INTERNATIONAL FOOD
POLICY RESEARCH INSTITUTE
sustainable solutions for ending hunger and poverty

Supported by the CGIAR

IFPRI Discussion Paper 00816

November 2008

Evaluating the Impact of Social Networks in Rural Innovation Systems

An Overview

Ira Matuschke

International Service for National Agricultural Research Division

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

The International Food Policy Research Institute (IFPRI) was established in 1975. IFPRI is one of 15 agricultural research centers that receive principal funding from governments, private foundations, and international and regional organizations, most of which are members of the Consultative Group on International Agricultural Research (CGIAR).

FINANCIAL CONTRIBUTORS AND PARTNERS

IFPRI's research, capacity strengthening, and communications work is made possible by its financial contributors and partners. IFPRI receives its principal funding from governments, private foundations, and international and regional organizations, most of which are members of the Consultative Group on International Agricultural Research (CGIAR). IFPRI gratefully acknowledges the generous unrestricted funding from Australia, Canada, China, Finland, France, Germany, India, Ireland, Italy, Japan, Netherlands, Norway, South Africa, Sweden, Switzerland, United Kingdom, United States, and World Bank.

AUTHOR

Ira Matuschke, Food and Agriculture Organization of the United Nations
Associate Professional Officer, Global Perspective Studies Unit

Notices

[1] Effective January 2007, the Discussion Paper series within each division and the Director General's Office of IFPRI were merged into one IFPRI–wide Discussion Paper series. The new series begins with number 00689, reflecting the prior publication of 688 discussion papers within the dispersed series. The earlier series are available on IFPRI's website at www.ifpri.org/pubs/otherpubs.htm#dp.

[2] IFPRI Discussion Papers contain preliminary material and research results. They have not been subject to formal external reviews managed by IFPRI's Publications Review Committee but have been reviewed by at least one internal and/or external reviewer. They are circulated in order to stimulate discussion and critical comment.

Copyright 2008 International Food Policy Research Institute. All rights reserved. Sections of this material may be reproduced for personal and not-for-profit use without the express written permission of but with acknowledgment to IFPRI. To reproduce the material contained herein for profit or commercial use requires express written permission. To obtain permission, contact the Communications Division at ifpri-copyright@cgiar.org.

Contents

List of Tables

List of Figures

ACKNOWLEDGMENTS

This paper was developed while the author was a consultant at the ISNAR Division in Addis Ababa, Ethiopia. The author thanks David J. Spielman, Kristin Davis, and Frank Hartwich for their valuable comments on earlier drafts of this paper. The support and supervision of Melinda Smale and Kwadwo Asenso-Okyere during the consultancy is gratefully acknowledged.

The views expressed in this document are those of the author, not necessarily those of the Food and Agriculture Organization of the United Nations, FAO

ABSTRACT

In light of an increasing focus on new demand-driven extension approaches that aim at accelerating the adoption of innovative technologies by smallholder farmers in developing countries, greater analysis is needed of the role of rural social networks and their impact on technology adoption. This paper contributes to this topic by reviewing selected studies on rural social networks and by outlining a research approach that combines social network analysis with econometric estimation techniques in one coherent framework to strengthen the study of technology adoption by smallholders. If applied, such a framework could help establish which network characteristics have the greatest impact on technology uptake, thereby lending support to and improving the design of new extension approaches.

Keywords: social network analysis, econometric modeling, adoption of innovations, farmer-to-farmer knowledge transfer

1. INTRODUCTION

> *Agricultural extension services, after a period of neglect, are now back on the development agenda, with a sense of excitement about many of the emerging institutional innovations. Clearly there is still much to do in bringing needed extension services to smallholders around the world, especially the poorest groups.*
>
> —*World Development Report 2008* (World Bank 2008, p. 175)

The adoption of agricultural production technologies and natural resource management (NRM) practices is an essential means of boosting agricultural productivity in developing economies. The importance of developing appropriately targeted technologies and promoting their subsequent uptake by farmers is a fundamental component of this process.

Over the past decade, rural extension and advisory services, established for precisely these purposes, have suffered a precipitous decline. Underinvestment, mismanagement, and top-down approaches in many developing countries have led to inefficient service provision, weak responsiveness to farmers' needs, and low rates of new technology adoption (World Bank 2008). Yet in spite of this situation, policymakers, donors, and development practitioners have come to recognize once again that extension and advisory services are critical to improving agricultural productivity, promoting agricultural development, and alleviating poverty.

This shift is driven by the emphasis that not only more investments but also new forms of extension are necessary. Novel approaches should consider farmers' specific needs, constraints, and capacities, while also encouraging farmers to act collectively and entrepreneurially (Birner and Anderson 2007). Empowering farmers in such a way is seen as a promising new avenue for stimulating the adoption of innovative technologies and NRM practices. Among new extension approaches discussed are partnerships between agricultural research institutes and farmers' organizations, a decentralization of extension services, new information and communication technologies, a general privatization of extension services, fee-for-service extension, and farmer field schools (see, e.g., Davis 2008; Rivera and Alex 2004). These approaches are being tested in different contexts, and their impact on innovation uptake is currently being evaluated by researchers and policymakers alike (e.g., Alex et al. 2004; Benin et al. 2007). Early evidence suggests that successes and failures are often very context specific and that more empirical evidence (and more rigorous evaluation tools) is needed.

One of the keys to success in the design of novel extension schemes is more intensive leveraging of farmers' social networks. Such networks are essentially informal communication channels, which farmers employ to receive and share information on new technologies or NRM practices. These networks are more complex than conventional "model farmer" approaches, which rely on a single actor as the focus of an information network, and they form spontaneously or organically among communities and subgroups within communities to bridge information gaps and reduce uncertainties about the application or local appropriateness of an innovation (Goyal 2005).

Such networks are particularly important to small-scale, resource-poor farmers, who tend to rely more on informal than formal sources of information, as well as to women farmers, whose information needs are often not addressed by formal extension services. Indeed, numerous empirical studies demonstrate that social networks do significantly influence the adoption decision of individual farmers (e.g., Baerenklau 2005; Conley and Udry 2001; Matuschke, Mishra, and Qaim 2007).

The effectiveness of farmers' networks in spreading information becomes vividly evident when one arrives as an outsider in a remote village of a developing country. Depending on the size of the village, within hours, everyone seems to know about the arrival of the outsider. The question that naturally emerges from this experience is how can these networks be leveraged to increase innovation adoption? Leveraging farmers' networks would not only be very time efficient, as this example illustrates;

it could also be very cost effective, in the sense that these social structures already exist and would not have to be constructed artificially, as in other extension approaches.

To design suitable policies that leverage social networks, further evidence on network construction and its impact on adoption behavior are warranted. Research to date has focused primarily on proof of concept—that is, whether the networks matter. Few studies push further by looking at the questions of (a) *Which* network *characteristics* matter? and (b) *How* do specific network characteristics matter? Moreover, few studies push the methodological frontier by examining the complexities of networks in relation to farmer and household characteristics, the landscape of rural institutions within which farmers operate, and other localized conditions, such as agro-ecology.

This paper intends to narrow these knowledge gaps by outlining an approach that combines social network analysis (SNA) with econometric estimation techniques in one coherent framework to improve the study of smallholder technology adoption. By joining SNA tools with standard econometric analysis, an impact assessment of network characteristics on economic outcomes becomes possible. Ultimately, this can strengthen the design of novel extension schemes and leverage social networks to promote more rapid adoption of technologies by smallholders.

The paper proceeds as follows. Section 2 provides an introduction to SNA concepts and presents selected applications in developing economies. Section 3 discusses different empirical studies that estimate the impact of social networks on economic outcomes. Section 4 outlines ways in which relational and attribute data may be combined. Section 5 offers a practical guide to what analysts should consider when collecting network data, followed by concluding remarks in Section 6. Throughout the paper, focus is placed on technical innovations that change the production function of a farmer or farm household, about which there is some uncertainty in terms of their successful application (Feder and Umali 1993).[1]

1 Other types of innovations are organizational innovations, which change the organization of production, and institutional innovations, which induce changes in policies and/or norms.

2. SOCIAL NETWORK ANALYSIS: CONCEPTS AND SELECTED APPLICATIONS

Social network analysis (SNA) is the formal analysis of relationships among agents, groups, or entities. Relationships may be formed at one level (e.g., between two persons) or at different levels (e.g., between a person and an institution). Relationships may provide channels for resource transfers of a material nature, such as money or information, or a nonmaterial nature, such as love or friendship.[2] The relationships that agents form may present opportunities and/or barriers to an individual's actions (Wasserman and Faust 2005).

The historic roots of SNA lie in sociology, social psychology, and anthropology. Mathematical graph theory has contributed significantly to the formalization of SNA. Since the 1980s, interest in social networks increased rapidly, due mainly to the recognition of social capital as an important factor of production (Durlauf 2002). Social capital theorists acknowledge that actions take place within a social context and are driven by social norms, rules, and obligations. Social capital, like human or physical capital, may be productive and beneficial to individuals. Unlike physical or human capital, however, social capital is not embodied in one person; rather it is in the relations a person has with other individuals and with the socioeconomic institutions within which that individual operates (Coleman 1999). Social networks are thus an expression of social capital.

Applications of SNA can be found in diverse areas, such as the analysis of occupational mobility, patent applications, community decision-making processes, social support systems, corporate interlocking, and the adoption and diffusion of innovations (Canter and Graf 2005; Wasserman and Faust 2005). Social network analysts, like sociologists or economists, have developed their own language to illustrate and explain network structures. The next subsection gives an overview of major SNA concepts. Focus is on concepts that are further employed in this paper. The second subsection discusses selected applications of SNA in developing economies.

