Modeling Technology Adoption in Developing Countries

By TIMOTHY BESLEY AND ANNE CASE*

Perhaps one of the main reasons for studying economic development is to understand better how individuals are able to make the transition out of poverty. Technology may be viewed as a means to this end. Yet, while the development of higher-yielding varieties (HYV's) of many crops grown by poor farmers has enhanced this hope, it is essential to understand how new technologies are adopted in practice if their promise is to be fulfilled. In our collective understanding of technology adoption, many questions remain unanswered. For instance, to what extent are socially valuable technologies slow to realize their potential due to information constraints or to externalities that lead the private and social value of new technologies to diverge?

A prior step to answering these questions is to provide an adequate analysis of technology adoption decisions by poor farmers, and the purpose of this paper is to review some possible empirical models for studying technology adoption. In so doing, we will belabor the issue of theoretical consistency. Can researchers ensure that their empirical adoption models are consistent with an underlying choice-based model? What are the costs of such consistency, measured in terms of data needs and model complexity, and what are the benefits, measured in terms of understanding the microeconomic foundations of adoption?

The typical scenario we investigate is as follows. Each of M farmers within a village or a region must decide whether or not to adopt a particular agricultural technology. We are interested in understanding what determines adoption of the technology across space and time. We begin by reviewing three basic empirical approaches that vary according to the type of data available.[1]

I. Empirical Approaches to Analysis of Technology Adoption

A. *Time-Series Studies*

Much of what is known about the adoption of new technologies comes from time-series evidence. In these data, one observes only an aggregate measure of adoption, such as the percentage of farmers employing the new technology at each date. The study of hybrid corn in the United States by Zvi Griliches (1957) is a classic study of this kind. In general, the aim is to capture the shape of the time-series diffusion process, and these studies tend to model the pattern of adoption as a logistic-shaped function over time. Letting p_{it} denote the fraction of adopters in region i at date t, one can estimate equations of the form

$$(1) \qquad p_{it} = f(p_{it-1}) + \varepsilon_{it}.$$

While it may be possible to parameterize the function $f(\cdot)$ by regional characteristics, the main purpose in such studies is often to estimate the intertemporal component of the relationship in (1). While disaggregating by region and investigating the effect of regional characteristics on adoption give some insight into what might drive adoption, this approach is limited in what it can

[†]*Discussants:* Timothy Besley, Princeton University; Anne Case, Princeton University; Adam Jaffe, Harvard University.

*Research Program in Development Studies, Woodrow Wilson School of Public and International Affairs, Bendheim Hall, Princeton University, Princeton NJ 08544. We are grateful for helpful discussions with Angus Deaton and Ariel Pakes throughout this project and for the financial support of the National Science Foundation (SES-9121838).

[1]Our aim here is not to survey the existing theoretical and empirical literature. For that the reader is referred to Gershon Feder et al. (1985).

say about the underlying dynamic process at work.

B. *Cross-Sectional Studies*

There are also many studies of technology adoption that use cross-sectional data. Such data are broadly of two kinds. First, there are studies that take a snapshot of M farmers' technology use at some date. The gain to farmer i of using the new technology is typically parameterized as $\gamma x_i + u_i$, where x_i are farm and farmer characteristics and u_i is an independently and identically distributed farm specific *ex ante* shock. It is often assumed that these shocks are normally distributed, and the model is then run as a probit, so that

$$(2)\quad \text{Prob}\{\text{adoption by farmer } i\} = \phi(\gamma x_i / \sigma_u)$$

where $\phi(\cdot)$ is the distribution function of the standard normal. The intention of this line of research is often to measure the impact of x_i on adoption decisions.

However, this model is problematic if, as the time-series analysis suggests, there is some dynamic structure to the adoption decision. The cross section provides only a snapshot; at that point, the technology may be incompletely diffused through the population. This confounds the interpretation of the coefficients in (2). For example, there may be a time-dependent element in the adoption decision, so that the expected profits are $\gamma x_{it} + \psi_{it}$, where $\psi_{it} = g(\psi_{it-1}, D_{t-1}, x_{it}) + \varepsilon_{it}$, where D_{t-1} is a history of the new technology's use up to period $t-1$. There are many possible interpretations of ψ_{it}, including (i) a farmer's knowledge about the new technology or (ii) evolving costs of adoption, which vary with credit availability. Either way, it may depend upon farm and farmer characteristics at time t. In these cases, ψ_{it} will bias the parameter estimates of γ. Thus, upon completion of the adoption process, cross-sectional studies of this kind may be able to provide insight into the farm and farmer characteristics associated with ultimately accepting the new technology. However, these data are of limited use in exploring the adoption process itself.

