OLS would give us consistent estimators. The key here is instrument variables Z_{iv}, here we use a farmer's total kin adopters and the number of household in the village as instruments for identification. We can readily prove that these two variables are correlated with social learning, and we believe the number of household in the village is pretty exogenous to error terms given the birth control policy in China is strict and unified across country[12]. Therefore, can we test the other instrument (total kin adopters) is valid?

C. Over-identification Test—Hansen-Sargan Test

If we have more instruments than we needed to identify a model, we could test whether the additional instrument is valid in the sense that they are uncorrelated with error terms. The idea is to compare 2SLS estimator using all instruments to 2SLS sung a subset of instruments. In this study, we use Hansen-Sargan test to examine the validity of the instrument in Stata (Ivreg2).

[12] Migrants are still included in the village population even though they reside in cities in most time.

The Hansen-Sargan test for overidentification evaluates the entire set of overidentification restrictions. For example, if we run the test like this:

$$Ivreg2 \cdots Y \cdots N_i (= Z_{iv1}, Z_{iv2}) \cdots X \cdots Z \cdots D_1 \cdots D_2 \cdots orthog(Z_{iv1}) \qquad (39)$$

where Z_{iv1}, Z_{iv2} are excluded instruments. There are three statistic are reported in the test. The first one is called "Lagrange multiplier test of excluded instruments", which reports Hansen-J statistic for restricted equation (using entire set of instruments including Z_{iv1}, Z_{iv2}); the second one is called "Hansen J statistic for unrestricted equation", which report Hansen-J test for unrestricted equation (using subset of instruments, not including Z_{iv1}, Z_{iv2}), the last one is C statistic for specified instruments: in this case, it is Z_{iv1}. Here we focus on the last one: C statistic. If we can not reject null hypothesis, the C statistic tells us we have more confidence to say Z_{iv1} is exogenous given we believe Z_{iv2} are really exogenous. Otherwise, we say Z_{iv1} is endogenous given Z_{iv2} are exogenous.

It is easy to see that the power of C statistic depends on the exogeneity of other instruments, but it provide us a better way to test validity of instruments: if we have multiple instruments, we have some confidence in exogeneity of subset of them, and we can use the subset of instrument to test the validity of others. That is the best thing we can do about testing the validity of instruments.

5. Results

We have two models to examine the effect of social learning on greenhouse adoption. From the theoretic model and its numerical solutions, we expect social learning have positive impact on adoption because it reduce the parameter uncertainty. The first layer of uncertainty from stochastic project value (σ_v) would discourage the adoption, and we also expect the conditional expectation of mean of project return (m) discouraged the adoption as well because waiting value increase in higher m.

5.1 Linear Probability Model Results

Table 2 provides the estimation results of LPM model by OLS with cluster robust standard errors. The first two columns use the measure of social learning within a village, and the last two column reports the results using the measure of social learning both in the village and nearby villages. Generally speaking, the results are very similar between the two measures, which indicates the village border is not important to social learning in term of its effects on greenhouse adoption.

Let us use first two columns to illustrate the results in detail. The first row of table 2 is the key result for this study: social learning has significant impacts on greenhouse adoption. It means one more adopters in a farmer's social network, the probability of his adoption increase by 0.53% after we control other factors. In other words, if there are 20 early adopters in his social network, the probability to adopt greenhouse increase about 10%. Given the greenhouse adoption rate is still low in rural China, this amount of increasing probability is significant.

The second row of table 2 is about how conditional expectation of mean of return (m) affects adoption. From theoretic model, we know the farmers' conditional expectation of mean of return will converge to the average level of his social network. However, in practice, we can not observe people's expectation. In the empirical study, we use economic growth rate in last 3 year before adoption to approximate the expected return of project. Amazingly, the sign of estimate is consistent with the theoretic model, and higher expected return means less adoption because higher expected return induces higher waiting value.

We use market price volatility of vegetables in last 3 year before adoption to represent market volatility in the empirical study. The result indicates the first layer of uncertainty from stochastic project value (σ_v) discourage the adoption (row 3), which makes perfect sense because market volatility also increases the option value of waiting.

We use year of awareness of the technology to represent the continuous learning effects, however, it is not significant. The reason could be: farmers don't have a lot of chance to learn market information in rural China if they have not adopted the greenhouse because their main information resource is social learning from his social network. We use the ratio of output price index to input price index to represent the profitability of the technology, the results makes perfect sense: more profitability of the technology induces more adoption.