2.1. Concepts of SNA

Social networks are composed of actors (*nodes*) and their contacts. The relationship that connects them is called a *tie*. Networks can be considered as a whole, or comprising all actors within a population of interest plus all of their ties (i.e., *full network* or *sociocentric network*). Networks that consider only selected actors (*ego*) and their ties to other actors (their *alters*) within the population of interest are called *ego networks*. The *network size* is the number of actors in a given network and is naturally bound to the size of the population of interest.[3]

Ties can be measured at different levels. *Binary measures,* which show whether a relationship exists between two actors, are the most commonly used measures. For example, if individual i relates to individual j, and j relates to i, then $X_{ij} = X_{ji} = 1$. Binary measures can also account for the direction of a relationship: If i relates to j but the reverse is not true, then $X_{ij} = 1$ and $X_{ji} = 0$. One can also measure the strength of ties between actors. *Grouped ordinal measures* allow for tie ordering by asking actors to rank all their relationships—for example, whether they like very much, dislike, or are indifferent to a specified actor (Hanneman and Riddle 2005). Other measures, such as the *interval measure* of ties, summarize not only whom an actor likes but also by how much—for example, individual i likes individual j two times more than individual i likes individual k.

Having explained the components of a network, we now turn to the questions of how networks can be described and measured. There are graphical and numerical techniques to do so, and both provide valuable diagnostics on resource flows and on how embedded actors are in their network. More embedded actors are assumed to be more influenced by their network's actions (Uzzi 2000). The following

2 Naturally, relationships may be overlapping and dynamic. For example, relationships within a family may be primarily characterized by love, but parents may also support their children with material means during their early stages of life. Material relationships may also reverse, when children support their parents at later stages of life.

3 For further details, see Scott (2000), Hanneman and Riddle (2005), and Wasserman and Faust (2005).

paragraphs discuss measures for ego networks, because these network types consider variations of individuals' embeddedness within their networks, which may provide good indicators for further econometric analyses (Hanneman and Riddle 2005). Because these measures are fairly complex, a hypothetical example of innovation adoption in rural India introduces these network measures here.

Assume we wish to illustrate how information on a recently introduced agricultural innovation, such as a row seeder, flows through networks in rural India. A farmer called Vijesh is randomly selected. To get a picture of his social neighborhood, he is asked to name the persons with whom he discusses important agricultural matters. In a follow-up question, Vijesh is requested to point out the persons within this network whom he regards as important in receiving or giving information about the row seeder. With a two-stage snowball method of interviewing, all contacts in Vijesh's network are approached and asked to answer similar questions on information exchanges about the innovation with the other actors in Vijesh's network. Table 1 illustrates the data in matrix form.

Table 1. Hypothetical egocentric network data on innovation information in rural India

	Vijesh	**Subramanian**	**Manoj**	**Sundeep**	**Prakash**	**Sachin**	**Jagdish**
Vijesh	0	1	0	1	1	1	0
Subramanian	0	0	1	0	0	0	0
Manoj	1	1	0	1	0	0	0
Sundeep	1	0	0	0	0	0	1
Prakash	0	0	0	1	0	1	1
Sachin	1	0	0	0	0	0	0
Jagdish	0	1	1	1	0	0	0

The first column of Table 1 shows Vijesh and all of his alters. The second row of the table illustrates who Vijesh gave information to regarding the row seeder—namely, Subramanian, Sundeep, Prakash, and Sachin. These ties are defined as *out-directed ties*. The second column shows from whom Vijesh received information—namely, Manoj, Sundeep, and Sachin. These ties are *in-directed ties*. All other rows and columns are read similarly, providing the complete picture of information exchange within Vijesh's social neighborhood.

Networks can be characterized graphically or by using specific network measures. Both options are discussed below. Figure 1 is an illustration, called a *sociogram*, of Vijesh's ego network. The arrows indicate the flow of information. For example, Vijesh gave Prakash information, but the opposite is not true. Graphic tools provide a quick and comprehensive understanding of complex relationships within a network—for example, Vijesh's position is central in the network. Figure 2 presents an overall picture of how information flows in Vijesh's direct neighborhood. It also shows who is well connected and who is fairly isolated (Sachin). Both graphs are simple, though graphical techniques do allow for a large number of illustrative options. For example, the strength of ties or different attributes of actors (e.g., who has adopted the innovation already) can be indicated by the size, color, or shading of the lines and nodes.

Figure 1. Graphic presentation of information network

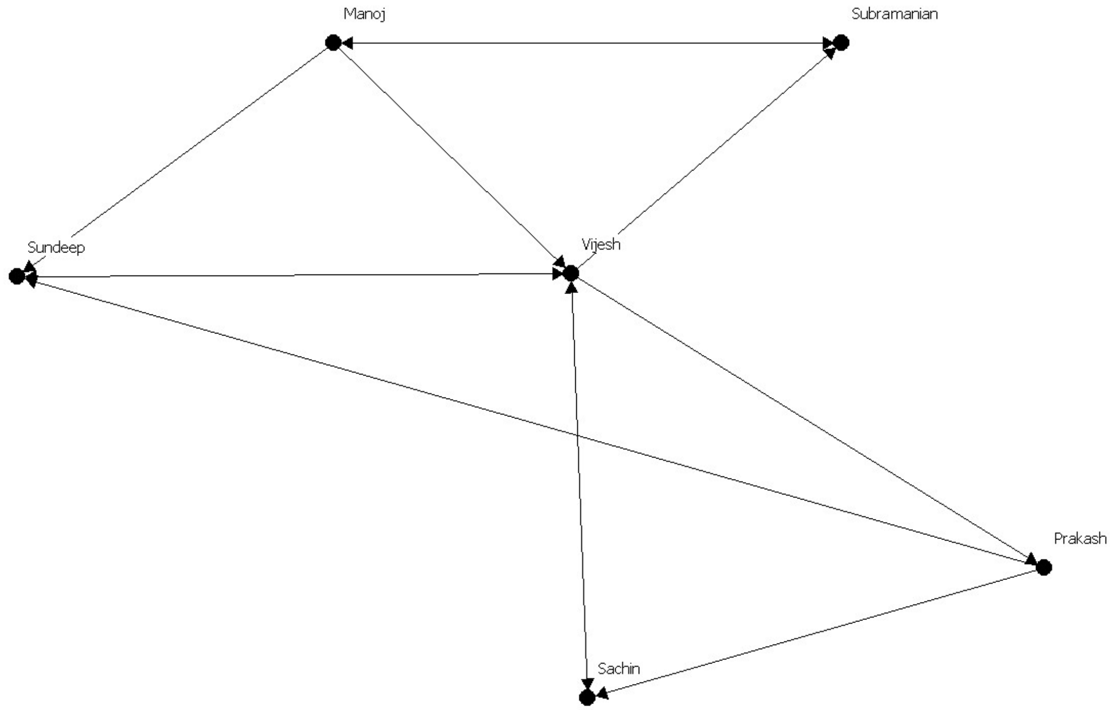
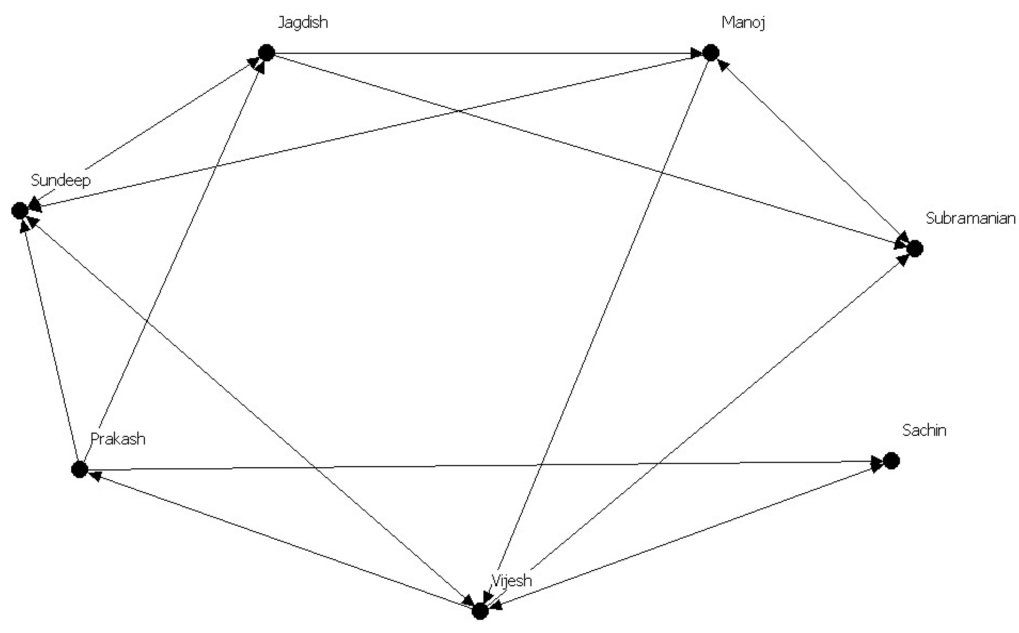


Figure 2. Graphic presentation of information flows within the social neighborhood



Specific network measures summarize relationship structures. For this hypothetical example, Ucinet, a statistical program for the analysis of social network data, was used to compute measures of Vijesh's ego network (Table 2).

