Feedback Relationships between New Technology Use and Information Networks: Evidence from Ghana

Emma C. Stephens
Pitzer College*

June 2009

Abstract

In this paper I examine the relationship between social network connections and profitability of a newly introduced agricultural commodity in Ghana. Using vector autoregression techniques, I illustrate the bidirectional causality between the dynamics of profits earned from a new agricultural crop and information obtained from one's social network. In particular, the data suggest both that greater information from the social network drives farmer profits in the new commodity and also that profitability drives information network development. This implies that less well-connected farmers have a smaller chance of success with the new crop and their information networks are subsequently impacted by these low profits. Spreading the benefits of new agricultural technologies may thus involve significant investments in information supply at the community level.
JEL Codes: O33, O12, O13, Q12, D83

*I would like to thank Chris Barrett, Harry Kaiser, Stefan Klonner, Jacqueline Vanderpuye-Orgle and participants at Pitzer's faculty seminar for excellent feedback and comments and Chris Udry for making the data available. All remaining errors are entirely my own responsibility.

1 Introduction

The introduction of new technologies into farming communities in developing countries has the potential to greatly reduce the incidence of poverty and improve the well-being of these populations. But farmers interested in using a new technology may not possess all of the relevant information about it (Bardhan and Udry, 1999; Evenson and Westphal, 1995; Feder, Just, and Zilberman, 1985). Eventual success with the new technology has been shown to partly depend on the availability of information about its specific characteristics (Bandiera and Rasul, 2006; Foster and Rosenzweig, 1995).

For farmers in developing countries, access to such information can be quite limited. In the case of a new crop that a farmer wishes to try, government supported extension services may not be available in all areas, and there may be little scope for contract farming arrangements, which are often implemented with strict guidelines on how to grow a new crop. In such cases, an important source of information is the farmer's own social network of family, friends and neighbors. Social networks have been shown in several studies to influence both the adoption and management of new crops (Bandiera and Rasul, 2006; Conley and Udry, 2009).

Given that farmers rely heavily on information from their social networks when learning about and using new technologies on their farms, differences in information availability can have a critical impact on farmer revenues. For example, Foster and Rosenzweig (1995) develop a target input learning model that relates farm profits to the number of 'experimental trials'of a new technology, either by the farmer on their own ('learning by doing') or from information gained through experiments conducted by information neighbors ('learning from others'). They show that both own and other's information experiments positively influenced the profitability of high-yield variety (HYV) seeds in India during the Green Revolution, as well as the likelihood of adoption.

In many of the existing studies of technology adoption and social networks and learning, analysts often have to assume that information availability is uniform throughout a farmer' s village, primarily because data on more detailed social network connections is not available (Foster and Rosenzweig, 1995; Munshi, 2004). However, Conley and Udry (2001) have demonstrated that not everyone in a village knows each other and that there is a lot of variability in access to information at the village level. Therefore, it is very likely that within a village, the technology adoption, diffusion and social learning processes will vary across households partly as a result of variable information availability. Using village level aggregates also does not provide information on how the benefits of new technologies are distributed among individual households, which may be of particular concern if the purpose of the new technology is poverty reduction.

Even when detailed social network data are available, estimating the effects of information obtained from a farmer' s social network on new technology adoption and profitability is complicated by the fact that social networks are clearly farmer choice variables.[1] Recent work shows that income and wealth status influence the likelihood of a social connection (Santos and Barrett, 2005).[2] Thus, information is very likely endogenous in the relationship between profits from a new technology and information about it from one' s social network. It may be difficult to determine whether farmers are successful because of access to information, or if they have access to information because they are successful. It may also be hard to find instruments for information from a social network structure that are not related to the profit outcome variable.

The bidirectional causality between social network formation and the profitability of new technology suggested by the existing research cited above, in addition to making estimation of the one-way relationships between these two variables difficult, may also be of interest in its own right. Given that information flows affect farmer profits in new crops, and farmer incomes in part determine social network connections, it is possible that farmers with strong social networks that undertake a new technology may have a better chance of success with the new technology which will enable them to maintain and even expand their own social network, while farmers who have few contacts may be less likely to profit from their experiments, further limiting their chances for obtaining critical information for future success. Social invisibility has been shown to be responsible for other aspects of endemic poverty, such as incomplete risk sharing (Vanderpuye-Orgle and Barrett, 2007), further hampering the ability of less well-connected farmers to lift themselves out of poverty. Similar feedback relationships have been suggested to explain persistent wage inequality between different sectors of the labor market (Calvó-Armengol and

[1] Social networks are often identified by responses to survey questions put to farmers about who they turn to for assistance or advice about farming (Bandiera and Rasul, 2006; Conley and Udry, 2009; de Weerdt and Dercon, 2006).

[2] Fafchamps and Gubert (2007) analyze the determinants of the formation of networks to share income risk in the Phillipines.

Jackson, 2007).

In terms of the prospects for poverty reduction via the introduction of new crops and other technologies, such social invisibility may thus lead to difficulty in adopting and using new technologies which may lead to even more limited social visibility. This possibility implies a different dynamic in technology diffusion than the one of 'strategic delay'suggested by Foster and Rosenzweig (1995) and Bandiera and Rasul (2006). Socially invisible farmers likely do not receive information spillovers from early adopters and thus may never adopt. And, if they do, they may not be able to achieve the same level of success as farmers with more extensive networks, with consequences for future social network development. This characteristic of technology adoption and diffusion has not been explored in depth, but may be important in explaining different rates of diffusion of new technologies and subsequent welfare outcomes. It has been noted by many authors that new technology diffusion is typically incomplete in developing countries (Feder, Just, and Zilberman, 1985). The analysis in this paper may provide insight into another reason this might be the case.

With time series data on technology use and social network formation, it is possible to explore these feedback relationships, as well as account for the endogeneity of social network formation, with the use of vector autoregression techniques. However, due to an absence of data on both of these variables for the same set of individuals, this has not yet been attempted in the technology adoption literature. But, with some recently collected, very detailed data on profits and social network formation, I am able to overcome this difficulty.