Some cross-sectional surveys contain information, based on *recall*, about when a farmer adopted a technology. Under some circumstances, recall data may provide a means for dealing with the above problems. If, for example, the dynamic structure was well represented by $\psi_{it} = \delta_t x_i$, we can augment equation (2) above. Creating for each farmer a set of discrete choice observations, d_{it}, equal to 1 if farmer i was using the technology at time t, $t \in [1,\dots,\tau]$, and zero otherwise, we estimate a probit:

$$(3)\quad \text{Prob}\{d_{it} = 1\} = \phi((\gamma x_i + \rho \text{T} + \delta[\text{T} \times x_i]) / \sigma_u).$$

where T is a set of $\tau - 1$ year indicators and $\text{T} \times x_i$ are interaction terms that allow the influence of field and farm characteristics to change over the diffusion process. While more flexible than (2), this structure is also extremely limiting. It is necessary to maintain the assumption that influential farm and farmer variables x_i do not change over time. In general, this seems unreasonable since farmer wealth and credit-worthiness are apt both to influence and to be influenced by adoption choices taken. If these data are available only at the time when the survey was done, their use in (3) will bias the parameters of interest. It also seems unlikely that the dynamics of adoption can be well captured by allowing time-varying coefficients on variables whose values are *assumed* to be constant over time. Thus, while one might be able to do better by having information about the year in which the technology was adopted, there are still problems.[2]

C. *Panel-Data Studies*

If panel data detailing farm and farmer characteristics and the adoption choices

[2]There are other possible ways of using recall data. For example, the researcher may wish to stratify the sample into individuals who always used the technology during the recall period, those who adopted during the period, and those who never used it. Predicting adopters in such circumstances may be interesting, but is still subject to the criticism that the x_i's may be endogenous.

made at each point in time are available, then a number of the criticisms raised for time-series and cross-sectional data can be met. Here, for example, the researcher can allow for household effects and state-dependence. Thus, consider a model in which the underlying dynamic component is well represented by

$$(4)\quad \psi_{it} = \beta \prod_{i \in N(i)} d_{it-1} + \alpha_t + \lambda_i + u_{it}.$$

James Heckman (1981) offers an excellent survey of methods for handling such models, where λ_i is either a random or a fixed effect.

However, the availability of panel data forces researchers to think harder about reasonable dynamic specifications for discrete choice. For, while there is a rich set of empirical models available for panel data, there is a real question about how these relate to the underlying choice problem that individuals face. For policy analysis, the ability to return to the underlying microeconomic foundations of adoption is imperative. This will lead us, in Section II, to think once again about theory.

II. Dynamic Choice

So far, we have said very little about the dynamic choices that generate the data. In this section, we do this in order to reappraise what empirical methods might be appropriate. We discuss two scenarios, distinguished by their allowance for externalities between farmers' choices.

A. *Dynamic Choices Without Externalities*

Let the dynamic process driving adoption be characterized by a vector of state variables $\mathbf{k}_t = (k_t^1, \dots, k_t^M)$ and let their transition function be $k_{t+1}^i = h(k_t^i, d_{it}; \varepsilon_{it})$, where ε_{it} is an independently and identically distributed shock experienced by farmer i at time t. We can write the probability of observing any value of the state variable in the future as a first-order Markov process: $F(k_{t+1}^i | k_t^i, d_{it})$. We can then consider dynamic choice as being governed by a recursive problem:

$$(5)\quad V_t^i(k_t^i) = \operatorname*{Max}_{d_{it}} \{ \pi_{it}(d_{it}, k_t^i) + \delta E[V_{t+1}^i(k_{t+1}^i) | k_t^i, d_{it}] \}$$

where $V_t^i(k_t^i)$ is farmer i's value function at time t, $\pi_{it}(d_{it}, k_t^i)$ is some current payoff, and δ reflects the discount rate. Several different interpretations of the state variable are possible:

(a) *Credit availability*.—The variable k_t^i may represent current assets that are available to pay for implementation of the new technology.
(b) *Learning*.—The variable k_t^i may represent the stock of knowledge about the new technology which evolves through time.
(c) *An Irreversible Investment*.—The variable k_t^i equals 1 if the investment was ever undertaken previously and 0 otherwise.