Table 2. Network measures for Vijesh's hypothetical ego network

	Size	Ties	Pairs	Density	Reach Efficiency	Betweenness
Vijesh	5.00	5.00	20.00	25.00	35.29	70.00
Subramanian	3.00	2.00	6.00	33.33	46.15	16.67
Manoj	4.00	6.00	12.00	50.00	37.50	29.17
Sundeep	4.00	4.00	12.00	33.33	35.29	33.33
Prakash	4.00	6.00	12.00	50.00	40.00	8.33
Sachin	2.00	1.00	2.00	50.00	0.50	50.00
Jagdish	4.00	4.00	12.00	33.33	4.00	33.33

The *size* of Vijesh's information network is 5.00, or the number of persons with whom he exchanged information on the innovation. The number of *ties* is the number of ties among all actors within Vijesh's information network. The number of *pairs* describes the number of all possible ties within Vijesh's information network and is equal to 20.00, or 5 * 4. The *density* of the network is the number of actual ties over the number of possible pairs in the information network. In a dense network, an individual's network partners also communicate with each other, which implies that information may spread faster (Valente 1995). The density of Vijesh's information network is 25.00, which means 25 percent of all possible ties are actually present.

Reach efficiency is a measure that shows how many secondary contacts can be reached for each unit of primary contact (Hanneman and Riddle 2005). It comprises the number of all actors in the ego network that are within two directed steps of the ego. For example, Vijesh and Jagdish did not exchange information about the row seeder, but Vijesh could give a message or learn from Jagdish via Sundeep or Prakash. Thus, reach efficiency essentially measures a farmer's outreach to secondary, or "weak," ties, which could be significant in transmitting information (Hogset and Barrett 2007). Granovetter (1973), in his seminal paper on the strength of weak ties, found that novel information on job opportunities tends to flow through weak, rather than strong, ties.

Betweenness, the final measure illustrated in Table 2, describes the extent to which an actor is between all other actors on their path to one another (Valente 1995). Actors with a high betweeness may possibly play an important intermediary role and therefore could act as brokers in their network (Coulon 2005). Table 2 shows that 70 percent of all relationships within the information network pass via Vijesh, an indication of high betweenness.

In recent years, social network analysts have increasingly focused on the development of statistical tools for SNA. For example, regression analyses were introduced to identify the determinants of network positions or to estimate the likelihood that a relation exists between two or more actors *within* one network (Breiger 2004).4 Though useful in their own right, these statistical tools do not accommodate for comparisons across different ego networks in an approach that combines SNA with econometric estimation techniques (as will be proposed in Section 4).

4 Hanneman and Riddle (2005) provide a detailed overview of statistical tools in SNA.

2.2. Selected Applications of SNA

Social network analysis has been widely applied in numerous areas of academic inquiry. Examples include the analysis of innovator networks in eastern Germany and Italy (Balconi, Breschi, and Lissoni 2004; Canter and Graf 2005), men and women networks within an organization (Brass 1985), and interorganizational collaborations in biotechnology (Powell, Koput, and Smith-Doerr 1996). This section focuses on selected studies that applied SNA in the context of developing-country agriculture.

Crona and Bodin (2006) analyzed the example of a community in a coastal seascape in Kenya that faced an overexploitation of fish resources and a lack of collective action to counter this problem. Their hypothesis was that this lack of collective action was due to different structures of communication networks. Using information on a full social network (i.e., a village census) and SNA techniques (i.e., a sociogram of the strength of relationships between different occupational groups), the authors established that communication about natural resource management (NRM) occurred within occupational networks: Fishermen, who used the same gear type, discussed NRM techniques, but other occupational groups in the community did not. Therefore, network structures within the community hindered information transfers and, consequently, collective community action. The authors concluded that homogeneity within groups may lead to faster knowledge transfers. Yet, if a group is too homogeneous, knowledge may not spread, because it is inaccessible by outside sources. The authors also looked at positions of influential leaders within groups and found that the characteristics of these leaders were essential in coordinating successful group action.

Overall, the authors showed that SNA offers valuable tools for identifying and illustrating positions within groups and may therefore be a useful method for analyzing collective actions and the bottlenecks to it. However, in looking only at the structure of networks, the study by Crona and Bodin (2006) did not establish the actual impact of network characteristics on NRM practices in the community. Moreover, they did not control for other characteristics, such as those found at the local level, that may influence adoption. Yet this information is particularly relevant for policy targeting, because some characteristics may be more significant than others.

Darr and Pretzsch (2006, 2007) presented another study on social networks based on full network data for two rural communities in Kenya and Ethiopia. In light of the fact that farmers' groups are increasingly favored by extension services as a means of innovation promotion, the authors aimed to determine what influence group characteristics have on the innovativeness of individual farmers. To do this, they analyzed the adoption of intercropping and farm woodlot innovations by members of different farmers' groups. Individual innovativeness scores were calculated based on the number and complexity of innovations adopted. Using SNA techniques, the authors calculated measures such as network density and the number of out-directed ties for farmers' groups. In a final step, these measures—in addition to other group measures and individual characteristics—were regressed on innovativeness scores using a general linear model.

The authors found that structural network variables had a significant and positive effect on the innovativeness of group members and, consequently, the group. Interestingly, the authors also established that in the case of farm woodlot adoption, more innovative groups tended to be characterized by a top-down leadership with powerful management boards and weak member participation in offsetting up the group's agenda. The authors concluded, in contrast to the Crona and Bodin study, that cohesive groups with active exchanges of information and collaboration among its members led to higher diffusion of innovations. However, even though the study by Darr and Pretzsch is informative and based on an extensive data set, it does not account for the fact that farmers may actually group together *because* they are innovative. If this is the case, a simultaneity problem, as described in Manski (1995), could evolve, and derived results may be biased and inconsistent.

Another study (Raini, Zebnitz, and Hoffmann 2005) aimed to determine why the adoption of integrated pest management (IPM) techniques within Kenya's tomato sector is low. To answer this question, the authors looked at social networks of IPM stakeholders—farmers, government, agrodealers, extension agents—in the tomato sector. Using illustrations of networks, as well as network measures, the authors established that the network of stakeholders exhibited a very low density. They concluded that for this reason, information flowed slowly through the network, thereby inhibiting a faster diffusion of IPM techniques. By using SNA, the authors were able to illustrate bottlenecks to information flows, but they were unable to include other factors that may have significantly influenced adoption. Moreover, the study did not estimate the actual impact of social networks on the adoption of IPM. Therefore, the results are indicative rather than conclusive.

Finally, a study by Hartwich et al. (2007) aimed to compare how different knowledge-management schemes influence innovation behavior of smallholder farmers in Bolivia. The authors compared a top-down approach with a more bottom-up approach that promotes innovation via a network of technology providers, farmers, and private sector agents. Using a case study approach and collecting quantitative and qualitative data from farmers and their information providers, the authors found that farmers who participated in network-related extension schemes had higher adoption rates of modern technologies than did farmers who participated in more traditional extension systems. The study is one of only a few that employs network characteristics in an estimation framework. Using Tobit models, the authors estimated the impact of a farmer's connectedness on adoption behavior. Yet farmers' networks were defined somewhat widely: They included not only other farmers, but also researchers, extension agents, nongovernmental organizations, input buyers, and transporters. Such a wide definition makes it difficult to interpret estimation results and to pin down the actual impact of each network agent on adoption. As a consequence, the design of policies that aim to stimulate adoption becomes very complex.

3. ECONOMIC APPROACHES TO NETWORK ANALYSIS: AN OVERVIEW

Like social network analysts, economists have long recognized that individual decision-making processes and economic outcomes are correlated with the behavior of other agents (see Brock and Durlauf 2001 for a review of interaction-based models). Coleman, Katz, and Menzel (1957), for example, showed that doctors in professional medical networks were more likely to prescribe a new drug than doctors who were not in such networks.

Economists, however, often paid less attention to relationships and structural network measures because of the inherent difficulties in combining relational and attribute data. Only recently have the schools of social network analysis and economics started drawing from each other to consider the impact that other agents or groups may have on individual behavior. This is particularly true in the study of agriculture, where socioeconomic researchers, concerned with barriers to the adoption of modern production technologies and NRM practices, have become increasingly aware of the importance of social networks in individual learning processes.

The underlying assumption that most studies in this vein have made is that farmers learn about the characteristics and risks of an innovation from three main sources: their own experience, the experience of others, and their interactions with formal sources, such as extension agents or seed dealers. Yet quantification of these learning effects has only been pursued in the past two decades. The approaches taken by empirical studies are diverse, and they are also becoming more and more complex, reflecting more accurately the realities that rural decision makers face.

3.1. Selected Economic Approaches to Network Analysis

This section reviews selected studies that incorporated learning from others into their analytical framework and discusses the empirical problems these analyses faced. The review is restricted to studies that considered rural areas in developing economies. The studies are ordered by the proxies that were used for estimating network effects. Whereas early studies used village-level variables as network proxies, later studies focused on the number of ties in smaller and individual social networks. Only in recent years have researchers started to include network characteristics to improve the estimation of social network impacts.

3.1.1. Village-Level Network Variables

Foster and Rosenzweig (1995) were among the first scholars who incorporated the learning behavior of farmers in an econometric framework. The basis of their analysis was a target-input model. With such a modeling framework, it is assumed that farmers do not know the optimal input level associated with a new technology. After the harvest, farmers update their beliefs on the optimal input level by combining the experiences of the past season with the experiences from all other previous periods. The inclusion of prior experiences and/or beliefs into a modeling framework is defined as a Bayesian updating method. The study also assumed that farmers update their beliefs by learning from experiences of neighboring farmers.

Using a panel data set on the adoption of high-yielding varieties (HYVs) of rice and wheat during India’s Green Revolution, the authors established that learning from one’s own experiences was significant. They also found that learning from neighbors mattered and led to increases in the profitability of farming operations. By using farm profits associated with HYVs as a dependent variable, the authors showed that farmers actually learned from each other and did not just mimic each other because of similar circumstances.