In particular, I apply the vector autoregression for panel data methods of Holtz-Eakin, Newey, and Rosen (1988) to data on profits from pineapple farming and changing information about pineapple crop management (based on a measure of experience with the new crop) from the social networks of pineapple farmers in three villages in Ghana over six distinct time periods between January and August 1998.[3] The methods of Holtz-Eakin, Newey, and Rosen (1988, 1989) are well-suited to systems of bidirectional causality, and have been used previously to look at the relationship between labor hours and wages (Holtz-Eakin, Newey, and Rosen, 1988) and government expenditures and revenues (Holtz-Eakin, Newey, and Rosen, 1989). The results for the Ghanaian pineapple farmers clearly show the bidirectional relation-

[3]One time period equals to six weeks. A full description of the data can be found in Goldstein and Udry (1999).

ship between profitability with a new technology and information network dynamics, with information levels from the farmer' s social network affecting profitability and profitability affecting the level of information available to the farmer. Granger causality tests show further that information flows can be said to Granger cause profitability over time, in line with other work on social learning and new technology use (Foster and Rosenzweig, 1995) and also that profitability drives information flows in a similar manner. Information from one' s social network changes both through the total experience of farmers in the network as well as through the changing composition of the network itself, brought on by farmers making different choices on information network partners over time. The results in this paper indicate that new technologies can affect information network dynamics, perhaps through both of these channels. This will in turn have an impact on the overall success with the technology itself in later periods.

2 Modeling Profit Growth and Information Dynamics

To motivate the vector autoregression (VAR) model of profits and information flows, I briefly outline the structural and reduced form equations for profitability and information network structure that can be used to generate the system of time-varying variables for the VAR.

2.1 Own profits

Farmer profits from a new agricultural crop are uncertain due to the fact that farmers may lack useful information about the management of the crop during the growing period. Foster and Rosenzweig (1995) characterize this uncertainty with a target input model adapted to agricultural production. In their model, the currently appropriate level of inputs for one unit of the agricultural crop is a random variable ($\bar{\theta}_{it} = \theta^* + u_{it}$), with expected value θ^* and is unknown to farmer i until the time of harvest. Profits π_{it} per unit of the new crop planted are therefore:

$$\pi_{it} = \eta - (\theta_{it} - \bar{\theta_{it}})^2 \tag{1}$$

with θ_{it} representing the actual level of inputs applied to the crop and η representing the maximum profits possible for the new crop.

Foster and Rosenzweig (1995) show that the expected profits for the new crop are a function of the farmer's posterior beliefs about the variance of the optimal level of inputs, θ^*. The farmer employs Bayesian updating to minimize this uncertainty by observing repeated trials with the new technology. These trials can either represent the farmer's own cumulative experience O_{it} or in the sum of all of the trials with the new crop of their information neighbors I_{it}. Following Foster and Rosenzweig, in reduced form, I represent profits for farm i at time t (π_{it}) as:

$$\pi_{it} = F(O_{it}, I_{it}, \mu_i, \epsilon_{it}) \tag{2}$$

where O_{it} represents the cumulative amount of information gained through the farmer' s own experience with the new technology, and I_{it} represents the cumulative information available to farmer i from farmers in $i\prime s$ information network. μ_i is an unobserved farmer effect that may influence profits, and ϵ_{it} is an idiosyncratic year/farm shock to profits. Lagged own profits in part determine the ability to experiment with the new crop. In my final model, own information is therefore simplified and represented by m lags of profits from growing pineapple. Thus, $O_{it} \equiv (\pi_{it-1}, \ldots, \pi_{it-m})$.

2.2 Information flows

Santos and Barrett (2005) and Fafchamps and Gubert (2007) outline models of social network formation that use social distance to estimate the probability of a link between two individuals i and j. These link probabilities (P_{ijt}) depend upon the characteristics of each individual (x_i, x_j) both directly and in terms of a function describing some measure of the differences between relevant characteristics to describe social distance $f(x_i, x_j)$. I assume that the following considerations are important for farmer i in obtaining information from potential social network partner j:

1. Household j has 'good information ' ($\pi_{jt-1} \geq \pi_{it-1}$)
2. Household j is willing to share information with household i ($P_{ijt} > 0$)

I thus model the total amount of information available to farmer i at time t (I_{it}) as the sum of all of the information (N_{jt}) from farmer i's potential information neighbors at time t ($j = 1, 2, \ldots, NB_I$)[4] multiplied by the probability

[4] NB_i is the total number of individuals in farmer i's information neighborhood. Thus, NB_i varies in number over each individual.

of a linkage between farmer i and neighbor j at time t, P_{ijt}. In this paper, I am interested in whether profitability from growing pineapples, π_{it}, is a factor in determining information flows. From a strategic perspective, relative profitability with the new crop matters as it determines which individuals a farmer will seek out to obtain additional information. Relative profitability may also be considered as one dimension of an individual' s characteristics that determine social distance. Generally, the relationship between information levels and profits can be represented as:

$$I_{it} = \sum_{j}^{NB_i} P_{ijt}(\pi_{it}, \pi_{jt}, x_{it}, x_{jt}, f(\pi_{it}, \pi_{jt}, x_{it}, x_{jt})) \bullet N_{jt} \tag{3}$$

or, in reduced form:

$$I_{it} = F(\pi_{it}, \pi_{jt}, x_{it}, x_{jt}, N_{jt}, NB_i) \tag{4}$$

Thus, drawing upon the works cited above, I can test the extent of the feedback between farmer profits (π_{it}) and information (I_{it}) by estimating the following vector autoregression model:

$$\begin{aligned} \pi_{it} &= \alpha_{0t}^{\pi} + \sum_{k=1}^{m+1} \beta_{kt}^{\pi} \pi_{i,t-k} + \sum_{k=1}^{m+1} \gamma_{kt}^{\pi} I_{i,t-k} + \Psi_t^{\pi} \mu_i + \epsilon_{it}^{\pi} \\ I_{it} &= \alpha_{0t}^{I} + \sum_{k=1}^{m+1} \beta_{kt}^{I} \pi_{i,t-k} + \sum_{k=1}^{m+1} \gamma_{kt}^{I} I_{i,t-k} + \Psi_t^{I} \mu_i + \epsilon_{it}^{I} \\ &(i = 1, \ldots, N, t = 1, \ldots, T) \end{aligned} \tag{5}$$

with m representing the number of time lags appropriate for the VAR and (μ_i) indicating farmer fixed effects. Profits and information are thus both assumed to be endogenous.[5] Analysis of the α, β and γ parameters provides information on the effects of changing farmer profits and information on each other. As well, the system in (5) provides a framework for conducting causality tests between each of these variables.

[5]As individual characteristics, x_i are not changing over time, they have been excluded from the I_{it} equation in the model and are assumed to operate on profits and information levels exclusively through the unobserved individual effect, μ_i. The N_{jt} and NB_i variables are implicitly included in the model through the method of constructing the final information variable, which is described in detail in a subsequent section.