In any of these cases, current choices have *future* consequences, and any current decisions ought to weigh these. In designing an econometric specification for studying technology adoption, we would formulate a model whose likelihood function comprised

$$(6)\quad \operatorname{Prob}\{d_{it} = 1\} = \operatorname{Prob}\Big\{ d_{it} \in \operatorname*{argmax}_{\omega} (A | \omega \in \{0,1\}) \Big\}$$

where

$$A = \pi_{it}(\omega, k_t^i) + \delta E\big[V_{t+1}^i(k_{t+1}^i) | k_t^i, \omega\big].$$

The parameters to be estimated would depend upon the exact structure of the underlying choice problem, which we could specify by giving functional form to the Markov transition kernel describing the evolution of k_t^i. To our knowledge, specifying and estimating models of this kind have not been attempted in the literature on agricultural technology adoption, although there is now a considerable literature focused on estimating investment decisions in this way. Two key contributions are Ariel Pakes's (1986) option-value model of the decision to invest in or renew a patent and John Rust's (1987) model of bus engine replacement. Surveys of the literature can be found in Pakes (1993) and Rust (1993).

It is clear from (6) that panel-data models estimable using equation (5) are not necessarily good representations of optimal dynamic choices. The main problem is that optimal choices are *forward-looking*, and we need to find some way of specifying the *future* of some action when specifying the model. The models discussed in (5) above use only lagged values of variables to do this.

B. *Dynamic Choices with Externalities*

Externalities may play an important role in technology-adoption decisions. The literatures cites a number of relevant sources, including:[3]

(i) *Network Externalities.*—Adopters care about how many other individuals adopt because there is some public-good element to the technology. In agriculture, the most common form of externalities arises in the need to build a marketing infrastructure for a new crop.

(ii) *Market Power Externalities.*—Adopters with market power will care about adoption by others if adopting early implies some advantage in market power. We do not know of examples in agriculture where this is important.

(iii) *Learning Externalities.*—Farmers may care about others' adoption decisions if early adopters teach late adopters something. For example, if a technology is of uncertain profitability, some potential adopters may wait until they observe whether others have fared well by using it. We believe that such externalities are potentially important in agricultural technology adoption. This has long been recognized by rural sociologists (see e.g., Everett Rogers [1983] for a review).

The canonical decision problem to be solved in a model with externalities involves conditioning one individual's behavior on that of others. To model this, suppose that the state variables for individual i evolve in a way that depends upon d_{-it}, the choices made by farmers other than i. The interdependent decision-making that ensues is potentially quite complicated. Moreover, a strategy for farmer i, s_{it}, can potentially depend upon all current and past behavior as well as current and past values of the state variables. A common simplification, first suggested in Eric Maskin and Jean Tirole (1988), and used in empirical work by Pakes and Paul McGuire (1991) and Besley and Case (1992), is to confine strategies to depend only upon the vector of current state variables, denoted by k_t. Strategies that comprise a Nash equilibrium at each date are then referred to as *Markov perfect*. The optimal decision for farmer i conditioning in d_{-it} is then characterized by the recursion

$$(7)\quad \hat{V}_t^i(k_t^i; d_{-it}) = \max_{d_{it}} \{\pi_{it}(d_{it}, k_t^i) + \delta E\{V_{t+1}^i(k_{t+1}) | k_t^i, d_{it}, d_{-it}\}\}$$

where $V_{t+1}^i(k_t) \equiv \hat{V}_t^i(k_t^i; s_{-it}(k_t))$, so that the equilibrium value function now depends upon the whole vector of state variables.[4] Studying problems like (7), even as a pure

[3] For review of the theoretical literature, see Jennifer Reinganum (1990). We are not aware of an empirical literature which has investigated these issues in industrial organization.