Using the same data set as Foster and Rosenzweig, Munshi (2004) added an interesting nuance to the analysis of learning behavior by analyzing the diffusion of HYVs for wheat and rice farmers separately. Wheat growers were characterized as being a more homogenous group than rice growers. Estimating the effect of social learning on the acreage allocated to HYVs, the author found that more social learning took place in the homogeneous population. Munshi concluded that the reason was that in heterogeneous populations, farmers may not rely on learning from the experiences of their neighbors, because the neighbors' characteristics that led to the outcome can vary widely and may therefore be hard to observe for the individual. This could be one reason that apparently attractive innovations remain unpopular. The overall conclusion that Munshi drew was that in highly heterogeneous populations, social learning may simply break down. The network proxy that both Munshi (2004) and Foster and Rosenzweig (1995) used was the average village adoption rate of HYVs.

Numerous scholars followed the same approach to estimate the impact of farmers' networks on adoption behavior. Among them are Pomp and Burger (1995), who studied the adoption of cocoa among Indonesian smallholders, and Isham (2002), who considered the adoption of fertilizer by Tanzanian farmers. The underlying assumptions of this approach are that all farmers in a village influence an individual farmer in the same way and that learning takes place along geographical lines. These assumptions may not necessarily be correct, particularly in villages characterized by social or cultural stratification. In such cases, there may be heterogeneity with respect to resource flows, trust, norms, or other institutional factors. Consequently, farmers may be more influenced by smaller networks formed on a personal basis than by the whole village (De Weerdt and Dercon 2006). Moreover, in using such a proxy, it becomes impossible to disentangle effects that are due to network behavior from those that are due to correlated unobservable variables at the village level. This may lead to biases in estimation results.

3.1.2 Social Network Variables

One reason many of these studies considered village-level variables instead of variables describing individual networks could be that specific data on farmers' attributes *and* relations are required. Yet these data are normally not collected in a standard household survey. In recent years, several studies addressed this problem by using specifically collected network data sets. Examples include informal insurance networks among farmers (e.g., De Weerdt and Dercon 2006; Hoddinott, Dercon, and Krishnan 2005), information networks (e.g. Bandiera and Rasul 2006), and informal labor-sharing networks employed in times of labor shortages (Krishnan and Sciubba 2007).

Conley and Udry (2000, 2001) produced one of the first network studies that combined data on farmers' attributes with data on their relations. Based on an input-target model with Bayesian updating and using comprehensive longitudinal data on pineapple cultivation in Ghana, the authors showed that farmers did not learn from all farmers in the village; rather, learning took place within smaller networks. More specifically, in estimating the impact of network partners on changes in input use, the authors concluded that (a) farmers' input usage depended on the input usage of their network partners, and (b) this impact was larger for inexperienced farmers. The authors also found that farmers responded more to network partners that are similar to them, thereby supporting the point made by Munshi (2004).

These studies raise the issue of endogenous, exogenous, and correlated effects and the techniques needed to control for these effects in an analysis of how social networks influence individual outcomes. We describe each below.

Endogenous effects describe the impact that the adoption decision of a network partner has on the individual. A potential problem in this context is simultaneity: The mean behavior of the group influences the individual, who in turn influences the group. Manski (1993) coined this as the *reflection problem*. Exogenous effects describe the idea that an individual may be part of a group because of characteristics of the group itself. For example, coming back to our hypothetical case, farmers with a row seeder may wish to group with other farmers who own a row seeder in order to share information on how to operate and repair the instrument. If this is the case, then group membership itself becomes endogenous. Correlated effects describe the idea that network members may be similar in their choices due to similar individual characteristics or because they face the same local, institutional, or cultural conditions (Manski 1993).

Failing to control for these three different effects may lead to inference biases, because the analyst is unable to disentangle impacts of social networks on adoption behavior from the impacts of local conditions. Conley and Udry (2001) controlled for these effects because their study was based on extensive information on individual and village characteristics, as well as group composition. Moreover, the authors considered intertemporal input choices to account for endogenous effects.

Bandiera and Rasul (2006) provided further evidence on the impact of social networks by using a case study on sunflower seed adoption in Mozambique. Also based on a target-input model with Bayesian updating, the study used data on the first year of the adoption. The authors constructed a social network proxy based on the number of sunflower adopters in an individual's network of friends and family. Using a linear probability model, the study found that networks play a significant role in adoption decisions. Moreover, networks of family and friends were more relevant than were other social cohorts (e.g., religious groups) for the adoption decision. The authors established further that (a) more informed farmers relied less on their social network in the decision-making process, and (b) the relationship between the adoption choice and the number of adopters in the network was shaped as an inverse U. This latter point means that as the number of adopters in a network rose, farmers delayed their adoption decision in order to observe other adopters in their network.

To encounter the potential estimation problems described above, the study used village-fixed effects to control for village unobservables. To control for unobservables in an individual context, the authors added variables of individual characteristics. Yet, the study did not address the reflection problem. Instead, Bandiera and Rasul wrote that their estimation procedures only informed whether the adoption decisions within social networks were correlated.

Two other studies that showed the significance of small social networks and that employed the number of ties as a network proxy are Miguel and Kremer (2003) and Behrman, Kohler, and Cotts Watkins (2002). The former study examined the impact of social networks among parents on the adoption of deworming medicine in schools of rural Kenya. This study, also based on an analytical framework of Bayesian updating and an extensive data set, used the number of social contacts whose children had received the drug as a network variable. Using probit estimations, the study showed that social networks inhibited adoption behavior. In the case of deworming medicine, more adopting links led to fast disadoption of the drug, because the drug did not offer sustainable protection.

The latter study, by Behrman, Kohler, and Cotts Watkins (2002), added to the empirical literature on social networks by looking at the adoption of contraceptives in rural Kenya. Using a fixed-effect logit model and employing the number of network partners with whom a woman chatted about using contraceptives, the authors found that social networks are important mechanisms in providing information on innovations—particularly on sensitive innovations like contraceptives.

Both studies considered the number of network ties rather than network characteristics, allowing the authors to establish whether social networks influence an individual's adoption behavior. However, this approach did not permit the authors to determine which network characteristics influence adoption.

3.1.3. Including the Characteristics of Social Networks

Whereas the studies previously described mainly considered the number of ties in order to study network impacts, few studies to date have pushed further to incorporate network structures into the analysis. An exception is a study by Krishnan and Sciubba (2007), which emphasized that the number of ties and the architecture of the network must be considered when evaluating the impact of social networks on economic outcomes. Looking at the number of ties only may lead to biased results. Using extensive panel data from the Ethiopia Rural Household Survey, conducted in 1994, the authors modified a model by Jackson and Wolinsky (1996) and looked at the formation of labor-sharing networks in rural areas. Krishnan and Sciubba defined two kinds of labor-sharing networks: symmetric and asymmetric labor-sharing agreements. Whereas the former consisted of farmers who had approximately the same number of ties, the latter were formed between farmers with different numbers of ties. The authors found that symmetric networks were stable and homogenous, because they comprised farmers with equal labor endowments. Asymmetric networks, on the other hand, consisted of farmers with different labor endowments. The authors differentiated asymmetric networks further into hubs and spokes: Hubs were farmers who had more links than their network partners, whereas spokes were farmers that had fewer links as compared with their network partners.

The variables on network architecture and the number of links were then used in a Cobb-Douglas production function estimation to test whether labor-sharing networks had an impact on individual productivity. To control for network endogeneity (i.e., the possibility that more productive farmers club together), the authors used an instrumental variable approach. Instruments that did not directly influence productivity, but that were assumed to influence network formation, were chosen. Examples of such instruments are the number of blood relatives in the village and whether or not the farmer was born in the village. The authors claimed that these variables account for the embeddedness of a farmer and his or her relative role and position in the village. In this case, measures of SNA could have been potentially better instruments to account for the embeddedness of a farmer in his or her network. The authors' regression results established that the number of links had an impact on productivity but that the architecture of the network had an even bigger impact on the economic outcome.

One of the underlying assumptions of this study was that individuals rationally chose whether to form or delete a tie with other agents in their network. This assumption may be applicable to labor-sharing networks, which are formed on an annual basis in the growing season, but it is doubtful whether this assumption holds for other network types, such as information or insurance networks. Insurance networks, for example, can be assumed to be based on long-standing (family) ties characterized by mutual trust. In these types of arrangements, the rational deletion of a tie may not be feasible.

De Weerdt and Dercon (2006) considered insurance networks and tested the role of networks in insuring health shocks. Based on a full census of all insurance networks in a rural village in Tanzania, the authors estimated the impact of health shocks and changes in network consumption on changes in individual consumption. Networks in the village were defined using SNA measures. The networks were assumed to be present if households had a geodesic distance of one or two, which means that households either were directly connected or were connected via one other household in the village. To account for the endogeneity of changes in network consumption, the authors used an instrumental variable approach. Instruments used were changes in the demographic characteristics of the network members over the study period. The authors concluded that networks did have a significant impact on changes in individual nonfood consumption but not on changes in food consumption. They also concluded that endogeneity is relevant in insurance networks and has to be controlled for.

Summing up, this section discussed economic approaches to network analysis. It demonstrated that even though the impact of networks on economic outcomes had been acknowledged, the integration of network measures into econometric analyses has been increasingly pursued only in the past two decades. Village-level variables were often used as proxies for network effects—an approach that was contested by more recent studies, which showed that resources flow through smaller networks that are not necessarily based on geographic proximity. Few studies went further to introduce not only the number of social network ties but also network architecture into their impact analysis.

