in (5.5), however, one can still prove a separation result between production and hedging (speculative) decisions. Specifically, in such a case the optimal output level q^* does not depend on the parameters of the producer's subjective distribution of futures prices [Lapan et al. (1991)], although it does depend on the agent's degree of risk aversion and on the parameters α and β, which define the expectation of the cash price (conditional on the futures price).

The results just outlined pertain to a static problem and, more crucially, pertain to a competitive producer who faces only price risk. For most commodities, however, the hedging problem needs to consider the fact that farmers typically are exposed to both price and production uncertainty. An early attempt at allowing both price and production risk was that of McKinnon (1967), who considered the hedging problem of minimizing the variance of profit for a given planned output level. Because of the complications generated by the joint presence of price and production risk, efforts to extend McKinnon's risk-minimization analysis to EU maximization often have relied on the assumption that producers maximize an objective function increasing in the mean and decreasing in the variance of revenue/profit. This approach was followed by Rolfo (1980), Newbery and Stiglitz (1981, Chapter 13), and Anderson and Danthine (1983), among others. In these studies it is shown that the correlation between the random production and random price is crucial for determining the optimal hedging strategy. Because demand considerations suggest the correlation is typically negative, a "natural" hedge is already built into the price system and the optimal strategy is to hedge an amount lower than expected output.

Such a mean-variance approach usually is justified on the grounds that it is exact for a CARA utility function if wealth/profit is normally distributed. But profit typically is not normally distributed when output is uncertain because it entails the product of two random variables [Newbery (1988)]. Indeed, the need to analyze our hedging problem in a general framework is clearly illustrated by noting that, under production uncertainty, the optimal hedge in general is less than expected output even when output and price are independent [Losq (1982)], a result that cannot be established by mean-variance analysis. Of course, the difficulty is that it is not possible to establish useful general hedging results that hold for arbitrary concave utility functions and arbitrarily jointly distributed random prices and quantities. If one assumes a CARA utility function, however, an exact solution to the hedging problem under production uncertainty may be possible, as illustrated by Bray (1981), Newbery (1988), and Karp (1988).

A model that captures the essence of a typical farmer's planting hedge was presented in Lapan and Moschini (1994), who consider futures hedging for a competitive producer who faces both production (yield) and price risk and whose only available hedging instrument is a futures contract (with basis risk). Following Newbery and Stiglitz (1981), stochastic output is represented in terms of a production function with multiplicative risk, i.e., $\tilde{Q} = \tilde{y}q(x)$, where x denotes the vector of inputs, $\tilde{y}$ is a random variable with mean $\bar{y}$, and $\tilde{Q}$ is random output. As noted earlier, with multiplicative production risk, input choices can still be represented by a standard cost function, say $C(q)$ where q denotes the scale of production.[32] With input prices assumed constant (they are typically known at the time production and hedging decisions are made) and subsumed in the function $C(\cdot)$, realized total profits are[33]

$$\tilde{\pi} = \tilde{p}\tilde{y}q - C(q) + (f_0 - \tilde{f})h. \tag{5.7}$$

Thus, the producer knows f_0 when q and h are chosen, but the realizations of the random variables $\{\tilde{f}, \tilde{p}, \tilde{y}\}$ are not known. The difference between $\tilde{f}$ and $\tilde{p}$ reflects basis risk.

Within this context, and assuming that producers maximize a CARA utility, and that the three random variables $\{\tilde{f}, \tilde{p}, \tilde{y}\}$ are jointly normally distributed, Lapan and Moschini (1994) derive and discuss the exact analytic solution to the optimal hedging problem. In particular, they show that the optimal futures hedge satisfies

$$h^* = \frac{f_0 - \overline{f}}{\lambda S_{11}} + q\left[\overline{y}\frac{S_{12}}{S_{11}} + \overline{p}\frac{S_{13}}{S_{11}}\right]. \tag{5.8}$$

Here S_{ij} are the elements of the matrix $S \equiv [\lambda q B + V^{-1}]^{-1}$, where λ is the coefficient of absolute risk aversion, V is the variance-covariance matrix of the three random variables, and B is an accounting matrix of zeros and ones. Hence, an important result here is that the optimal hedge does depend on the degree of risk aversion, even when the futures price is perceived as unbiased. This insight was not present in earlier mean-variance models of hedging under production uncertainty [e.g., Rolfo (1980), Newbery and Stiglitz (1981)]. For likely parameter values, this risk preference effect may be important and the optimal hedge may differ substantially from the mean-variance one. Furthermore, the optimal hedge under yield uncertainty depends on the conditional forecast of the harvest price ($\overline{p}$) and of the yield term ($\overline{y}$), even when the futures price is perceived as unbiased. Thus, in addition to precluding the separation result, production uncertainty also entails that the optimal hedge is inherently time-varying because conditional forecasts will be revised as harvest approaches.