In order to estimate the parameters in (5) and account for unobserved farmer heterogeneity (μ_i), I follow Holtz-Eakin, Newey, and Rosen (1988) and apply quasi-differencing to the model. Quasi-differencing is a method suggested by Chamberlain (1983) that allows for consistent estimation of parameters in a panel data model with lagged endogenous variables, as traditional fixed effects techniques, such as first differencing, result in biased estimates (Nickell, 1981). The method involves treating the parameter on the fixed effect as time-varying (shown in (5) as Ψ_t) and transforming the system by multiplying the system in time $t-1$ by the ratio of the fixed effect parameter in time t to the parameter in time $t-1$ (i.e. $r_t = \Psi_t/\Psi_{t-1}$). This procedure is outlined in more detail in the appendix.

After quasi-differencing, actual estimation is carried out on the following transformed model:

$$\begin{aligned}
\pi_{it} &= a_{0t}^{\pi} + \sum_{k=1}^{m+1} b_{kt}^{\pi}\pi_{i,t-k} + \sum_{k=1}^{m+1} d_{kt}^{\pi} I_{i,t-k} + e_{it}^{\pi} \\
I_{it} &= a_{0t}^{I} + \sum_{k=1}^{m+1} b_{kt}^{I}\pi_{i,t-k} + \sum_{k=1}^{m+1} d_{kt} I I_{i,t-k} + e_{it}^{I} \\
&(i = 1, \ldots, N, t = m+3, \ldots, T)
\end{aligned} \tag{6}$$

Identification of the original model parameters in (5) from estimates of the reduced form parameters in (6) is only possible if the total number of time period observations T is sufficiently large.[6] However, it is still possible to use the reduced form estimates to test Granger causality between profits and information (Holtz-Eakin, Newey, and Rosen, 1988). As this is of primary interest in this paper, and the length of the panel is relatively short, all subsequent analysis is conducted using the reduced form model, (6).

Note that, due to the differencing, I am not able to estimate parameters for the first $m+2$ time periods. Cameron (1999) points out that this may be a problem in estimating the effect of learning on technology adoption in the case of new crops where learning takes place relatively quickly. However, in the data, all of the farmers are already adopters (all the households are pineapple farmers), so I am actually examining the interaction between profitability and information flows on farmers who are already users of the new technology. In terms of fully managing all of the aspects of the pineap-

[6] Holtz-Eakin, Newey, and Rosen (1988) show that this is only possible if $T-m-2 \geq 2m$.

ple crop, including negotiating with exporters over sales, fertilizer application rates and timing and other management decisions, the learning process may continue for quite some time after the crop has been adopted, and new information may continue to arrive. And a primary advantage of using quasi-differencing is that I am able to account for time-varying unobserved heterogeneity in farmer learning. So, the benefits of using quasi-differencing may outweigh the drawbacks in this particular setting.

The actual estimation of the transformed system (6) is performed by GLS following Holtz-Eakin, Newey, and Rosen (1988). They demonstrate that appropriately selected lags of all of the endogenous variables can be used as instruments (Z_t) which can correct for the endogeneity introduced by the quasi-differencing necessary to remove the unobserved fixed effect. The model in (6) allows for complete non-stationarity of the parameters on profits and information over time. In other formulations of the panel vector autoregression, it is possible to estimate models with stationary parameters for each of the right-hand side variables by imposing cross-equation restrictions, but these restrictions were not appropriate for this data. The procedure for constructing the final estimator is outlined in the appendix.

3 Data and Panel Construction

The data come from a panel household survey among pineapple farmers in Ghana between 1997 and 1999. Data were collected approximately every six weeks, for a total of 15 rounds of data, so the relevant time period t for the model in (6) is six weeks. Households in the survey area undertake a wide variety of farming activities, with many taking up the production of pineapple fruit for export to European markets. There was limited contract farming among pineapple farmers at the time of the first round of the survey.[7] Pineapple farmers therefore were learning about pineapple production mostly through social networks and extension services.

The survey collected detailed information on household farm composition, including timing of agricultural activities by plot such as harvests, sales and input costs. Crops on specific plots are also enumerated over the 15 rounds, with farmers changing plot composition over time. Households can be observed harvesting pineapple on multiple plots and in multiple rounds.

[7] In subsequent rounds, contract farming appears to have become more prevalent (Suri, 2008).

It is this variation that allows the construction of the panel used for this paper. Table 1 contains household and round specific summary statistics of the data used in this paper.

The survey is also designed to collect detailed information about farmers's information neighborhoods. Information about who farmers know and who they talk to about farming activities, as well as frequency of interaction with these contacts were recorded. In addition, a general roster of farmer contacts was collected and updated for the last six rounds of the survey. For this information roster, farmers are asked in each round with whom they had significant conversations in that particular round. This list of information contacts is recorded for each of the last six rounds and is used to construct the time varying level of information available to farmers in the panel necessary to evaluate the model in (6).

In order to estimate model (6), I created time series variables on farmer profits from pineapple and information on pineapple production from information neighbors. For farmer profits, the variable used is average profits per plant per round. Average profits are calculated by subtracting labor, input and land costs for each pineapple harvest from revenues earned from pineapple sales and dividing by the total number of pineapples currently in production.[8] Revenues, inputs[9] and land costs are recorded in cedis, the Ghanaian currency. Hired labor is also given in monetary terms for the cedis paid to workers, and household labor is evaluated at the prevailing gender-specific wage rate, under the assumption that labor markets in the area function relatively well.

3.1 Per round profits

With multiple harvests on a single plot, it becomes difficult to disentangle input usage across different harvests on the plot.[10] In this situation, the

[8] Pineapple revenues include sales of ripe fruits. The number of pineapples in production is estimated from the total number of pineapples harvested in each round. For plots with multiple harvests in multiple rounds, it is assumed that the harvested fruits have been in production on the plot for the past 10 rounds.