In term of household characteristics, age has negative impacts on adoption because growing greenhouse is a heavy labor duty. But the coefficient is small (0.3%), which means age is not a big problem to most of farmers. It is easy to understand more land, the probability of adoption increase a little bit. The impacts of other characteristics of household are not statistically significant, however, the sign of coefficients are consistent with theory in most cases. Some of them almost reach to the significant level, for example, the family size, the irrigation ratio and greenhouse costs, which indicate these factors could play some roles in adoption practices.

It is worthy to mention that the R-square of this regression is pretty high (0.90). It implies that the model includes most of factors which could affect adoption, and it also implies that the irreversible investment model is appropriate to interpret the greenhouse adoption behavior in reality.

We use Hausman test (regression-based two steps method) to examine the exogeneity of social learning, and P-value (0.26 and 0.71) indicate that we can not reject null hypothesis, in other words, we have more confidence to believe that social learning is indeed exogenous. The instruments we use to identify the equation are: total kin adopters and the number of household in a village. As we discuss in foregoing section, the number of household in a village is pretty exogenous to error terms because the strict and unified birth control policy in rural China. But we need test the validity of the other one: total kin adopters. The Hansen-J test provide a good way to test the validity of an instrument given we believe the other one is exogenous. We report all the Hansen-J test results to provide the whole pictures of the test. But the key statistics to us is the C statistic (last row of table 2), which indicate the other instrument (total kin adopters) is valid to be exogenous variable because the P-value indicate we can not reject the null hypothesis. We also could use Hansen-J test to examine the exogeneity of social learning given the two excluded instruments are valid, the result also proves that the social learning is exogenous.

5.2 Probit Model Results

Table 3 provides the results for Probit model estimated by MLE. We expect that Probit model provides us more efficient estimators than LPM provided normal distribution is correctly specified for the empirical model.

As we discussed before, there is no much difference between two different measures of social learning, and several points worthy to be mentioned here: First, the results for key variables (social learning, market volatility etc) are close to LPM, which verifies robustness of the results. Second, most coefficients in Probit are reasonable greater than coefficients in LPM, which is even better to us because for example, social learning has more significant impact (0.75%) on adoption. Third, conditional expectation of mean of project return becomes insignificant statistically even though the coefficient is greater and sign is still consistent with the theory.

Most household characteristics become significant statistically, which could be due to the efficiency gain from MLE (smaller asymptotic variances). In the Probit model, the off-farm income, wealth of household (we use house value to approximate the wealth of household) and major land reallocations have significant negative impacts on greenhouse adoption. It is easy to understand impact of these factors, for example, if farm can earn a lot of money from off-farm jobs, it is reasonable not to adopt greenhouse given they can not do these two things simultaneously. The same logic explains why wealthier farmers don't want to adopt greenhouse. The major land reallocation is interesting, and it means land tenure security could be an important factor in greenhouse adoption.

Another thing is worthy to be mentioned here: we can not use year dummies in the Probit model because the year dummies would perfectly predict the results due to the data structure. To avoid inducing endogenous problems by this, we omit the years of awareness of technology as well because it is closely related with year dummies.

As we discussed before, Probit model could be more efficient than LPM, however, it comes at expense of less robustness because nobody can test whether the error term is normal distribution. On the contrary, OLS can provide us consistent estimators after we take care of endogenous problems. The conditions for OLS consistency is much weaker than conditions for Porbit model. Therefore, LPM is the safer bet if we care for robustness of the results.

6. Panel Data Approach

As we mention before, farmers adopted greenhouse in different year, which means it is very difficult to us to observe enough new adoption in any particular year if we do randomly sampling. That is the reason our data is pooled data and we use time dummies to control heterogeneities over time. After we carefully deal with endogeneity and test instrument variables using Hausman and Hansen-Sargan tests in previous models, we are pretty confidence in robustness of our results. However, we could still get spurious results if some small probability things happen. For example, the number of household in a village is correlated with error terms for some strange reasons even though it is hard to image. We can not solve this problem by finding better instrument variables because there are no perfect instruments in most cases. So the questions is “can we prove our results from a new angle?”.