4. COMBINING SNA AND ECONOMIC APPROACHES: POTENTIAL IDEAS

This section presents a simple analytical framework for the adoption of modern innovations that incorporate social learning. This is followed by a general discussion of potential estimation approaches. A case study is then presented that applies one of the estimation approaches in order to analyze the adoption of hybrid seeds in India.

4.1. Analytical Framework

The simple framework presented here is based on Brock and Durlauf (2001) and Durlauf (2001). Assume that individual *i* is faced with the decision of adopting an innovation. The choice is described by indicator variable c_i. Each individual strives to maximize his or her utility function, which can be expressed as

$$c_i = \mathrm{Max} U(c, Z_i, \varepsilon_i)\,. \tag{1}$$

The utility function (1) is a function of the innovation choice *c*, a vector of individual characteristics Z_i, and a vector of unobserved individual characteristics ε_i, the latter of which are known to the individual but unknown to the analyst. Interactions with others can be incorporated into the *Z*-vector. In this way, the innovation choice becomes directly dependent on the behavior of others. In line with Manski (1995), the analyst should account for three effects: correlated, exogenous, and endogenous effects. With this in mind, X_i is defined as a vector of individual characteristics. $Y_{n(i)}$ is a vector of variables that are common to all members of individual *i*'s network *n*(*i*). Endogenous effects μ_i are expressed as

$$\mu_i\,(c_{n(i)}\,|\,F_i)\ , \tag{2}$$

where μ_i is a function of the individual's beliefs on the adoption choices made by his or her network partners $c_{n(i)}$, given a set of information F_i available to the individual. Accordingly, the utility function (1) is rewritten as

$$c_i = \mathrm{Max} U\left[c, X_i, Y_{n(i)}, \mu_i(c_{n(i)}|F_i), \varepsilon_i\right]. \tag{3}$$

In a next step, it is assumed that the utility function (3) is additively separable, such that

$$U\left[c, Xi, Y_{n(i)}, \mu_i(c_{n(i)}|F_i), \varepsilon_i\right] = u(c, X_i, Y_{n(i)}) + S\left[c_i, X_i, Y_{n(i)}, \mu_i(c_{n(i)}|F_i)\right] + \varepsilon_i(c)\,, \tag{4}$$

where the first term on the left side expresses the private deterministic utility, the second term denotes the social deterministic utility, and the third term signifies the private random utility. The difference to standard choice models is the social deterministic utility function, which implicitly includes the choices of the individual's network partners.

In a further step, Durlauf (2001) assumed that the social utility function follows a specified quadratic function:

$$S(\bullet) = -E\ \sum_{i \neq j}\frac{J_{ij}}{2}\left[c - E(c_j|F_i)\right]^2 . \tag{5}$$

The term $J_{ij}/2$ measures the interaction weight of *i*'s and *j*'s choices. Most modeling approaches assume that the interaction weight is positive, but this is not necessarily true (Brock and Durlauf 2001). If $J_{ij} > 0$, then individual *i* derives a greater utility from making the same choice that he or she perceives network partner *j* is making. If $J_{ii} < 0$, then individual *i* will make a different choice than the perceived choice of network partner *j*, because individual *i* would experience a lower utility otherwise. A frequently made assumption in modeling choices is that the private random utility is logistically distributed

$$\mu\left[\varepsilon_i(c) - \varepsilon_i(-c) \leqslant z\right] = \frac{1}{1+\exp(\beta_i X)} \quad \beta_i \geq 0. \tag{6}$$

Furthermore, it is assumed that the private deterministic utility function is linear:

$$u(\bullet) = h_i c + k_i\,, \tag{7}$$

where *h* is the slope and *k* is the intercept term. Terms *h* and *k* are chosen so that

$$h_i + k_i = (1, X_i, Y_{n(i)})\,. \tag{8}$$

With these assumptions, all elements of the model [equation (4)] are fully defined, and the model that includes social interactions is closed (Durlauf 2001). Individual *i* will adopt an innovation if the utility derived from adopting the innovation is greater than the utility derived from the status quo:

$$U\left[(1, X_i, Y_{n(i)}, \mu_i(c_{n(i)} | F_i)\right] > U\left[(0, X_i, Y_g, \mu_i(c_{n(i)} | F_i)\right]. \tag{9}$$

4.2. Potential Empirical Estimation Procedures

A number of empirical approaches are feasible when estimating the impact of social networks on economic outcomes. The choice of the most suitable approach depends on the analyst's specific needs and interests, the technological innovation to be analyzed, and the availability of data. Therefore, potential methods are briefly discussed here before a case study is presented that applies one particular approach.

For cross-sectional data, several estimation approaches could be applied. *Input-target models*, for example, are popular for network studies that focus on agricultural innovations in developing economies (see review in Section 3). A precondition for using these models is that the innovation in question be divisible. Input-target models assume that farmers are insecure about the use of a modern input. Through their own experiences and network partners, farmers learn about optimal input levels. Input-target models can be estimated using ordinary least-squares estimations. If input use is censored—that is, if some farmers do not use the input in question at all—a Tobit estimation approach is viable.

Binary choice models are generally based on maximum likelihood estimations, and the dependent variable measures whether a farmer adopted an innovation. A proxy for social networks is added as an independent variable to estimate the impact of networks on adoption decisions.

Production function models measure the impact of social networks on outputs using a Cobb-Douglas production function (see Krishnan and Sciubba 2007). By using output as a dependent variable, one would also exclude the possibility of mimicry, as explained in Foster and Rosenzweig (1995).

If panel or time-series data are available to the analyst, other estimation approaches would be possible. Duration models, for example, could be estimated that determine not only the reasons a farmer adopted but also the timing of the adoption decision. Other possible approaches include treatment models that look at adopters and nonadopters of an innovation over a longer period. Brock and Durlauf (2001), in their study on interaction models, offered an excellent overview of potential models to be used when time-series or panel data are available.

4.3. Case Study: Analyzing the Adoption of Modern Seeds

This case study is based on Matuschke and Qaim (2007), in which the authors employed a binary choice model to analyze the impact of social networks on the adoption of hybrid wheat in India. It should be stated up front that this case study is not a perfect example of how to incorporate SNA and econometric estimation techniques into one coherent framework. Nonetheless, the study presents a valuable step in the right direction and helps further outline the requirements and potentials.

Data collection for the study took place in 2004 in three districts of the Indian state of Maharashtra. Using stratified random sampling methods, the authors collected information on 284 wheat farmers—87 adopters and 197 nonadopters of hybrid wheat. Next to information on household characteristics, specific information on individual networks was gathered. Farmers were asked about their information network, and individual characteristics of network members were requested. The authors did not inquire about the ties between network members themselves; therefore, direct network measures, as illustrated in Section 2, could not be retrieved from the data. The individual data collection was complemented by village-level surveys. Based on a simple framework of learning, the authors estimated the following equation:

$$a_{iv} = \beta X_{iv} + \gamma a_{n(i)} + \delta X_{n(i)} + \kappa G_v + e_{iv} \, . \tag{10}$$

The adoption decision of individual i in village v was assumed to depend on the farm household's characteristics X, as well as on the adoption decision of the social network partners $a_{n(i)}$ and their characteristics $X_{n(i)}$. The authors also included village-fixed effects G_v to control for unobservables at the village level that may influence adoption. Term e_{iv} was defined as an error term and was assumed to be normally distributed and uncorrelated with any of the variables. Adoption of hybrid wheat was modeled as a binary choice problem $a_{iv}\{0,1\}$.

After setting up their model, the authors proceeded in three steps. In their first estimation (Model 1), they only included the village adoption rate as a social network proxy, which resembles early network models presented in Section 3. In the second estimation (Model 2), the share of adopters in a farmer's individual network was used as a network proxy. In addition, village-fixed effects were added to control for village unobservables. In the third estimation (Model 3), group characteristics were added to account for the exogenous network characteristics. Table 3 displays the probit regression results.

Table 3. Analyzing the adoption of hybrid wheat (probit model estimation results)

Explanatory variable	*Model 1*		*Model 2*		*Model 3*	
	Coefficient	*z*-value	Coefficient	*z*-value	Coefficient	*z*-value
Individual characteristics						
Education (in years)	1.52E-03	0.25	-5.37E-04	-0.08	-4.54E-04	-0.07
Experience (in years)	-2.25E-03	-0.86	-1.51E-03	-0.49	-1.53E-03	-0.48
Farm size (in acres)	3.77E-03	1.23	3.82E-03	1.10	4.33E-03	1.18
Irrigation (share of irrigated land in total land)	-6.47E-03	-0.07	0.03	0.23	0.02	0.21
Household expenditures (in 000 Rupees)	0.01***	2.70	0.01*	1.77	0.01*	1.79
Information constraint (dummy, 1 = yes)	-0.25***	-3.36	-0.26***	-2.95	-0.25***	-2.89
Credit constraint (dummy, 1 = yes)	-0.03	-0.60	-0.04	-0.69	-0.05	-0.82
Association membership (dummy, 1 = yes)	0.11*	1.87	0.11*	1.77	0.11*	1.80
Village and regional characteristics						
Village adoption rate 2003/2004 (in %)	0.15	0.51				
Distance to input dealer (in km)			0.01	1.39	7.22E-03	1.42
Distance to output market (in km)			-2.68E-03	-0.72	-3.07E-03	-0.81
Number of households in the village			2.30E-04	1.67	2.43E-04*	1.79
Average soil quality[a]			0.05	0.64	0.05	0.58
Poor soil quality[a]			0.06	0.71	0.07	0.72
Yavatmal[b]	0.08	0.94	0.11	1.10	0.14	1.33
Aurangabad[b]	0.02	0.32	0.13	1.36	0.15	1.47
Network characteristics						
Share of adopting network members (NMs) (in %)			0.38***	4.13	0.39***	4.13
Age of NM					1.26E-03	0.37
Caste of NM[c]					0.08	1.00
Farm size of NM (in acres)					-1.70E-03	-0.56
Communication with NM (times per month)					-3.01E-04	-0.08
Distance to NM (in m)					4.90E-04	0.05
Log – likelihood	-150.81		131.40		-130.66	
Pseudo (R^2)	0.13		0.20		0.20	

Source: Matuschke and Qaim (2007)

Note: The number of observations is n = 282 in all models. Coefficients can be directly interpreted as marginal effects on the probability to adopt (evaluated at sample means). Standard errors are robust.