[9] Inputs refer to any purchased fertilizer or pesticide used during a given round on the pineapple plot

[10] This is also due to the fact that pineapple growth occurs over 10 rounds and the survey length is only 15 rounds long. Therefore harvests on a given plot at intervals smaller than 10 are likely from the same batch of pineapple plants under cultivation which may ripen at different times due to natural variability in pineapple growth.

average profits on each pineapple harvest in the plot are calculated by adding up all revenues and costs and dividing by total number of plants in production as of each successive harvest. For example, if a farm had two observed pineapple harvests on a given plot, one in round 6 and one in round 10, the recorded value for the first average profit would take the total profit on the pineapples up to round six and divide by the number of pineapples harvested in round 6 plus the estimated additional pineapples that will subsequently be harvested in round 10. This average profit value is attributed to the household for rounds 6, 7, 8, and 9 (i.e., until the next realized pineapple harvest). The average profit on the second harvest in round 10 would sum up the revenues and costs for the pineapples in round 6 and round 10 and divide by the total number harvested in round 10. This average profit calculation is then the observed average profit for the farmer for rounds 10, 11, 12, 13, 14 and 15. Across plots, per round profits are averaged across the time varying, round specific total numbers of pineapple plants in each plot. Although this method of averaging profits tends to dampen the changes in profitability in each round, it avoids a more serious issue of extreme positive spikes in profitability that occur in rounds when input costs in that round are minimal and farmer profits are mostly composed of revenues from sales. For a crop like pineapple, which takes a year to grow and can be harvested year round, evaluating profits round by round and not accounting for previous outlays is not an appropriate way to represent the total profits from the new crop. Better aggregation of revenues and costs into profits would be possible with data from a longer panel, but the data available cover only two years of activity, with many of the pineapple plantings occurring during the second year. Thus the averaging technique above is judged to be the best alternative under these limitations.

Table 1: Round-specific summary statistics on pineapple profits and information network variables ($N = 88$)

	Round[a]					
	10	11	12	13	14	15
Own Profits (cedis per plant)	3372.177	4140.109	4041.862	4038.31	4149.255	4152.449
	(11165.300)	(12137.810)	(12231.990)	(12372.090)	(12467.780)	(12472.79)
Own cumulative information O_{it} (plants)	11028.560	18437.75	30115.170	39473.300	48848.800	59068.87
	(20778.24)	(31880.87)	(55075.200)	(71693.91)	(87086.04)	(105031.200)
Number of Contacts	1.475	0.275	0.550	0.150	0.225	0.363
	(1.331)	(0.573)	(1.282)	(0.480)	(0.551)	(0.641)
Average Cumulative Neighbor Info I_{it} (plants)	14876.140	16706.510	21630.760	22220.300	25626.470	27781.500
	(18540.730)	(20231.980)	(28121.000)	(28432.740)	(30351.940)	(34983.080)
Number of HHs in Village 1 (N_1)	10					
Number of HHs in Village 2 (N_2)	38					
Number of HHs in Village 3 (N_3)	40					

[a]Standard deviations in parentheses.

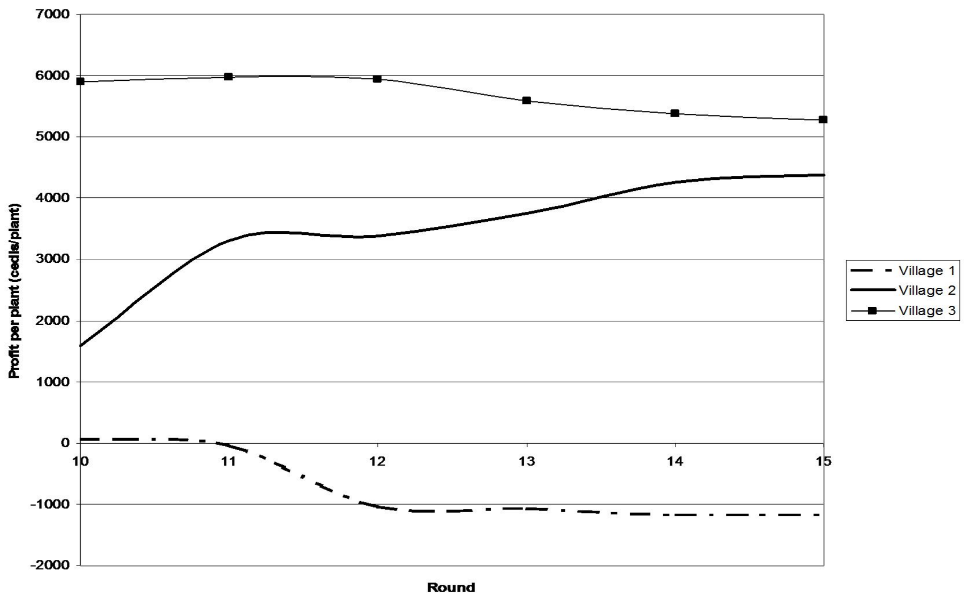

Figure 1: Average per round profits by village ($N_1 = 10, N_2 = 38, N_3 = 40$)

Figure 1 shows the average per round profits for each village in the sample over all of the rounds of the panel, as calculated in the manner described above. As can be seen, per plant profits are growing in Village 2, but declining and stabilizing for Villages 1 and 3. Average profits are also only positive for Villages 2 and 3. This may be due to the fact that pineapple farming is more established in these villages.

3.2 Per round information availability

The round-specific measure of information about pineapple production used is the total number of pineapples ever harvested by members of the household's information network, as of that round. The household uses these observations on pineapple plant harvests to update their own information on overall plant production. For the amount of information available from a particular farmer, I use data on observed pineapple harvests that records numbers of pineapple fruits sold by each observed pineapple farmer in the sample. The round in which the sale takes place is also recorded. Therefore, for each farmer in the sample, it is possible to construct a round-by-round tally of the total numbers of pineapples they have ever harvested. This number changes as farmers make multiple harvests. For farmers with only one recorded harvest, the amount of information available from this farmer is zero up until the time of their first harvest, and then is equal to the numbers harvested in the harvest round, and remains at this number for all of the subsequent rounds until the end of the survey period. With multiple harvests, the amount of information is equal to the number of pineapples in the first harvest until the time of the second harvest. The information is then the sum of the total pineapples from both harvests, until either the end of the survey or the next observed harvest period.

Farmers were asked between rounds 10 and 15 to list all of the individuals with whom they had significant conversations in each round in order to construct a roster of information contacts over the survey period. Therefore, to determine the total amount of information on pineapple production from farmer $i\prime s$ social network available per round, I use this changing roster of contacts in combination with the round-specific measure of information on pineapples described above. More specifically, for each person listed in farmer $i\prime s$ roster in a given round, I add up the associated number of pineapples ever harvested for the individuals contacted as of the round in which the conversation occurred. Farmers are assumed not to forget information obtained in previous rounds. Therefore, in addition to the information obtained from having conversations with different pineapple farmers in different rounds, the farmer also carries forward any information gathered in previous rounds. Finally, farmers update their information if they speak with the same individual in multiple rounds of the survey. The information available to a farmer i in round t is therefore equal to the sum of the number of pineapples ever harvested as of round t by each person with whom farmer i speaks in that round plus the latest information gathered in previous rounds $(t-1, t-2, ...1)$ from other individuals in the farmer's social network (as determined by observing past conversations), but with whom the farmer has not spoken in round t. Changes in the composition of the list of conversation contact individuals reflect the changing decisions of farmers on to whom they turn for information, as well as the possibility that new individuals might become more accessible over time.[11]

[11]Information on the length of time over which farmers who have conversations have known each other is also available. The average length is approximately 17 years, but the range runs between one day and 80 years.