The panel data approach is the attempt from the new angle. The idea is as follows: instead of examining how social learning affects the adoption itself, we examine how social learning affects greenhouse area or the number of greenhouses owned by households, and we also use number of adopters in a production team to approximate the amount of social learning. Production teams are smaller units compared to villages, a village could include several to dozens of production teams which were basic production unit in the People’s Commune period. Even people are not doing farming together nowadays, but farmers in the same production team know each other well, that is the rational we could use adopters in the same production team to approximate the amount of social learning. By these two new measures, panel data is available from randomly sampling and it provides us a new angel to examine how social learning affect greenhouse adoption. In addition, in term of econometric perspective, Panel data is a better way to deal with heterogeneity problem because fixed effect method can eliminate the unobserved heterogeneity pretty well.

We have panel data in 2001 and 2006, and fixed effect method and first difference produce identical estimates and inference for two time periods panel data. We adopt the first differencing in our study because it is easy to implement for two time periods data.

Table 4 and table 5 provide the results using first differencing data estimated by OLS. The table 4 provides the results using current number of adopters in production team and table 5 uses lagged (in 1996) number of adopters in production team because of the concern of simultaneous endogeneity.

Let us discuss the table 4 firstly. The first two columns are results using greenhouse area as the dependent variable, and the next two columns use number of greenhouse as the dependent variable. In both cases, social learning has significant impact on greenhouse production. The coefficients are similar in both cases (0.4% and 0.3%) because the usual greenhouse size is one mu (60 meters long, ten meters wide). It is interesting to note that the coefficient from the new angel (0.4%) is close to coefficients in LPM or Probit (0.5% or 0.75%) because of the same reason. It is another evidence that the result of social learning impact are robustness. We also did Hausman test for endogeneity in panel data case, the P-Values indicate that social learning variables are exogenous[13].

Table 5 uses lagged number of adopters in a production team (in 1996) as the measure of social learning, and social learning becomes statistically more significant, but the coefficient become smaller, which implies that more recent social learning have greater impacts.

7. Conclusion

This paper introduces social learning into irreversible investment theory through parameter uncertainty, and shows that social learning could reduce parameter uncertainty and facilitate technology adoption with irreversible investment. We also use household level data from energy saving greenhouse adoption in northern China to test the theory, here are our findings:

First, social learning has significant positive impacts on technology adoption through reducing parameter uncertainty. The parameter uncertainty is the second layer of uncertainty, which increases total uncertainty when some parameters are unknown to agents. Second, empirical data also verifies the conventional theory about irreversible investment: the higher uncertainty from stochastic process (first layer of uncertainty), the higher trigger value, hence less adoption. Third, social learning also could affect technology adoption through updated belief about mean of returns, and it depends on agents' initial belief and average level of belief in his social network. The empirical data

[13] Panel data approach has less explanatory variables because some variable are dropped out because of no variations over time such as age, some variables are dropped out because of no variations across people.

shows some evidences about this effect, but the evidence is not so strong compared to uncertainty part.

This paper also provides an answer to this research question: how could small farmers in developing countries deal with risk from irreversible investment and incomplete information when the product-backup system is not available? The results from this study indicate social learning is an effective solution. Therefore, the policy implications from this paper are clear: when small farmers face technology adoption such as tube well or machinery, helping several successful adoptions in their villages maybe be the best way to induce more adoption. In addition, stable market environment is also helpful to this kind of technology adoption.

Figure 1: The Distribution of Tomato in Shandong, 2005[14]