The empirical application reported by Lapan and Moschini (1994), based on a generalization of Myers and Thompson's (1989) hedge ratio estimation procedure, showed that the optimal hedge is considerably less than the full hedge, and that the amount sold forward declines as risk aversion increases. Of course, CARA, joint normality, and multiplicative production risk are rather restrictive assumptions, but nonetheless this model is useful because it can relax the straitjacket of the mean-variance framework and provide insights into the EU maximizing optimal hedge. Although analytical results based

---

[32] Thus, for any level of inputs, $q = q(x)$. In this setting, q aggregates planted acreage and other inputs, and $\tilde{y}$ reflects random yield.

[33] Of course, simultaneous use of crop insurance contracts (discussed later) would alter the nature of this problem.

on more general assumptions are difficult to obtain, empirically it is easy to consider alternative risk preference structures and stochastic distributions. For example, Lapan and Moschini (1994) solve numerically for the optimal hedge for CRRA preferences and log-normally distributed $\{\tilde{f}, \tilde{p}, \tilde{y}\}$, and find that the conclusions obtained under CARA and normality are reasonably robust.[34]

5.1.2. Options on futures

Among the instruments traded on commodity exchanges, futures contracts arguably have the most direct relevance to risk management for farmers. With the introduction of options on futures for many commodities in the 1980s, however, the possibility of trading put and call options has attracted considerable attention.[35] The use of options as hedging devices when the producer faces only price (and basis) risk (but not production risk) was considered by Lapan, Moschini and Hanson (1991). They emphasize that the inclusion of commodity options in a decision maker's portfolio leads to a violation of the two main conditions for a mean-variance representation of expected utility: (i) options truncate the probability distribution of price (so that the argument of the utility function, profit or wealth, is not normally distributed even if the random price is normal), and (ii) the use of options generally means that the argument of utility is not monotonic in the random attributes. The model essentially entails adding another hedging instrument (options) to the payoff in Equation (5.7). A basic modeling issue here is that, given the presence of futures, one of these basic types of options is redundant (for example, a put can always be constructed using a futures and a call). Hence, for modeling purposes attention can be limited to any two of the three types of assets (futures, puts, and calls). Equivalently, as emphasized by Lapan, Moschini and Hanson (1991), one can consider futures and a combination of puts and calls such as straddles.[36] The use of futures and straddles is fully equivalent to allowing the use of futures and calls (or puts), but it has the analytical advantage of illuminating the interpretation of a number of results because the payoff of a straddle is essentially orthogonal to the payoff of a futures contract.

Lapan, Moschini and Hanson (1991) show that, when the futures price is unbiased (from the producers' own point of view), then options are redundant hedging instruments. The key insight here is that, unlike futures contracts, options allow the construction of payoffs that are nonlinear in the realized futures price. But when futures prices

---

[34] Whereas the discussion here has emphasized price-contingent contracts, some yield futures have traded on the Chicago Board of Trade. Clearly, such contracts are potentially useful for farmers (provided enough liquidity exists). A mean-variance analysis of the hedging problem with both price and yield futures is presented by Vukina, Li, and Holthausen (1996).

[35] A "put" conveys to the buyer the right to sell the underlying futures contract at a given price (the "strike price"). This right can be exercised over a certain period of time (the life of the option), and for this right the buyer must pay a "premium" (the price of the option) to the seller (the underwriter). Similarly, a "call" conveys to the buyer the right to sell the underlying futures at the strike price during the life of the option. See [Cox and Rubinstein (1985)] for more details.

[36] A (short) straddle can be constructed by selling one call and one put at the same strike price (or, because of the redundancy just mentioned, it can be constructed by buying one futures and selling two calls).

and options premiums are perceived as unbiased (such that the only reason to trade these instruments is to hedge the risky cash position), the relevant payoff of the producer is linear in the futures price. Hence, the optimal hedging strategy involves using only futures contracts, which provide a payoff that is linear in the price of interest (the option payoff is uncorrelated with the risk that remains after the optimal futures hedge). If futures prices and/or options premiums are perceived as biased, however, then there is a speculative motive to trade futures and options, and options are typically used along with futures.