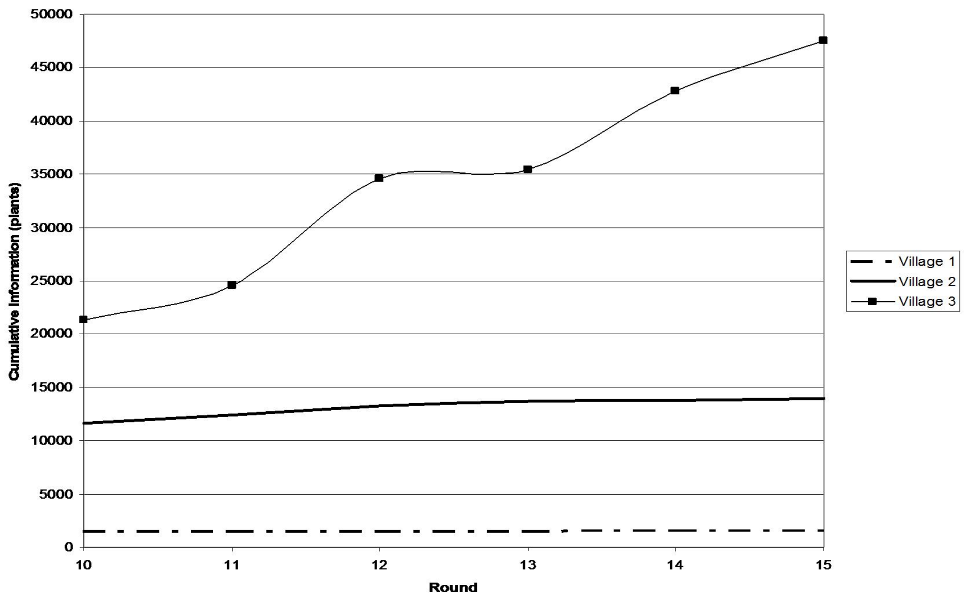

Figure 2: Average per round information by village ($N_1 = 10, N_2 = 38, N_3 = 40$)

The household's information time series thus consists of the information available from the social network, which is changing over time both in composition of network contacts and in total cumulative numbers of pineapples harvested by network contacts. It is possible that not all 'significant conversations' recorded concerned pineapple farming. Information is available in the sample on who specifically farmers speak to about pineapple production; however this information was only collected in one round. Therefore I use the more general measure of information network given by the changing list of people with whom farmers talk in a given period. Other information network measures, like networks of family and friends, are also not specific to a particular enterprise but have been shown to be useful in making predictions about agricultural production behavior (Bandiera and Rasul, 2006). Figure 2 presents average information flows per round for each village. Village 3 dominates in the overall sample in terms of the prevalence of pineapple farming (Goldstein and Udry, 1999), which may explain the overall greater amount of information availability for farmers in this village. There is a sample truncation issue involved in using the roster of individuals with whom a farmer has conversations to construct the information network, as not all of the individuals listed in the roster are included in the sample of individuals in the dataset (i.e., the household may indicate conversations with other individuals, like family members living out of town or other villagers who have not been included in the sample.). Indeed, for the pineapple farmers in this study, only approximately 20% of their recorded information exchanges were with other individuals in the sample. But, as the information measure used is cumulative, the information levels recorded with sample households represents a lower bound on the total amount of information available to the farmer. Therefore, further information from other sources is likely merely to reinforce any results presented here.

After creating the necessary time series variables, I combined them into a panel dataset for as many farmers as possible in the sample. Given that the information roster data for each household is only available for the last six rounds of the survey, the panel only covers these periods (i.e., from round 10 to round 15). In the sample, there are 209 plots on which pineapple is grown, spread among 91 farmers. However, a record of at least one pineapple fruit harvest is missing for 11 farmers, leaving a sample of 80 farmers with six rounds of data on profits and information.

4 Results

The estimates of the transformed model parameters (from equation 6) are shown in Tables 3 and 4.

4.1 Dynamic analysis of profits and information

Table 3 shows the estimates with own average profits as the dependent variable, while Table 4 has information as the dependent variable. The estimates from the full (reduced form) model (as in equation (6)) are shown in column (1) in each table, while the results from estimation of a restricted model, (used to conduct Granger causality tests) are shown in column (2). In the course of some preliminary tests of various lag lengths, the best fit was obtained for models with two lags (i.e., $m = 2$).[12] Therefore, the subsequent

[12]Both the Akaike Information Criterion and a comparison of the sum of squared residuals between models confirmed that at least two lags are best. The short length of the panel limits the lag lengths that can be tested to a maximum of two.

analysis is performed with two lags included.

4.2 Own profits, model (1)

In the model with farmer's own average profits as the dependent variable, the estimates using the full model (5) with own profits and information seems to fit the data quite well. The sum of squared residuals for this model is equal to 0.143 (this is the Q statistic shown at the bottom of the column (1) estimates). As demonstrated by Holtz-Eakin, Newey, and Rosen (1988), this Q statistic has a chi-square distribution. Estimation of the parameters of equation (5) is performed using lags of the endogenous variables as instruments for the right-hand side variables, which become correlated with the error term through the quasi-differencing procedure. The number of these instruments minus the number of parameters estimated determines the degrees of freedom of the Q statistic and can be used to perform a number of hypothesis tests on the appropriateness of the model specification, as well as tests for Granger causality in any of the different right hand side variables. For model (1), the Q statistic has 2 degrees of freedom. As a conservative test of fit, the 90% critical value for a chi-square random variable with 2 degrees of freedom is 4.61. Therefore, with a Q statistic of 0.143, I cannot reject the hypothesis that the full model as shown in column (1) is appropriate for the underlying data generating process.

The parameter estimates for the full model show that several of the lags of both own profits and neighbor information are significant in determining current profits with pineapple. However the influence of the longer lags (which represents 12 and 18 weeks of time, respectively) is mixed. The difference in effects of earlier versus later lags of the own profits and information variables may suggest that older information, both from one's own experiments (given by lagged profits) and others (given by lagged information) is less useful in achieving higher profits than newer information. The information about pineapple farming, either from own profits or from harvests of information neighbors in earlier periods may not remain relevant for farmers using it as a tool to improve their profits from the new crop in the current period, as marketing conditions may change.