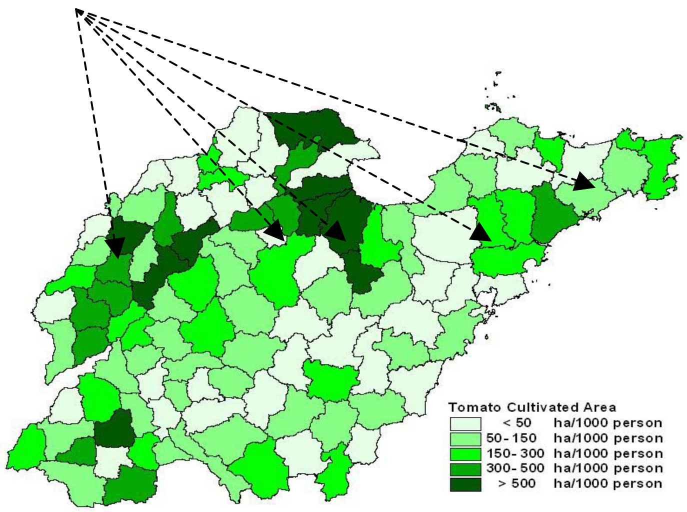
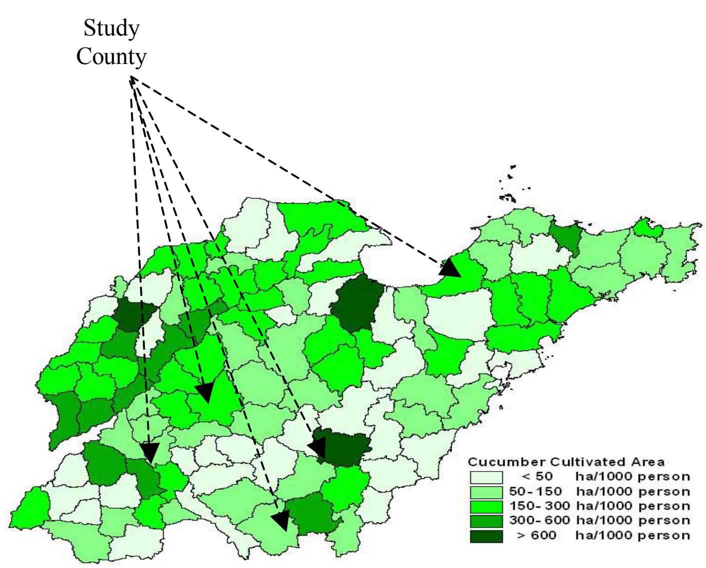


Figure 2: The Distribution of Cucumber in Shandong , 2005



[14] Data source: Provided to the authors by the Shandong Agricultural Bureau, Jinan.