*, **, *** Coefficients are significantly different from 0 at the 90%, 95%, and 99% confidence level, respectively.

[a] Reference variable is high soil quality. [b] Reference variable is Nashik district. [c] Share of network members with the same caste as the individual (in %).

The first column of the table displays the explanatory variables used in the regression analysis (for a detailed discussion, see Matuschke and Qaim 2007). Variables are divided into three sections: individual characteristics, village and regional characteristics, and network characteristics. Among the individual characteristics are two dummy variables that describe the individual's perceived access to information and credit. Association membership is a dummy that captures whether the individual was a member in one or more village associations. Among village characteristics, dummies were added to control for districts and soil quality characteristics. The network proxies are the share of adopting network members, as well as exogenous characteristics of the network, such as average age or education of network members.

Matuschke and Qaim (2007) drew four main conclusions from their analysis. First, village adoption rates may not be suitable proxies for network effects, because they may underestimate the impact of social networks on adoption behavior. Model 1 shows that the village adoption rate is insignificant. Second, individual social networks matter. Model 2 shows that the share of adopting network members is highly significant. Third, exogenous network characteristics are insignificant in this regression setup (Model 3). Finally, information constraints are important barriers to adoption in all regression setups.

The authors also aimed to address the problem of simultaneity—that is, the fact that a farmer is influenced by his or her group and at the same time influences the group. To do this, the authors looked more closely at the timing of the adoption decisions within the networks. Having information on when the network partners had adopted (i.e., before, after, at the same time as the individual), the authors constructed a new network variable that included only those members who had adopted with a lag in time. In this way, the authors excluded the possibility that a farmer influenced and was influenced by the network at the same time. The probit regression results, which include the new network proxy, confirmed the overall conclusions. An instrumental variable approach may have been more suitable in this context, but there was a lack of suitable instruments.

4.4. Further Extensions and Requirements

The case study in the previous section serves as a good example on how SNA measures could further improve the econometric analysis. If the number of adopting network members depends on network structure (e.g., density or betweenness), then SNA measures could prove to be valuable instruments. In this case, a two-step estimation procedure could be proposed. In the first step, the impact of network structure on the share of adopting network members would be estimated. Based on this, in a second step, the impact of adopting network members on the adoption of hybrid wheat would be estimated.

As a result, two analytical outcomes would emerge. First, the social network "blackbox" would be open, in that it would become possible to estimate which network structures matter and how they influence adoption. Second, the simultaneity problem would be addressed in a more efficient way.

Furthermore, SNA measures could add detail to the introduction of unique network actors, such as formal extension services. Specifically, extension agents could be introduced into the regression analysis as an additional actor in the network, or they could be proxied by such measures as the frequency or quality of contact. By including formal extension services into the analysis, the landscape of rural institutions within which farmers operate could be accounted for more accurately.

5. A PRACTICAL GUIDE TO SOCIAL NETWORK DATA COLLECTION

Section 4 showed how studies of technology adoption could be improved by collecting more accurate social network data. Indeed, data collection is one of the most interesting and debated topics in discussions of how to combine relational and attribute data. The two main questions an analyst might ask are what kind of data to collect and how to collect data on an individual's contacts. We address these questions below.

An analyst can approach the first question by choosing to collect census data or data from randomly selected egos. Census data collected for all individuals within a selected population would offer a complete picture of all relationships and allow for the calculation of complex network measures. Yet census data collection is often an expensive and time-intensive undertaking (Hanneman and Riddle 2005). It may also be difficult to collect such data if the contacts named are not consistently resident within the geographic boundaries of the selected population.

One solution to this problem could be to select a small sample population (e.g., 50 tomato-growing farmers). In this case, however, the estimation results would be valid only in the case of the small census population, and it might be inappropriate to draw wider generalizations for a larger population. Consequently, policy recommendations on how to leverage social networks may be misleading. For this reason, we advocate the collection of ego network data. Data collected from randomly selected egos provide an informative picture on the embeddedness of an individual within his or her network and allow for the application of econometric techniques that lead to unbiased estimation results. In addition, egos can be selected across a wide range of individuals and/or locations within the population of interest, thereby ensuring a greater representativeness and applicability of the analysis.

Having decided on the type of data to be collected, the analyst must then choose how to phrase the question that generates information on an individual's contacts (i.e., the name generator). There are several potential pitfalls to consider. First, one must decide whether the individual should be restricted in the number of contacts to be named. If the individual is not restricted in his or her choice, the information provided will theoretically cover all strong and weak ties. In reality, however, individuals may not be able to come up with all their contacts at the time of the interview. To assist individuals in this case, at the end of each household questionnaire, a list of persons in the village (or any other geographic boundary) may be given so that the respondent can select contacts in addition to those already named (Wasserman and Faust 2005).

Nonetheless, even though unrestricted contacts are preferable from a theoretical point of view, individuals may not be able to name a large number of contacts, because keeping up good ties is resource intensive (Cox and Fafchamps 2006). For this reason, there may be a natural limit to the number of possible contacts—smaller for more resource-intensive networks (e.g., credit networks) and larger for less resource-intensive networks (e.g., information networks). Sufficient pretesting of household questionnaires should provide the analyst with a good idea of the natural boundary of networks and the ties within these networks.

The second pitfall is related to the phrasing of the question on contacts itself. For example, as Scott (2000) pointed out, asking individuals for their close friends is difficult, because close friendship may be perceived entirely differently by respondents. If the analyst decides to classify closeness as the frequency of contacts, then a definition of friendship is imposed that may be artificial to the respondent. This, in turn, may lead to an inappropriate naming of contact partners.

Another option is whether real or potential ties should be studied (Santos and Barrett 2007). A question on a real tie would, for example, be phrased, "Who *did* you change information with?" A question on a potential tie would be phrased, "Who *would* you change information with?" Real links are stated to be more appropriate when studying past behavior, such as information exchanges, whereas potential links may be more interesting to explore in the case of future behavior, such as insurance networks. The analyst should be aware of these fine differences and should decide in accordance with the objectives of the study.

A third pitfall is associated with the collection of information on network partners. To save time, the analyst could ask the respondent to provide information on the names generated. However, as studies by Behrman, Kohler, and Cotts Watkins (2002) and Hogset and Barrett (2007) showed, the information given may be incorrect. Hogset and Barrett, in their study on the adoption of natural resource management (NRM) techniques in Kenya, established that respondents do not exactly know about the behavior of their peers. For example, if farmers were adopters of one NRM technique, they were more likely to assume that their contacts had adopted the same technique. Comparing regression results for self-reported and cross-checked peer behavior, the authors found that estimation results and statistical interferences based on self-reported behavior were highly misleading. A secondary snowballing methodology could avoid this problem. Within this secondary methodology, generated network partners are personally contacted, and their information is directly retrieved. However, the counterargument to this approach is that individual adoption decisions are not based on the actual behavior of network partners, but only on behavioral perceptions (Hoddinott and Schiffer, pers. comm., 2008). If the analyst takes this standpoint, secondary snowballing becomes redundant.

Finally, data collection for census and ego networks may suffer from missing links (Santos and Barrett 2007). Missing data may arise if contacts generated by individuals live outside the geographical area in which the survey takes place. There may also be persons who are very hard to reach for the analyst or who are simply unwilling to participate in a survey. All of the studies discussed in Section 3 suffer from this problem, and missing links may make up to one-third of the names generated by individuals (Krishnan and Sciubba 2007). Missing links may lead to a systematic bias in the regression results, which is hard to correct for. To avoid this problem, Conley and Udry (2001), as well as Santos and Barrett (2007), applied a random matching technique, in which individuals were confronted with a random set of persons within their village and were asked whether they would be willing to form a relationship with these individuals. When using random matching, full information on egos and their alters becomes available. Moreover, data collection is simplified because the random matches may be selected within the same geographical boundary. The possibility that weak ties are neglected is also reduced when applying random matching (Santos and Barrett 2007).

Using Monte Carlo analysis, a study by Santos and Barrett (2007) compared random matching and secondary snowballing techniques by estimating whether a link would be formed between two persons. Based on a sample of 120 pastoralists in rural Ethiopia, they found that random matching was superior to secondary snowballing. Yet random matching may be more applicable in the case of potential rather than real ties. As an example, if the analyst wants to study past network behavior, randomly matched links will not provide an accurate picture of exchanged information.

Taking into consideration some of the aspects discussed above, Appendix presents and discusses a mock-up survey on how one could potentially retrieve data on information ego networks. This mock-up survey is based on the hypothetical example of innovation adoption in rural India introduced in Section 2.

6. CONCLUSIONS

In the light of an increasing interest in novel approaches to extension and advisory service provision, this paper provides a background on how social networks can be analyzed and how relational and attribute data may be combined to quantify the effect of social networks on smallholder technology adoption. The paper reviews selected studies on rural social networks and describes a research approach that combines social network analysis and econometric estimation techniques into a single framework that might strengthen the study of technology adoption by smallholders. The application of such a framework would contribute to the current discussion of modern extension approaches by providing new insight into which network characteristics have the most significant impact on adoption. Having more precise knowledge on the impact of social networks is a significant contribution to informed and effective policy formulation and implementation.