4.3 Granger causality test of the effect of information on profits (model(2))

The parameter estimates in column (2) in Table 3 represent estimates from a restricted model that excludes information flows. By comparing the Q statistic from the restricted model (Q^R) with that of the unrestricted model (Q), following Holtz-Eakin, Newey, and Rosen (1988, 1989), I can conduct a Granger causality test of whether or not information flows can be said to cause farmer profits from growing pineapples. The appropriate test statistic, L, is a criterion function statistic that takes the difference in the sum of squared residuals in the restricted model and the unrestricted model (Wooldridge, 2002, pg. 202)) (i.e., $L \equiv Q^R - Q$.) This statistic has a limiting chi-square distribution, with degrees of freedom equal to the difference in the degrees of freedom of the restricted and unrestricted statistics. By eliminating information flows, the difference in the sum of squared residuals between model (2) and model (1) is $L = 213.945 - 0.143 = 213.802$ and the degrees of freedom of L are: $d.o.f(L) = 8 - 2 = 6$. The 90% critical value for a chi-square random variable with four degrees of freedom is 10.64. Therefore, I can strongly reject the hypothesis that information flows do not cause farmer profits. This is in line with other results on learning and technology adoption, like Foster and Rosenzweig (1995).

Table 2: Unconstrained and constrained estimates of the panel vector auto regression. Dependent variable: Average own profits per round (cedis/plant) (in grey) $N = 88$, $T = 6$

		(1)			(2)		
Round	**Parameter Name**	**Estimate**	**s.e.**		**Estimate**	**s.e.**[a]	
14	Constant($a^{\pi}_{0,14}$)	-54.422	20.733	***	-54.798	3.869	***
14	Own ave. profits, t-1 ($b^{\pi}_{1,13}$)	0.268	1.351		0.234	0.214	
14	Own ave. profits, t-2 ($b^{\pi}_{2,12}$)	1.537	1.697		1.577	0.259	***
14	Own ave. profits, t-3 ($b^{\pi}_{3,11}$)	-0.808	0.354	***	-0.814	0.046	***
14	Information, t-1 ($d^{\pi}_{1,13}$)	177.399	27.566	***			
14	Information, t-2 ($d^{\pi}_{2,12}$)	-175.787	26.642	***			
14	Information, t-3 ($d^{\pi}_{3,11}$)	-3.627	0.981	***			
15	Constant($a^{\pi}_{0,15}$)	-3.391	1.91		-3.363	0.176	***
15	Own ave. profits, t-1 ($b^{\pi}_{1,14}$)	1.33	0.015	***	1.331	0.001	***
15	Own ave. profits, t-2 ($b^{\pi}_{2,13}$)	-0.333	0.016	***	-0.334	0.002	***
15	Own ave. profits, t-3 ($b^{\pi}_{3,12}$)	0.002	0.001		0.002	0.001	***
15	Information, t-1 ($d^{\pi}_{1,14}$)	0.203	0.373				
15	Information, t-2 ($d^{\pi}_{2,13}$)	-5.803	0.463	***			
15	Information, t-3 ($d^{\pi}_{3,12}$)	5.648	0.365	***			
	Q (sum of sq. res.)	0.143			213.945		
	d.o.f.(Q)	2			8		
	AIC[b]	22.159			34.631		

[a] ***,**,* indicates significant at the 10, 5 and 1% level respectively.
[b] AIC=Akaike Information Criterion.

4.4 Information flows, model (1)

Table 4 presents estimates of the model in equation (6) with information flows as the dependent variable. Lagged information is significant in explaining current information availability in round 14 and lagged profits and information are both strongly significant in explaining current information levels in round 15. Lags of information generally positively affect current information flows. Recent lags of profits negatively affect information, while longer lags positively affect current information availability. This may reflect changes in the substitutability of information between oneself and one's information neighbors. In some periods, larger profits may limit the need to seek out information from others, while at other times, larger profits may encourage greater communication between individuals by reducing the relative social distance between them. However, a more structured analysis than the one presented here is necessary to further examine these kinds of dynamics and is beyond the scope of this paper. The overall fit of the full model with information as the dependent variable is also not strong and this full model is rejected at the 3% level of significance. Further specification tests or additional covariates may be necessary to better identify the information equation.

4.5 Granger causality test of the effect of profits on information (model(2))

Column (2) in Table 4 shows the estimates of a restricted model without profits in order to test Granger causality between profits and information. Using the same test statistic as described in section 4.2, the test statistic L is: $L = 227.002 - 6.743 = 220.259$ and its degrees of freedom are $d.o.f(L) = 8 - 2 = 6$. With this value for L, I can still strongly reject the hypothesis that profits do not cause information, and they appear to have a large amount of explanatory power in the information equation in round 15. Therefore, this demonstrates that profits with a new technology may also impact information networks. This fact may be critical in understanding patterns of technology diffusion and the subsequent impact of the introduction of new technologies on farmer welfare. The information from the social network responds to profitability, and given the results in section 4.2 profitability eventually responds to changes in information levels.

These results may add to the discussion presented in Moser and Barrett (2006) on social conformity and technology adoption and disadoption decisions by providing further empirical evidence of the interaction between technology use patterns and social dynamics. For example, among the results presented in this paper, it is the case that in some periods, such as round 15, that there appears to be a reinforcing process at work between profitability with a new technology and information flows. When results from both the profit and information equation are taken simultaneously, the positive parameter estimates on the one period lagged information variable in the profit equation as well as a positive parameter estimate on the one period lagged profit variable in the information equation indicate that successful pineapple farmers have access to more information, which brings them greater profitability with the new crop, while relatively unsuccessful farmers obtain less information overall about pineapple farming from their network, leading to lower profitability. If profitability were to fall to an extremely low level, one might imagine that a farmer may decide to disadopt the crop altogether. The results in this paper suggest that this decision would in part be due to the impact of changes in the information available from the social network. Less successful farmers may become more socially invisible, with information decreasing due to fewer contacts over time, or may be shut out of better information networks, with the information levels from existing network contacts declining over time. Moser and Barrett (2006) suggest that social conformity effects influence the disadoption decision. Social pressures on unsuccessful farmers through increasing costs of network maintenance may also be a part of this process.