In this context it is clear that a hedging role for options is likely when there is a nonlinear relation between profit and the futures prices, such as the presence of nonlinear basis risk or the presence of production uncertainty together with price uncertainty. The latter situation is obviously of great interest to farmers, and has been analyzed by Moschini and Lapan (1995). They study the problem of a farmer with end-of-period profit given by

$$\tilde{\pi} = \tilde{p}\tilde{y}q - C(q) + (f_0 - \tilde{f})h + \left(r - |\tilde{f} - k|\right)z, \tag{5.9}$$

where z is a short straddle with strike price k and premium r (note that the payoff of the straddle depends on the absolute value of the difference between realized futures price and strike price). The producer knows f_0, k, and r when q, h, and z are chosen, but the realizations of the random variables $\{\tilde{f}, \tilde{p}, \tilde{y}\}$ are not known. Under the assumption of CARA and normality, Moschini and Lapan (1995) provide analytic solutions for the optimal use of futures and straddles. If futures and options prices are perceived as unbiased, then the optimal hedging strategy entails selling futures and buying straddles. Of course, because of the simultaneous presence of price and production uncertainty, the optimal use of the hedging instruments depends on the agent's degree of risk aversion, and in general the optimal hedge is less than the full hedge. For example, for a representative soybean producer with a local relative risk aversion of $R = 2$, and after translating optimal levels of futures and straddles into futures and puts, the optimal hedge is to sell futures in an amount of about 63 percent of the expected output and to buy puts in an amount of about 15 percent of expected output.

If the producer perceives the futures and straddle prices as being biased, then there is a speculative motive to trade these assets. An interesting result here is that, if the agent perceives only the options price to be biased, then only the straddle position is affected, whereas if only the futures price is perceived as biased, both futures and options positions will be affected.[37] This result is reminiscent of the speculative hedging role of options illustrated by Lapan, Moschini and Hanson (1991, 1993), and in particular, cannot be obtained by using the special mean-variance framework.

---

[37] Thus, options are useful to provide insurance against the risk of speculating on the futures price because the nonlinearity of their payoffs can compensate for the speculation outcome of extreme price realizations. But futures are not useful to hedge the speculative risk induced by the optimal option use under biased option prices.

5.1.3. The time pattern of hedging

The discussion so far has dealt with a simple version of the hedging problem, a one-period (two-dates) model. At the beginning of the period, when the risky cash position is incurred (say, when corn is planted or when feeder cattle are bought and placed in the feedlot), the farmer hedges by trading futures and other derivatives (options). At the end of the period, when the cash position is liquidated (because the crop is harvested or the cattle are sold), the financial positions are closed. But what if the farmer were free to adjust the futures hedge after it is established and before it is closed? Two questions are relevant here. Does the possibility of revising the optimal hedge affect the initial hedging decision? And, if it is optimal to revise the hedge over time, how is the hedge revised? These problems have been addressed, in different contexts, by (among others) Anderson and Danthine (1983), Karp (1988), and Myers and Hanson (1996). It turns out that the answer to these questions depends crucially on, among other things, whether the producer believes that futures prices are biased or unbiased, and whether or not there is production uncertainty in the model.

Because our focus is on risk reduction (hedging), suppose that futures prices are unbiased. Also, consider first the pure price and basis risk case (no production risk), and suppose that there are T periods, with the initial hedge being taken at $t = 0$, and the last hedge being lifted at $t = T$, and that the terminal profit of the producer is

$$\tilde{\pi}_T = \tilde{p}_T q + \sum_{t=1}^{T} (1+i)^{T-t} (\tilde{f}_t - \tilde{f}_{t-1}) h_{t-1} - C(q), \tag{5.10}$$

where i is the per-period interest rate. If the producer maximized the EU of terminal profit, $\mathrm{E}[U(\tilde{\pi}_T)]$, then the optimal hedging problem (for any given level of output q) can be solved by backward induction. Suppose first that $i = 0$ and that the linear basis assumption made earlier is rewritten as

$$\tilde{p}_T = \alpha + \beta \tilde{f}_T + \tilde{\theta}_T. \tag{5.11}$$

Then, it is easily shown that the optimal hedge is to sell an amount $h_t^* = \beta q$ for all $t = 0, \ldots, T-1$. Thus, if futures prices are unbiased, the static optimal hedge solution gives the optimal hedging strategy at any time based upon the conditional moments available at that time. In particular, the myopic hedging rule (i.e., the hedge that does not take into account that later revisions in the hedge positions are possible) is the same as the optimal dynamic hedging strategy [Karp (1988)].