Figure 3: Greenhouse Diffusion Curve at Household Level

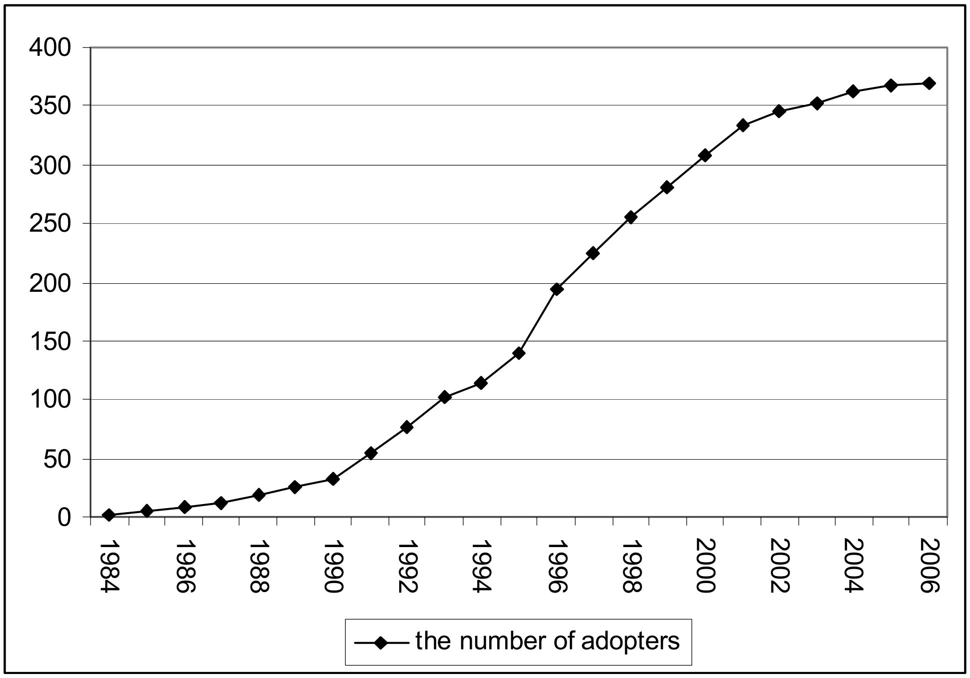


Table 1: Descriptive statistics: Household Level Data

Basic characteristics (all variables are measured in the year before adoption)	**Non-Adopter (mean)**	**Adopter (mean)**	**Test of Equality Ho: diff = 0 (P-value)**
Number of adopters in your village	4.7 (0.7)	6.9 (0.67)	0.027**
Number of adopters in your 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)	0.00***
Off-farm employment	0.8 (0.054)	0.24 (0.022)	0.00***
Age of family head	46.4 (0.6)	35 (0.46)	0.00***
Education of family head	7.0 (0.17)	7.24 (0.14)	0.25
Off-farm income (yuan)	8420 (649)	1643 (182)	0.00***
Farm size (mu)	5.6 (0.19)	6.01 (0.16)	0.09*
Irrigation ratio	0.80 (0.019)	0.89 (0.013)	0.00***
Major land reallocations since 1980	1.44 (0.067)	0.79 (0.05)	0.00***
Minor land reallocations since 1980	4.29 (0.26)	3.19 (0.19)	0.00***
House value (yuan)	8773 (539)	4294 (413)	0.00***
Maximum lend	862 (104)	368 (66)	0.00***
Maximum borrow	1352 (146)	925 (102)	0.01**

*diff= mean(0) - mean(1), and we use ttest to performs t tests on the equality of group means. The numbers in brackets are standard errors

Table 2: Greenhouse Adoption and Social Learning –LPM Model
Dependent variable: 1=adopt, 0=not adopt

Explanatory variables	Coefficient	Robust std error	Coefficient	Robust std error
Social Learning				
Social learning within village	0.0053	0.0018***		
Social learning within village and nearby villages			0.0054	0.0019***
Conditional mean of market return	-0.024	0.0075***	-0.023	0.007***
Market volatility	-0.0015	0.0007**	-0.0015	0.0006***
Years of awareness of the technology	-0.0004	0.002	-0.0011	0.0023
Output price/input price	1.03	0.29***	1.03	0.28***
Household Characteristics				
Family size	0.015	0.013	0.016	0.012
Age of family head	-0.0037	0.0016**	-0.0038	0.0016**
Education of family head	0.0024	0.0033	0.0019	0.0032
Off-farm labor	-0.002	0.024	-0.0086	0.023
Off-farm income	-0.0058	0.007	-0.0042	0.007
Farm size	0.006	0.003*	0.0063	0.0034**
Irrigation ratio	0.052	0.034	0.053	0.036
House value	-0.0023	0.002	-0.0024	0.0021
Greenhouse construction cost	-0.01	0.009	-0.0093	0.0091
Times of major reallocations	0.004	0.014	0.0011	0.014
Times of minor reallocations	-0.008	0.0049	-0.0073	0.0049
Dummies and constant terms				
Crop dummy	-0.026	0.038	-0.02	0.038
County dummies	Yes		Yes	
Year dummies	Yes		Yes	
Constant terms	-0.45	0.26**	-0.46	0.25**
Observations	625		625	
Adjusted R-squared	0.90		0.91	
Hausman test for endogeneity	**P-value**	**0.26**	**P-value**	**0.71**
Hansen J tests for over-identification	**Statistics**	**P-value**	**Statistics**	**P-value**
Lagrange multiplier test of excluded instruments	1.9	0.38	0.34	0.84
Hansen J statistic for unrestricted equation	0.27	0.60	0.22	0.63
C statistics for specified IV	1.6	0.20	0.18	0.73

Excluded instruments: total kin adopters and village household number which are used as IVs in Hausman test for endogeneity, and specified IV: total kin adopters . *** denotes significance at 1%, ** 5% and * 10%.

Table 3: Greenhouse Adoption and Social Learning –Probit Model
Dependent variable: 1=adopt, 0=not adopt

Explanatory variables	dF/dx	Robust std error	dF/dx	Robust std error
Social Learning				
Social learning within village	0.0075	0.003***		
Social learning within village and nearby villages			0.0077	0.0028***
Conditional mean of market return	-0.061	0.055	-0.057	0.054
Market volatility	0.005	0.002***	-0.005	0.0018***
Output price/input price	2.39	1.05***	2.32	1.03***
Household Characteristics				
Family size	0.073	0.039**	0.071	0.038**
Age of family head	-0.03	0.01***	-0.03	0.01***
Education of family head	-0.02	0.014	-0.02	0.013
Off-farm labor	0.21	0.08***	0.02	0.081***
Off-farm income	-0.19	0.057***	-0.18	0.05***
Farm size	0.037	0.018***	0.038	0.017***
Irrigation ratio	-0.035	0.10	-0.043	0.10
House value	-0.023	0.008***	-0.023	0.008***
Greenhouse construction cost	-0.088	0.033***	-0.084	0.032***
Times of major reallocations	-0.11	0.05***	-0.11	0.05***
Times of minor reallocations	-0.007	0.015	-0.005	0.015
Dummies and constant terms				
Crop dummy	-0.99	0.011***	-0.99	0.01***
County dummies	Yes		Yes	
Observations	625		625	
Pseudo R-squared	0.78		0.79	