Because this paper is not based on a data set that was specifically designed to illustrate this framework, concluding comments can only outline potential benefits and beneficiaries. Potential beneficiaries may be national-level researchers in the social sciences who, through such an analysis, could obtain a better understanding of technology adoption decisions and could enrich their impact assessment studies by adding a further nuance and complexity. The results from a social network study could also provide extension agencies and agents with a new set of diagnostic tools that could fit well with the new extension emphasis on participatory and demand-driven extension approaches. By quantifying the role of social networks in innovation processes, such an analysis could further support the promotion of farmer-centered innovation processes, such as farmer field schools. Finally, by combining relational and attribute data into one framework, and by considering network architecture in econometric estimations, such a study could contribute to current academic research pursued in this field.

Further research on assessing the impact of social networks is certainly warranted. Although theoretical network models are developed very well, empirical studies that apply these models are not sufficiently available. This may be related to the complex data requirements associated with such an analysis. Nonetheless, more empirical studies could help evaluate more effectively the impact of social networks on economic outcomes.

APPENDIX: MOCK-UP SURVEY FOR EGOCENTRIC INFORMATION NETWORKS

This appendix is based on the hypothetical example, introduced in Section 2, of the adoption of a modern row seeder in rural India. A questionnaire, which could be used to retrieve data on information ego networks, is presented, and different elements of the questionnaire are discussed. It is assumed that the questionnaire presented here is part of an overall household survey. Questionnaire questions are based on Conley and Udry (2001), Hoddinott, Dercon, and Kirshnan (2005), and Bandiera and Rasul (2006). Explanations are given below the questionnaire.

Questions That Could Be Included in a Survey

Q1: Please name six people with whom you discuss important agricultural matters.	Q2: To whom did you give information about the row seeder?	Q3: From whom did you receive information about the row seeder?	Q4: What is the age of each person? (in years)	Q5: What is the education level of each person? (in years)
1 Subramanian	X		34	8
2 Manoj		X	36	6
3 Sundeep	X	X	50	3
4 Prakash	X		25	10
5 Sachin	X	X	41	0
6 Jagdish			32	5

Q6: How many bullocks does this person own?

Q7: What is the size of the land owned by this person? (1: Bigger than mine, 2: Same as mine, 3: Smaller than mine)

Q8: Is this person's land allocated next to yours?

Q9: Where does this person live? (1: Next house, 2: In the village, 3: Outside the village)

Q10: Approximately how far does this person live from you (in meters)?

Q11: Is this person a relative of yours? (1: Yes, 2: No)

Q12: Is this person of the same caste as you? (1: Yes, No: 2)

Q13: How often do you talk with this person per month?

Q14: Who told you about the row seeder first?

Q15: Who convinced you to buy the row seeder?

Q16: When did these persons adopt the row seeder (1: Not yet, 2: Before me, 3: At the same time as me, 4: After me)

Q17: Do you exchange more than agricultural information with this person? (Open question; Potential answers: Borrowing/lending, Seasonal labor, Seeds, etc.)

Q18: In your opinion, who does Subramanian discuss important agricultural matters with?

Manoj? (1: Yes, 2: No)

Sundeep? (1: Yes, 2: No)

Prakash? (1: Yes, 2: No)

Sachin? (1: Yes, 2: No)

Jagdish? (1: Yes, 2: No)

Q19: In your opinion, who does Manoj discuss important agricultural matters with?

Subramanian? (1: Yes, 2: No)

Sundeep? (1: Yes, 2: No)

Prakash? (1: Yes, 2: No)

Sachin? (1: Yes, 2: No)

Jagdish? (1: Yes, 2: No)

Q20: In your opinion, who does Sundeep discuss important agricultural matters with?

Subramanian? (1: Yes, 2: No)

Manoj? (1: Yes, 2: No)

Prakash? (1: Yes, 2: No)

Sachin? (1: Yes, 2: No)

Jagdish? (1: Yes, 2: No)

Q21: In your opinion, who does Prakash discuss important agricultural matters with?

Subramanian? (1: Yes, 2: No)

Manoj? (1: Yes, 2: No)

Sundeep? (1: Yes, 2: No)

Sachin? (1: Yes, 2: No)

Jagdish? (1: Yes, 2: No)

Q22: In your opinion, who does Sachin discuss important agricultural matters with?

Subramanian? (1: Yes, 2: No)

Manoj? (1: Yes, 2: No)

Sundeep? (1: Yes, 2: No)

Prakash? (1: Yes, 2: No)

Jagdish? (1: Yes, 2: No)

Q23: In your opinion, who does Jagdish discuss important agricultural matters with?

Subramanian? (1: Yes, 2: No)

Manoj? (1: Yes, 2: No)

Sundeep? (1: Yes, 2: No)

Prakash? (1: Yes, 2: No)

Sachin? (1: Yes, 2: No)

Q24: Please name two formal sources from which you received information on the row seeder. (open question)	Q25: How often did you talk to these formal sources in the past 12 months?	Q26: How satisfied were you with the information provided by these services? (1: very satisfied, 2: somewhat satisfied, 3: not satisfied, 4: totally unsatisfied
Potential answers: 1 Extension service		
2 Producing company		

Question 1 is the name generator with which the respondent indicates his or her social neighborhood. As explained in Section 5, the maximum number of alters to be named should be based on careful pretesting. Questions 2 and 3 allow for determining the direction of the tie. Questions 4 through 13 are constructed to receive background information on the alters' characteristics. When visualizing networks, these characteristics may play significant roles in determining the position and power distributions within the network. Questions 14 through 16 are intended to receive more information on the adoption process within the network. Farmers who know many adopters can be assumed to be more exposed to the innovation, and they can assess its suitability from the experience of their network partners (Valente 1995). Question 17 provides information on the different layers (financial, labor, etc.) that a network may have. Questions 18 through 23 aim to provide information on information flows within the entire social neighborhood. Questions 24 through 26 provide information on formal extension supplies, which may have had an impact on the decision to buy a row seeder. These questions allow for incorporating not only farmers' networks but also other sources of information.

REFERENCES

Alex, G., D. Byerlee, M. Helene-Collion, and W. Rivera. 2004. Extension and rural development. Converging views on institutional approaches? Agriculture and Rural Development Discussion Paper 4. Washington, DC: The World Bank.

Baerenklau, K. A. 2005. Toward an understanding of technology adoption: Risk, learning, and neighborhood effects. *Land Economics* 8(1): 1–19.

Balconi, M., S. Breschi, and F. Lissoni. 2004. Networks of inventors and the role of academia: An exploration of Italian patent data. *Research Policy* 33: 127–145.

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

Behrman, J. R., H.-P. Kohler, and S. Cotts Watkins. 2002. Social networks and changes in contraceptive use over time: Evidence from a longitudinal study in rural Kenya. *Demography* 39(4): 713–738.

Benin, S., E. Nkonya, G. Okecho, J. Pender, S. Nahdy, S. Mugarura, et al. 2007. Assessing the impact of the National Agricultural Advisory Services (NAADS) in the Uganda rural livelihoods. IFPRI Discussion Paper 724. Washington, DC: IFPRI.

Birner, R., and J. Anderson. 2007. How to make agricultural extension demand-driven? The case of India's agricultural extension policy. IFPRI Discussion Paper 729. Washington, DC: IFPRI.

Brass, D. J. 1985. Men's and women's networks: A study of interaction patterns and influence in an organization. *The Academy of Management Journal* 28(2): 327–343.

Breiger, R. L. 2004. The analysis of social networks. In *Handbook of data analysis,* eds. M. Hardy and A. Bryman, 505–526. London: Sage.

Brock, W. A., and S. N. Durlauf. 2001. Interaction-based models. In *Handbook of econometrics,* eds. J. J. Heckman and E. Leamer. Amsterdam: Elsevier.

Canter, U., and H. Graf. 2005. *The network of innovators in Jena: An application of social network analysis*. Paper presented at the 4th European Meeting on Applied Evolutionary Economics, Utrecht, The Netherlands (May 19–21).

Coleman, J., E. Katz, and H. Menzel. 1957. The diffusion of an innovation among physicians. *Sociometry* 20(4): 253–270.

Coleman, J. S. 1999. Social capital in the creation of human capital. In *Social capital: A multifaceted perspective.* eds. P. Dasgupta and I. Serageldin. Washington, DC: World Bank Publications.

Conley, T. G., and C. R. Udry. 2000. *Learning about a new technology: Pineapple in Ghana* (Center Discussion Paper 817). New Haven, CT: Economic Growth Center, Yale University.

Conley, T. G., and C. Udry. 2001. Social learning through networks: The adoption of new agricultural technologies in Ghana. *American Journal of Agricultural Economics* 83(3): 668–673.

Coulon, F. 2005. The use of social network analysis in innovation research: A literature review. Unpublished paper. Lund University, Lund, Sweden.

Cox, D., and M. Fafchamps. 2006. Extended family and kinship networks: Economic insights and evolutionary directions. Unpublished paper. Forthcoming in *Handbook of Development Economics*.

Crona, B., and Ö. Bodin. 2006. What you know is who you know? Communication patterns among resource users as a prerequisite for co-management. *Ecology and Society* 11(2). http://www.ecologyandsociety.org/vol11/iss2/art7/.

Darr, D., and J. Pretzsch. 2006. *The spread of innovations within formal and informal farmers' groups: Evidence from rural communities of semi-arid Eastern Africa*. Paper presented at the Tropentag 2006, Bonn, Germany (October 11–13).