Table 3: Unconstrained and constrained estimates of the panel vector auto regression.
Dependent variable: Information per round (Number of plants) (in grey) $N = 88$, $T = 6$

		(1)			(2)		
Round	**Parameter Name**	**Estimate**	**s.e.**		**Estimate**	**s.e.**[a]	
14	Constant($a^I_{0,14}$)	891.752	692.612		1997.912	116.924	***
14	Own ave. profits, t-1 ($b^I_{1,13}$)	-6295.09	4170.242				
14	Own ave. profits, t-2 ($b^I_{2,12}$)	6328.708	5270.611				
14	Own ave. profits, t-3 ($b^I_{3,11}$)	-67.727	1126.221				
14	Information, t-1 ($d^I_{1,13}$)	88.307	11.269	***	-13.830	0.514	***
14	Information, t-2 ($d^I_{2,12}$)	38.81	9.548	***	40.734	0.474	***
14	Information, t-3 ($d^I_{3,11}$)	0.813	0.326	***	0.596	0.098	***
15	Constant($a^I_{0,15}$)	-122.216	18.333	***	-72.018	8.575	***
15	Own ave. profits, t-1 ($b^I_{1,14}$)	-169.328	15.18	***			
15	Own ave. profits, t-2 ($b^I_{2,13}$)	161.718	16.074	***			
15	Own ave. profits, t-3 ($b^I_{3,12}$)	9.519	1.604	***			
15	Information, t-1 ($d^I_{1,14}$)	1.08	0.037	***	0.928	0.009	***
15	Information, t-2 ($d^I_{2,13}$)	-0.391	0.053	***	-0.267	0.018	***
15	Information, t-3 ($d^I_{3,12}$)	0.352	0.031	***	0.373	0.015	***
	Q (sum of sq. res.)	6.743			227.002		
	d.o.f.(Q)	2			8		
	AIC	33.726			32.275		

[a] ***,**,* indicates significant at the 10, 5 and 1% level respectively.

5 Conclusions

The results of this work have shown complex nature of technology use and information diffusion among small farmers undertaking a new agricultural crop. There are many possible interactions that affect profitability that may be important in considering the potential of new crops to lift poorer rural households out of poverty.

First, information about a new technology from a farmer's social network is important in the overall profitability of a crop. This may be critical in determining farmers management of a new crop.[13] Farmers without adequate information networks will not be as successful with the new crop as those with high quality information networks. Socially isolated farmers may need an injection of information inflows from an external source (perhaps from an extension service) to ensure initial success with a new crop.

Second, information flows in this sample are also affected by current crop profitability. Although the effects of profitability change over time, in certain periods, unsuccessful farmers may find that they are able to obtain less useful information about the crop, with potentially serious results on future profitability. On the other hand, farmers with high levels of success with a new crop may be able to eventually secure better information network contacts for themselves, and the need for external assistance and information support may become less critical over time. It may be important to identify and more specifically characterize social network structures, in order to better ensure wide-spread diffusion of a beneficial technological intervention among both well-connected and less visible farmers.

Third, there appear to be tradeoffs between learning by doing and learning from others that change over time, with own experimentation and learning by doing substituting for learning from others in certain periods, and profitability and learning by doing facilitating expansion of the information available from the social network in others.

Analysis of the dynamics of profitability and information flows has been conducted in past studies (Foster and Rosenzweig, 1995; Munshi, 2004) but usually under some strong assumptions that a farmer's information network is composed of their entire village, which has been shown not to be the case (Conley and Udry, 2001). In looking at this process with better information about the farmer's network contacts, it is possible to begin to better

[13] This has been shown to be the case in Bandiera and Rasul (2006).

distinguish farmers who are likely to be successful with a new technology, by virtue of their improved access to information, from those that may not be able to change their agricultural crop portfolio due to a lack of critical information from others. Further, profits may eventually affect information flows, suggesting that the process of technology diffusion may never reach certain members of a village and that those who do not adopt might also lose opportunities for increasing the size of their social network, which may have other consequences as shown in the risk-sharing literature (Townsend, 1994; de Weerdt and Dercon, 2006). Knowledge of the different possible relationships between information flows and new technology dynamics through the use of time series analysis may help improve understanding of the process of technology diffusion as well as the ingredients necessary for new adoption and the consequences for existing social network structures.

A The Transformed Panel Vector Autoregression Model

In order to eliminate the individual fixed effect, μ_i from the model shown in equation 4, I employ the process of quasi-differencing proposed by Chamberlain (1983). This process involves multiplying the model in equation 4 at time $t-1$ by $r_t = \Psi_t/\Psi_{t-1}$, the ratio of the time-varying parameters on the fixed effect, and then subtracting this transformed equation from the equation representing behavior at time t. This removes the fixed effect, and allows for unbiased estimation of the model parameters, given that it includes lags of endogenous variables.

In the model I estimate in this paper, with two lags ($m = 2$) and two endogenous variables, the quasi-differencing procedure results in the following transformation of the model parameters:

$$
\begin{aligned}
\pi_{it} &= \alpha^{\pi}_{0t} + \beta^{\pi}_{1t}\pi_{it-1} + \beta^{\pi}_{2,t}\pi_{it-2} + \gamma^{\pi}_{1t}I_{i,t-1} \\
&\qquad + \gamma^{\pi}_{2t}I_{i,t-2} + \Psi^{\pi}_{t}\mu_i + \epsilon^{\pi}_{it} \\
-r_t\pi_{it-1} &= r_t\alpha^{\pi}_{0,t-1} + r_t\beta^{\pi}_{1t-1}\pi_{i,t-2} + r_t\beta^{\pi}_{2t-2}\pi_{i,t-3} + r_t\gamma^{\pi}_{1t-1}I_{i,t-2} \\
&\qquad \underline{+ r_t\gamma^{\pi}_{2t-2}I_{i,t-3} + r_t\Psi^{\pi}_{t}\mu_i + r_t\epsilon^{\pi}_{it-1}} \\
\pi_{it} &= a^{\pi}_{0t} + b^{\pi}_{1t}\pi_{i,t-1} + b^{\pi}_{2t}\pi_{i,t-2} + b^{\pi}_{3t}\pi_{i,t-3} + d^{\pi}_{1t}I_{i,t-1} \\
&\qquad + d^{\pi}_{2t}I_{i,t-2} + d^{\pi}_{3t}I_{i,t-3} + e^{\pi}_{it}
\end{aligned}
\tag{7}
$$

with:

$$\begin{aligned}
a_{0t} &= \alpha_{0t} - r_t \alpha_{0t-1} \\
b_{1t} &= \beta_1 + r_t \\
b_{2t} &= \beta_{2t} - r_t \beta_{1t-1} \\
b_{3t} &= -r_t \beta_{2t-2} \\
d_{1t} &= \gamma_{1t} \\
d_{2t} &= \gamma_{2t} - r_t \gamma_{1t-1} \\
d_{3t} &= -r_t \gamma_{2t-2} \\
e_{it} &= \epsilon_{it} - \epsilon_{it-1}
\end{aligned} \tag{8}$$

Note that the quasi-differencing results in time-varying coefficients in the transformed model and also implies cross equation restrictions across time periods. The new error term on the transformed model is a function of errors from time t and time $t-1$. This is the reason why the instruments are necessary for estimation of the transformed model and why they must have at least a lag length of two to be appropriate.