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

Table 4: Robustness Check –Panel Data Approach

	Dependent variable: Area of greenhouse		Dependent variable: Number of greenhouse	
Explanatory variables (first difference)	**Coefficient**	**Robust std error**	**Coefficient**	**Robust std error**
Number of adopters in production team	0.004	0.002*	0.003	0.0016*
Family size	0.028	0.038	0.016	0.037
Sum of education in family	0.004	0.004	0.0049	0.0052
Total labor in the family	0.05	0.058	-0.035	0.049
Total Ag labor	0.0006	0.0009	0.0003	0.0007
Non-farm labor	-0.07	0.058	-0.028	0.043
Non-farm income	0.0025	0.062	0.026	0.06
Farm size	0.013	0.017	0.002	0.01
Irrigation area	0.022	0.026	0.015	0.015
House value	-0.038	0.034	-0.037	0.025
Construction cost of greenhouse	0.167	0.11	0.12	0.16
Constant terms	0.14	0.052***	0.14	0.63
Observations	632		632	
Adjusted R-squared	0.11		0.11	
P-value for Hausman endogeneity test	0.92		0.54	

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

Table 5: Robustness Check –Panel Data Approach

	Dependent variable: Area of greenhouse		Dependent variable: Number of greenhouse	
Explanatory variables (first difference)	**Coefficient**	**Robust std error**	**Coefficient**	**Robust std error**
Lagged number of adopters in production team	0.0016	0.0007**	0.002	0.0005***
Family size	0.02	0.04	0.009	0.038
Sum of education in family	0.0057	0.005	0.006	0.005
Total labor in the family	0.06	0.06	-0.022	0.048
Total Ag labor	0.0005	0.001	0.001	0.0007
Non-farm labor	-0.09	0.06	-0.039	0.042
Non-farm income	-0.013	0.061	0.011	0.054
Farm size	0.024	0.02	0.015	0.015
Irrigation area	0.009	0.03	0.0025	0.02
House value	-0.03	0.034	-0.036	0.024
Construction cost of greenhouse	0.19	0.11	0.15	0.16
Constant terms	0.13	0.05	0.13	0.062**
Observations	632		632	
Adjusted R-squared	0.11		0.09	
P-value for Hausman endogeneity test	0.98		0.75	

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

References

1. Abasov, T.M. (2005):"Dynamic Learning Effect in Corporate Finance and Risk Management", Unpublished Ph.D Dissertation, University of California, Irvine.
2. 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.
3. Bradiera,O. and Rasul, I. (2005): "Social Network and Technology Adoption in Northern Mozambique", Manuscript, London School of Economics.
4. Bardhan, P. and Udry, C. (1999): "*Development Microeconomics*", Oxford: Oxford University Press.
5. Bertola, G. and Caballero, R. (1994): "Irreversibility and Aggregate Investment," *Review of Economic Studies*, 61, 223-246.
6. Besley, T. and Case, A. (1997): "Diffusion as a Learning Process: Evidence from HYV Cotton", mimeo, *Princeton University*.
7. Brennan, M.J. (1998): "The Role of Learning in Dynamic Portfolio Decisions", *European Economic Review* 1, 295–306.
8. Caplin, A. and Leahy, J. (1993): "Sectoral Shocks, Learning and Aggregate Fluctuations", *Quarterly Journal of Economics*. 60(4): 779-794.
9. Chavas, J. P. (1993): "On Sustainability and the Economics of Survival," *American Journal of Agricultural Economics*, 75(1), 72–83.
10. Chinese Agriculture Yearbook. (2006). Chinese Agriculture Press.
11. Conley, T. and Udry, C. (2003): "Learning About a New Technology: Pineapple in Ghana," mimeo, *Yale University*.
12. Demers, M. (1991): "Investment under Uncertainty, Irreversibility and the Arrival of Information over Time," *Review of Economic Studies,* 58:333-350.
13. Dixit, A.K and Pindyck, R.S. (1994): "Investment under Uncertainty," *Princeton University Press*.
14. Fafchamps, M. and Lund, S. (2000): "Risk Sharing Networks in Rural Philippines," mimeo, *University of Oxford*.