Darr, D., and J. Pretzsch. 2007. *The influence of structural and functional network properties on the spread of agroforestry innovations within farmers groups: Evidence from Eastern Africa*. Paper presented at the 27th International Sunbelt Social Network Conference, Corfu, Greece (May 1–6).

Davis, K. 2008. *Extension in sub-Saharan Africa: Overview and assessment of past and current models and future prospects*. Paper presented at the 24th Annual Conference of the Association for International Agriculture and Extension Education, EARTH University, Costa Rica, March 10–15.

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

Durlauf, S. N. 2001. A framework for the study of individual behavior and social interactions. Unpublished paper. University of Wisconsin, Department of Economics.

Durlauf, S. N. 2002. On the empirics of social capital*.* *The Economic Journal* 112: F459–F479.

Feder, G., and D. L. Umali. 1993. The adoption of agricultural innovations: A review. *Technological Forecasting and Social Change* 43: 215–239.

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

Goyal, S. 2005. Learning in networks: A survey. In *Group formation in economics: Networks, clubs, and coalitions,* eds. G. Demange and M. Wooders. Cambridge, United Kingdom: Cambridge University Press.

Granovetter, M. S. 1973. The strength of weak ties. *American Journal of Sociology* 78(6): 1360–1380.

Hanneman, R. A., and M. Riddle. 2005. *Introduction to social network methods.* Riverside: University of California Press.

Hartwich, F., M. M. Perez, L. A. Ramos, and J. L. Soto. 2007. *Knowledge management for agricultural innovation within the Bolivian agricultural technology system: Insights from the analysis of rural networks*, Paper submitted in the special November issue of Knowledge Sharing and Knowledge Management For Development in Latin America and the Caribean.

Hoddinott, J., S. Dercon, and P. Krishnan. 2005. *Networks and informal mutual support in 15 Ethiopian villages*. Unpublished manuscript. http://www.economics.ox.ac.uk/members/stefan.dercon/hodd_der_kr.pdf.

Hogset, H., and C. B. Barrett. 2007. *Imperfect social learning among Kenyan smallholders.* Discussion paper. Cornell University, Department of Applied Economics and Management.

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

Jackson, M. O., and A. Wolinsky. 1996. .A strategic model of social and economic networks. *Journal of Economic Theory* 71(1): 44–74.

Krishnan, P., and E. Sciubba. (2007). *Links and architecture in village economies*. Cambridge Working Paper in Economics 0462. Cambridge University, Cambridge, United Kingdom*.*

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

Manski, C. F. 1995. *Identification problems in the social sciences*. Cambridge, MA: Harvard University Press.

Manski, C. F. 2000. Economic analysis of social Interactions. *Journal of Economic Perspectives* 14(3): 115–136.

Matuschke, I., R. R. Mishra, and M. Qaim. 2007. Adoption and impact of hybrid wheat in India. *World Development* 35(8): 1422–1435.

Matuschke, I., and M. Qaim. 2007. What you grow is who you know? The impact of hybrid seed adoption in India. Unpublished discussion paper. University of Hohenheim, Stuttgart, Germany.

Miguel, E., and M. Kremer. 2003. *Networks, social learning and technology adoption: The case of deworming drugs in Kenya*. Working paper. University of California, Berkeley, Department of Economics.

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

Pomp, M., and K. Burger. 1995. Innovation and imitation: Adoption of cocoa by Indonesian smallholders. *World Development* 23(3): 423–431.

Powell, W. W., K. W. Koput, and L. Smith-Doerr. 1996. Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative Science Quarterly* 41: 116–145.

Raini, R. K., C. P. W. Zebnitz, and V. Hoffmann. 2005. *Integrated pest management and information flow: Case study tomato stakeholders' practices in Kenya*. Paper presented at the Tropentag 2005. Stuttgart-Hohenheim, Germany (October 11–13).

Rivera, W., and G. Alex (eds.). 2004. Extension reform for rural development (vol. 1–3). Agriculture and Rural Development Discussion Paper 8–10. Washington, DC: The World Bank.

Santos, P., and C. B. Barrett. 2007. *Understanding the formation of social networks.* Discussion paper. Cornell University, Department of Applied Economics and Management.

Scott, J. 2000. *Social network analysis. A handbook*. London: Sage.

Uzzi, B. 2000. The sources and consequences of embeddedness for the economic performance of organizations: The network effect. *American Sociological Review* 61(4): 674–698.

Valente, T. W. 1995. *Network models of the diffusion of innovations.* Cresskill, NJ: Hampton Press.

Wasserman, S., and K. Faust. 2005. *Social network analysis. Methods and applications*. Cambridge, United Kingdom: Cambridge University Press

World Bank. 2008. *World Development Report 2008.* Washington, DC: The World Bank.

RECENT IFPRI DISCUSSION PAPERS

For earlier discussion papers, please go to www.ifpri.org/pubs/pubs.htm#dp.
All discussion papers can be downloaded free of charge.

815. *Migration and technical efficiency in cereal production: Evidence from Burkina Faso*. Fleur S. Wouterse, 2008.

814. *Improving farm-to-market linkages through contract farming: A case study of smallholder dairying in India*. Pratap S. Birthal, Awadhesh K. Jha, Marites M. Tiongco, and Clare Narrod, 2008.

813. *Policy options and their potential effects on Moroccan small farmers and the poor facing increased world food prices: A general equilibrium model analysis*. 2008. Xinshen Diao, Rachid Doukkali, Bingxin Yu, 2008.

812. *Norway: Shadow WTO agricultural domestic support notifications*. Ivar Gaasland, Robert Garcia, and Erling Vårdal, 2008.

811. *Reaching middle-income status in Ghana by 2015: Public expenditures and agricultural growth.* Samuel Benin, Tewodaj Mogues, Godsway Cudjoe, and Josee Randriamamonjy, 2008.

810. *Integrating survey and ethnographic methods to evaluate conditional cash transfer programs*. Michelle Adato, 2008.

809. *European Union: Shadow WTO agricultural domestic support notifications.* Tim Josling and Alan Swinbank, 2008.

808. *Bt Cotton and farmer suicides in India: Reviewing the evidence.* Guillaume P. Gruère, Purvi Mehta-Bhatt, and Debdatta Sengupta, 2008.

807. *Gender, caste, and public goods provision in Indian village governments.* Kiran Gajwani and Xiaobo Zhang, 2008.

806. *Measuring Ethiopian farmers' vulnerability to climate change across regional states.* Temesgen Deressa, Rashid M. Hassan, and Claudia Ringler, 2008.

805. *Determinants of agricultural protection from an international perspective: The role of political institutions*. Christian H.C.A. Henning, 2008

804. *Vulnerability and the impact of climate change in South Africa's Limpopo River Basin.* Sharon Shewmake, 2008.

803. *Biofuels, poverty, and growth: A computable general equilibrium analysis of Mozambique*. Channing Arndt, Rui Benfica, Finn Tarp, James Thurlow, and Rafael Uaiene, 2008.

802. *Agricultural exit problems: Causes and consequences*. Derek Headey, Dirk Bezemer, and Peter B. Hazell, 2008.

801. *Cotton-textile-apparel sectors of India: Situations and challenges faced*. Jatinder S. Bedi and Caesar B. Cororaton, 2008.

800. *Cotton-textile-apparel sectors of Pakistan: Situations and challenges faced*. Caesar B. Cororaton, Abdul Salam, Zafar Altaf, and David Orden, with Reno Dewina, Nicholas Minot, and Hina Nazli, 2008.

799. *Race to the top and race to the bottom: Tax competition in rural China*. Yi Yao and Xiaobo Zhang, 2008.

798. *Analyzing the determinants of farmers' choice of adaptation methods and perceptions of climate change in the Nile Basin of Ethiopia*. Temesgen Deressa, R.M. Hassan, Tekie Alemu, Mahmud Yesuf, Claudia Ringler, 2008.

797. *Economic transformation in theory and practice: What are the messages for Africa*? Clemens Breisinger and Xinshen Diao, 2008.

796. *Biosafety at the crossroads: An analysis of South Africa's marketing and trade policies for genetically modified products.* Guillaume P. Gruère and Debdatta Sengupta, 2008.

795. *Publish or patent? Knowledge dissemination in agricultural biotechnology*. An Michiels and Bonwoo Koo, 2008.

794. *Agricultural growth and investment options for poverty reduction in Malawi.* Samuel Benin, James Thurlow, Xinshen Diao, Christen McCool, Franklin Simtowe, 2008.

793. *China: Shadow WTO agricultural domestic support notifications*. Fuzhi Cheng, 2008.

792. *India: Shadow WTO agricultural domestic support notifications*. Munisamy Gopinath, 2008.

791. *Agricultural growth and investment options for poverty reduction in Zambia.* James Thurlow, Samuel Benin, Xinshen Diao, Henrietta Kalinda, and Thomson Kalinda, 2008.

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

www.ifpri.org

IFPRI HEADQUARTERS

2033 K Street, NW
Washington, DC 20006-1002 USA
Tel.: +1-202-862-5600
Fax: +1-202-467-4439
Email: ifpri@cgiar.org

IFPRI ADDIS ABABA

P. O. Box 5689
Addis Ababa, Ethiopia
Tel.: +251 11 6463215
Fax: +251 11 6462927
Email: ifpri-addisababa@cgiar.org

IFPRI NEW DELHI

CG Block, NASC Complex, PUSA
New Delhi 110-012 India
Tel.: 91 11 2584-6565
Fax: 91 11 2584-8008 / 2584-6572
Email: ifpri-newdelhi@cgiar.org