Because profits/losses of the futures position are marked to market in Equation (5.10), if the interest rate is positive then the optimal futures hedge at time t should be adjusted by a factor of $(1+i)^{T-t}$. This gives a first, albeit trivial, reason for the pure hedge to change as time t moves from 0 to T, as the amount sold forward will increase over time because of this pure discounting effect. As harvest approaches, the agent may revise her expectations about futures (and therefore cash) prices at T. However, there would be no need to adjust the futures position through the growing season due to these changed price expectations, provided the farmer continued to believe that the futures price was unbiased. A second reason to revise the hedge position arises if the moments of the distribution of cash and futures prices (for time T) change over time (as a result of new information), in which case the optimal hedge will be revised as time progresses from t to T, as illustrated by Myers and Hanson (1996). Furthermore, in that situation the ability to revise the futures hedge does affect the initial (at time $t = 0$) hedge position, so that myopic and optimal dynamic hedges differ.

As illustrated by Anderson and Danthine (1983), Karp (1988), and Lapan and Moschini (1994), production uncertainty gives yet another fundamental reason for the optimal hedge to change over time. Because production uncertainty implies that the futures market cannot provide a perfect hedge, the hedge itself depends on the agent's forecast of realized cash price (realized futures price) and realized yield, even when the futures price is unbiased [recall Equation (5.8)]. Clearly, changes in expectations of realized yields (and hence output) will lead to revisions in the futures position. Even if yield forecasts do not change, however, changes in the futures price (and therefore in the expected cash price) will lead to changes in the optimal hedge if the realizations of yields and price are correlated.

A somewhat different dynamic hedging problem arises when the production setting allows for some inputs to be chosen after the uncertainty is resolved, as in the ex-post flexibility models of Hartman (1976) and Epstein (1978). This hedging problem has been studied by Moschini and Lapan (1992), who emphasize that in this model the ex-ante profit of the firm is nonlinear (convex) in the risky price (hence, once again, the mean-variance framework is unlikely to be very useful unless one is willing to assume that the utility function is quadratic). They derive a special case of the separation result for this instance of production flexibility (without basis and production risk, of course), which attains when the shadow price of the quasi-fixed input (the input that is chosen ex-ante) is linear in the output price. This linearity means that the incremental risk due to changes in the quasi-fixed inputs can be fully hedged using futures (because the payoff of the futures position is also linear in price). The nonlinearity of profit in the risky price, however, means that not all income risk can be hedged via futures for the case of production flexibility, and thus there is a pure hedging role for options, over and above that of futures.

5.1.4. Hedging and production decisions

The hedging review so far has emphasized the optimal use of hedging instruments conditional on a given output or a given expected output. An important but distinct question concerns how the availability of these hedging opportunities affects the firms' choice of output. As mentioned earlier, in the special case where basis risk and production risk are absent, the availability of futures contracts allows a separation between production and hedging (speculative) decisions. Specifically, the futures price determines the optimal output level, irrespective of the subjective beliefs of the producer, and any difference between the agent's price expectations and the prevailing futures price only affects the hedging/speculative position. Even in this simple case, however, whether the hedging opportunity increases output depends crucially on whether the futures price is biased or not. If the futures price is perceived as unbiased, then the availability of futures hedging induces the risk-averse firm to expand output.

When we relax the restrictive assumptions that lead to the separation result, and allow for basis and production risk (in addition to futures price risk), in general the planned output of the risk-averse firm will depend on both the futures price and price expectations. The question of how hedging affects the choice of planned output, therefore, is only meaningful in the context of unbiased prices, but even in this context it turns out that general propositions are not possible. Some insights, however, are provided by Moschini and Lapan (1995) for the case of jointly normally distributed random variables and CARA preferences. In particular, they show that if the level of risk aversion is small or if the orthogonal production risk is sufficiently small and the futures price is unbiased, then the availability of futures hedging induces the risk-averse firm to produce a larger output level. Essentially, the ability to hedge effectively changes (increases) the risk-adjusted price the firm perceives for its output. Similarly, it is shown that, if the degree of risk aversion or the level of pure production risk is not too large and futures and option prices are unbiased, then the availability of options (in addition to futures) also causes the firm to increase output.