[4] We use the following shorthand notation:

$$s_{-it}(k_t) \equiv [s_1(k_t^1), \ldots, s_{i-1t}(k_t^{i-1}), s_{i+1t}(k_t^{i+1}), \ldots s_{Mt}(k_t^M)].$$

Even though current payoffs depend only on k_t^i, the value function depends upon the whole vector k_t, because individual j's ($j \neq i$) strategy depends upon k_t^j.

computational venture, is quite demanding. Consequently, few studies have explored dynamic models of equilibrium behavior in any context, including agricultural technology adoption. The next section presents one approach to solving such a problem.

III. Examples: Learning from the Behavior of Others

We focus here on how one might build a dynamic multiagent model of learning in the adoption of a new technology. The first example is based on Besley and Case (1992), to which the reader is referred for a full development and empirical implementation of the model. In that study, we apply the model to the diffusion of a new cotton seed in a south Indian village.

In a world of learning, we interpret k_t^i as the state of farmer i's knowledge about the new technology at time t. Our first simplification is to suppose that knowledge is a public good so that $k_t^i = k_t$ for all i. Suppose that each farmer has one field and is deciding whether or not to sow this with a new seed. Each farmer observes the yields other farmers obtain on their land. Knowledge evolves with the realization of yields from past planting decisions. We parameterize the model so that k_t is the expected gain from the new technology. The planting decision at time t by farmer i is based on expected gains $\pi_{it}^* = k_t + \nu_{it}$, where ν_{it} is a random shock not observed by the researcher. The realized gain in profits using the new seed, π_{it}, differs from the expected gain by a shock ε_{it}, assumed to be normally distributed with mean zero and variance σ_ε^2. Then the prior distribution of k, the change in profitability attributable to the new seed, is normal with mean k_{t-1} and variance $\sigma_{k,t-1}^2$, and we have the following standard Bayesian updating formula for normal distributions: if m_t farmers sow to the new seed at time t,

$$(8)\qquad k_{t+1} = k_t + \left\{m_t + \sigma_\varepsilon^2/\sigma_{k,t}^2\right\}^{-1} \times \sum_i d_{it}(\pi_{it} - \pi_{it}^*)$$

(see, e.g., Morris H. De Groot, 1970 p. 167.) This is the source of the externality, since updating depends upon the total number of plots sown.

As shown in Besley and Case (1992), equation (8) can be used to characterize the Markov transition kernel for this problem. It is normal and depends upon the variances $(\sigma_\varepsilon^2, \sigma_{k,t}^2)$. That paper also shows how the parameters $(\sigma_\varepsilon^2, \sigma_\nu^2)$ might be identified from farmer-level data. Finally, the estimated parameter values are used to simulate dynamic choices as suggested in equation (7). Each simulation gives an equilibrium set of predicted choices $\hat{d}_{it}$. Each set of such simulations is, however, conditioned on a set of starting values. To find these values, we take advantage of a technique, first suggested by Pakes (1986) and developed in Pakes and David Pollard (1989), of computing the pseudo-likelihood function from simulated runs from the model. There remain a number of issues to be dealt with, such as the treatment of multiple equilibria. The data may suggest appropriate refinement strategies; research in this area is still unfolding.

A different approach to learning and technology adoption has recently been put forward by Glenn Ellison and Drew Fudenberg (1991), who suggest some rules of thumb to capture adopters' behavior in situations in which individuals may learn from each other's past experience but do not solve the kinds of dynamic optimization problems given by (7). Given the complexity of the calculations required in structural dynamic models, this direction of research is appealing. But is such an approach useful? One way of appraising this would be by fitting such models to the data and comparing them with other possibilities available to the researcher, such as structural models. To illustrate how this might be done, consider Ellison and Fudenberg's (1991) simplest rule for time-series data:

$$(9)\qquad p_{it} = \begin{cases} (1-\alpha)p_{it-1} + \alpha & \text{with probability } q \\ (1-\alpha)p_{it-1} & \text{with probability } 1-q. \end{cases}$$

The parameter α represents the inertia in the population. Only a fraction α of the population considers switching technologies each period. The parameter q represents the probability that the new technology is better than the old one. The likelihood function for the model in (9) for a sample $(p_{i1},\dots,p_{iT})$, given some initial p_{i0}, is

$$(10)\quad \mathscr{L}(p_{i1},\dots,p_{iT}|p_{i0}) \equiv \prod_{\tau=1}^{T} \mathrm{Prob}\{p_{i\tau}|p_{i\tau-1}) = \prod_{\tau=1}^{T}\{(1-\alpha)p_{i\tau-1}+\alpha q\}.$$

It is, in principle, possible to find the values of (α,q) that maximize (10) for some data set. Computationally, there are certainly advantages in using rule-of-thumb models. There are limits, however, in the use of such models, especially for policy analysis. If learning is important, then even socially optimal paths would have gradual diffusion of a technology. One would ideally like to compare the private and socially optimal situations, and rules-of-thumb models provide little help in this respect.