15. Feder, G., Just, R.E., and Zilberman, D. (1985): "Adoption of Agricultural Innovations in Developing Countries: a Survey," *Economic Development and Cultural Change*, 33, 255–298.
16. 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.
17. Gennotte, G. (1986): "Optimal Portfolio Choice under Incomplete Information", *Journal of Finance,* 41, 733–746.
18. Goorbergh, R.W.J, Huisman, K.J.M and Kort.P.M. (2003): "Risk Aversion, Price Uncertainty, and Irreversible Investments," *CentER Working paper*, 2003-119.
19. Griliches, Z. (1957): " "Hybrid Corn: An Exploration in the Economics of Technological Change," *Econometrica,* 25, 501–522.
20. Hasset, K. A., and Gilbert E. M, (1992): "Energy Tax Credits and Residential Conservation Investment," *NBER Working Paper* No. W4020.
21. Huang, L. and Liu, H. (2007):"Rational Inattention and Portfolio Selection", *The Journal of Finance*, 62, 1999-2040.
22. Jovanovic, B. and Nyarko, Y. (1994): "The Bayesian Foundations of Learning by Doing", *NBER Working Paper* No.4739.
23. Jackson, M.O. and Watts, A. (2002a): "The Evolution of Social and Economic Networks," *Journal of Economic Theory*, 106, 265-295.
24. Liptser, R. and Shiryaev, A. (2001): Statistics of Random Processes, Springer-Verlag, Berlin.
25. Manski, C.F (1993): "Identification of Social Effects: Reflection Problem," *Review of Economic Studies,* 60, 531-542.
26. McDonald, R. and Daniel, S. (1984): "Option Price When the Underlying Asset Earns a Below-Equilibrium Rate of Return: A Note," *Journal of Finance*, March, 261-265.
27. Merton, R.C. (1980): "On Estimation the Expected Return on the Market", *Journal of Financial Economics,* 8, 323-361.

28. Munshi, K. (2004): “Social Learning in a Heterogeneous Population: Social Learning in the Indian Green Revolution,” *Journal of Development Economics*, 73, 185–213.
29. 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.
30. Olmstead, A.L. and Rhode, P (1993), “Induced Innovation in American Agriculture: a Reconsideration,” *Journal of Political Economy,* 101(1), 100–118.
31. Reardon, T. (1997): “Using Evidence of Household Income Diversification to Inform Study of the Rural Nonfarm Labor Market in Africa,” *World Development* 25, 735-748.
32. Sunding, D and Zilberman, D. (2000): “Research and Technology Adoption in a Changing Agricultural Sector,” *Draft for Handbook of Agricultural Economics.*
33. Schultz, W, (1961): “Investment in Human Capital,” *American Economics Review* 51, 1-17.
34. Summer, L.H. (1987):“Investment Incentives and the Discounting of Depreciation Allowance,” In *The effects of Taxation on Capital Accumulations*, Chicago University Press.
35. Twonsend, R. (1994): “Risk and Insurance in Village India,” *Econometrica,* 62, 539–91.
36. Wan, X. (2000): “The Chinese protection agriculture outlook and trend,” *Agricultural Machinery (in Chinese),* 2000, 4-6.
37. Wooldridge, J. (2002): “Econometric Analysis of Cross Section and Panel Data,” Massachusetts, The MIT Express.
38. Xia, Y. (2001): “Learning about Predictability: The Effects of Parameter Uncertainty on Dynamic Asset Allocation”, *Journal of Finance,* 56, 205-246.
39. Yamauchi, F. (2007): “Social Learning, Neighborhood Effects, and Investment in Human Capital: Evidence from Green-Revolution India,” *Journal of Development Economics*, 83, 37-62.

40. 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.
41. Zeira, J. (1987): "Investment as a Process of Search," *Journal of Political Economy,* 95(1) 204-210.