B GLS Estimation of the Transformed Panel Vector Autoregression System

Following Holtz-Eakin, Newey, and Rosen (1988) I estimate the transformed model 5 by general least squares. Given that the first four time periods cannot be included in the estimation directly, due to the quasi-differencing procedure described above, I estimate a two period system of equations for each endogenous variable (own profits and information flows).

Designating profits as the dependent variable for the purposes of illustration, the system of equations for profits (Y) and information (X) over $i \in \{1, 2, ..., N\}$ individuals, $Y_t = [y_{1t}, ...y_{Nt}]'$, $X_t = [x_{1t}, ...x_{Nt}]'$ can be described by the general linear model:

$$
\begin{aligned}
&Y = WB + U \\
&\text{with} \\
&Y = [Y'_{m+3}, ..., Y'_T]' \\
&((T-m-2)Nx1) \\
&B = [B'_{m+3}, ..., B'_T]' \\
&((T-m-2)(2m+3)x1) \\
&U = [U'_{m+3}, ..., U'_T]' \\
&((T-m-2)Nx1) \\
&W = diag[W_{m+3}, ..., W_T]
\end{aligned}
\tag{9}
$$

where W_t is the $(Nx(2m+3))$ matrix of right-hand side variables, $W_t = [1, Y_{t-1}, ..., Y_{t-m-1}, X_{t-1}, ...X_{t-m-1}]$ and $diag[]$ designates a block diagonal matrix with the given matrix entries along the diagonal. B is the vector of least squares parameters to be estimated.

GLS requires the use of a consistent weighting matrix, Ω. This can be estimated in a first stage 2SLS estimation of the round specific parameters which generates round specific residuals, $\hat{u}_t$ that can be used to construct $\hat{\Omega}$ as follows:

$$(\hat{\Omega}/N)_{rs} = \sum_{i=1}^{N} (\hat{u}_{ir}\hat{u}_{is}Z'_{ir}Z_{is})/N \tag{10}$$

with $\hat{u}_{it}$ estimated over the time periods $t = 5, 6$ $(t = r, s)$ for all farmers in the sample, $i = (1, \ldots, N)$.

Z'_{it} is the l_t dimensional vector of household/round specific instruments available at time t that can be used to estimate the transformed model, 3.5. Note that the number of instruments increases as time goes on, because all lags of the endogenous variables longer than two time periods can be used as instruments. Therefore, in t=5, there are 3 lags of each endogenous variable plus the constant that can be used (therefore $l_5 = 7$), while in t=6, there are 4 lags available (and $l_6 = 9$).

After estimating $\hat{\Omega}/N$, the final GLS estimator $\hat{B}$ is:

$$\hat{B} = [W'Z(\hat{\Omega})^{-1}Z'W]^{-1}W'Z(\hat{\Omega})^{-1}Z'Y \tag{11}$$

with standard errors $diag(\sqrt{\left[W'Z(\hat{\Omega})^{-1}Z'\right]^{-1}})$.

References

Bandiera, O., and I. Rasul. 2006. "Social networks and technology adoption." *Economic Journal* 116:869–902.

Bardhan, P., and C. Udry. 1999. *Development microeconomics*. Oxford University Press.

Calvó-Armengol, A., and M.O. Jackson. 2007. "Networks in labor markets: Wage and employment dynamics and inequality." *Journal of Economic Theory* 132:27–46.

Cameron, L. 1999. "The importance of learning in the adoption of high-yielding variety seeds." *American Journal of Agricultural Economics* 81:83–94.

Chamberlain, G. 1983. "Panel data." In Z. Grilliches and M. Intrilligator, eds. *Handbook of Econometrics*. Amsterdam: North Holland, vol. 2, pp. 1247–1320.

Conley, T., and C. Udry. 2009. "Learning about a new technology." Unpublished, Forthcoming, *American Economic Review.*

—. 2001. "Social learning through networks: The adoption of new agricultural technologies in Ghana." *American Journal of Agricultural Economics* 83:668–673.

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

Evenson, R.E., and L.E. Westphal. 1995. "Technological change and technology strategy." In T. N. Srinivasan and J. Behrman, eds. *Handbook of Development Economics*. Amsterdam: North Holland, vol. 3A, pp. 2209–2300.

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

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

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

Goldstein, M., and C. Udry. 1999. "Agricultural innovation and resource management in Ghana." Working paper, IFPRI.

Holtz-Eakin, D., W. Newey, and H.S. Rosen. 1988. "Estimating vector autoregressions with panel data." *Econometrica* 56:1371–1395.

—. 1989. "The revenues-expenditures nexus: Evidence from local government data." *International Economic Review* 30:415–429.

Moser, C.M., and C.B. Barrett. 2006. "The complex dynamics of smallholder technology adoption: The case of SRI in Madagascar." *Agricultural Economics* 35:373–388.

Munshi, K. 2004. "Social learning in a heterogeneous population: Technology diffusion in the Indian Green Revolution." *Journal of Development Economics* 73:185–213.

Nickell, S. 1981. "Biases in dynamic models with fixed effects." *Econometrica* 49:1417–1426.

Santos, P., and C.B. Barrett. 2005. "Interest and identity in network formation: Who do smallholders seek out for information in rural Ghana?" Unpublished, Working paper, Cornell University.

Suri, T. 2008. "Technological change and the evolution of markets: The case of pineapple in Ghana." Unpublished, Working paper. MIT Sloan.

Townsend, R. 1994. "Risk and insurance in village India." *Econometrica* 62:539–592.

Vanderpuye-Orgle, J., and C.B. Barrett. 2007. "Risk management and social visibility in Ghana." Unpublished, Working paper, Cornell University.

Wooldridge, J.M. 2002. *Econometric analysis of cross section and panel data*. Cambridge: The MIT Press.