IV. Conclusions

One of the key decisions in modeling technology adoption concerns the extent to which empirical estimation is consistent with an underlying theoretical model of optimizing behavior. If the adoption decision is motivated by models of the form (5) or (7), researchers may face a dilemma; it appears that the data and computation costs of research are rather high. Some researchers find structural models unconvincing, noting the lack of specification tests or well-defined null hypotheses to test the model against. Of course, this critique is not confined to dynamic structural approaches, and Angus Deaton (1992) broadens it to other studies in development economics.

At the other extreme, some models pay only lip service to the underlying theory. It is often difficult to interpret results from a model that does not correspond to an underlying decision process. To get the best of both worlds, one might try to combine both types of modeling on any given data set. Statistical models may suggest what is worth modeling structurally. The latter is then a good way of developing theoretical models that are better tailored to the data under consideration, thereby encouraging researchers to think more carefully about the underlying process generating the data.

REFERENCES

Besley, Timothy and Case, Anne, "Taking Learning Seriously: A Diffusion Model for HYV Cotton," unpublished manuscript, Princeton University, 1992.

Deaton, Angus, "Data and Econometric Tools for Development Economics," in Jere Berhman and T. N. Srinivasan, eds., *Handbook of Development Economics*, Vol. III, Amsterdam: North-Holland, 1993 (forthcoming).

De Groot, Morris H., *Optimal Statistical Decisions*, New York: McGraw-Hill, 1970.

Ellison, Glenn and Fudenberg, Drew, "Rules of Thumb in Social Learning," unpublished manuscript, Massachusetts Institute of Technology, 1991.

Feder, Gershon, Just, Richard E. and Zilberman, David, "Adoption of Agricultural Innovations in Developing Countries: A Survey," *Economic Development and Cultural Change*, January 1985, *33*, 255–98.

Griliches, Zvi, "Hybrid Corn: An Exploration in the Economics of Technological Change," *Econometrica*, October 1957, *25*, 501–22.

Heckman, James, "Statistical Models for Discrete Panel Data," in Charles Manski and Daniel McFadden, eds., *Structural Analysis for Discrete Data with Econometric Applications*, Cambridge, MA: MIT Press, 1981, pp. 114–78.

Maskin, Eric and Tirole, Jean, "A Theory of Dynamic Oligopoly I: Overview and Quantity Competition with Large Fixed Costs," *Econometrica*, May 1988, *56*, 549–70.

Pakes, Ariel, "Patents as Options: Some Estimates of the Value of Holding European

Patent Stocks," *Econometrica*, July 1986, *54*, 755–84.

______, "Estimation of Dynamic Structural Models: Problems and Prospects, Part II: Mixed Continuous—Discrete Models and Markets Interactions," in C. Sims and J. J. Laffont, eds., *Proceedings of the Sixth World Congress of the Econometric Society*, Cambridge: Cambridge University Press, 1993 (forthcoming).

______ **and McGuire, Paul,** "Computing Markov Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model," unpublished manuscript, Yale University, 1991.

______ **and Pollard, David,** "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, September 1989, *57*, 1027–57.

Reinganum, Jennifer, "The Timing of Innovation: Research, Development and Diffusion," in Richard Schmalensee and Robert Willig, eds., *Handbook of Industrial Organization*, Amsterdam: North-Holland, 1990, pp. 849–908.

Rogers, Everett M., *Diffusion of Innovations*, New York: Free Press, 1983.

Rust, John, "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, September 1987, *53*, 999–1033.

______, "Estimation of Dynamic Structural Models: Problems and Prospects, Part I: Discrete Decision Processes, in C. Sims and J. J. Laffont, eds., *Proceedings of the Sixth World Congress of the Econometric Society*, Cambridge: Cambridge University Press, 1993 (forthcoming).