5.1.5. The value of hedging to farmers

Whereas the foregoing cursory review suggests a potentially important role for futures and option contracts to manage farmers' risk, empirical surveys often find that use of such contracts by farmers is limited.[38] Many explanations for this situation have been offered. From a purely economic point of view, it is clear that existing futures markets do not complete the set of markets in the Arrow–Debreu sense, and thus futures are unlikely to provide a full hedge in a number of production situations. For example, as discussed earlier, consideration of basis and other risks may substantially affect (typically reduce) the optimal futures hedge. Furthermore, even abstracting from basis and other risks, one may note that the time horizon of existing futures is limited (i.e., the most distant delivery date for agricultural futures is often little beyond one year). Thus, producers who hedge optimally their one-period risk are still exposed to some intertemporal price

---

[38] A recent survey [U.S. General Accounting Office (1999)] finds that use of risk management tools by farmers is actually fairly common in the United States. In 1996, 42 percent of the United States' two million farmers used one or more risk management tool, and use of risk management strategies was even more frequent for larger farms. For example, among farmers with annual sales greater than \$100,000, 55 percent used forward contracts and 32 percent engaged in hedging with futures and/or options (52 percent of these farmers also purchased crop insurance, a risk management tool discussed below).

risk even after accounting for "rollover" hedging [Gardner (1989), Lapan and Moschini (1996)].

From a more practical viewpoint, certain costs of hedging that are typically neglected in the analysis, such as brokerage fees, initial deposit, and the requirement to mark to market, may deter hedging activities. Lence (1996) argues that such costs may make the net benefits of hedging almost negligible and may help explain why many farmers do not hedge. Also, limited use of futures by farmers may, to a certain extent, result from mistrust and lack of proper education on the working of such instruments, an observation that suggests a clear scope for extension activities. But one should also keep in mind that the futures markets may be indirectly quite important for agricultural risk management even when many farmers do not use futures contracts directly. For example, futures may be routinely used by country elevators to hedge the risk of storing grain, and these elevators may in turn offer forward contracts to farmers.

5.2. *Crop insurance*

Given the susceptibility of crop yields to weather fluctuations, there is obviously a latent demand for crop insurance. Although crop insurance markets have existed for a long time in some parts of the world (e.g., the United States, Canada, and Sweden), their existence has depended crucially on government support, and these governments often have seen fit to subsidize or even run crop insurance markets. Unsubsidized private insurance markets for agricultural risks have been confined mostly to single-peril insurance contracts. Wright and Hewitt (1990) express the belief that private agricultural insurance markets may fail because the costs of maintaining these markets imply unacceptably low average payouts relative to premiums. Furthermore, they suggest that the perceived demand for crop insurance may be overstated because farmers can use diversification and savings to cushion the impact of a poor harvest on consumption. Although Wright and Hewitt's conjectures are solidly motivated, little has been done to verify the claims empirically. It seems clear, however, that unsubsidized agricultural insurance may not be attractive to farmers because it may be too costly. In particular, the costs of private insurance contracts arise, in part, from information problems that are inherent in these insurance contracts, and it is to these problems that we now turn.

Almost invariably crop insurance markets that have benefited from government intervention, especially for multiple-peril contracts, have been either unexpectedly costly to maintain or unattractive to producers, or both. Consider, for example, the case of the U.S. Federal Crop Insurance Corporation (FCIC), which subsidizes insurance for U.S. crop growers. Below is a table of acreage participation rates and loss ratios for some of the major grain and oilseed crops over the ten-year period 1987 to 1996. The loss ratio is the ratio of indemnities to premium payments, and does not include premium subsidies.[39] When one notes that loss ratios of no more than 0.7 are deemed necessary for

---

[39] In addition to subsidizing premiums, the FCIC also absorbs the administrative costs.

Table 1
FCIC Coverage and Payouts 1987–1996

Crop	U.S. acres planted (millions)	Acres that are FCIC insured (percent)*	Loss ratio*
Wheat	71.0	46.8	1.53
Corn	73.6	38.3	1.22
Soybeans	59.9	35.3	1.06
Sorghum	11.4	37.9	1.37
Barley	8.9	44.0	1.44
Rice	2.9	29.5	2.42

* Averages reported are the annual numbers averaged over 10 years.

Sources: United States Department of Agriculture (1996) and Federal Crop Insurance Corporation (1997).

unsubsidized insurance to be viable given the administrative costs of running it [Wright (1993)], it is clear that the acreage premia would have to be raised substantially for the program to be self-sustaining. Even so, despite heavy government involvement, the subsidized programs are insufficiently generous to attract even a majority of acres planted to these crops. Indeed, the reported participation rates are artificially high because in 1989 and some subsequent years producers had to sign up to be eligible in the event of ad hoc relief, and in 1995 producers had to sign up in order to be eligible for very attractive target price programs. Knight and Coble (1997) provide a detailed overview of the multiple-peril crop insurance environment since 1980. Given that a good insurance policy should attract decision makers who are willing to lose money on average in order to have a less variable income, it is obvious that the FCIC programs have left much to be desired.

Not the least of the problems that arise in crop insurance markets is the existence of a strong political interest in their perceived success. Although the political aspects of these markets are many and varied, the following provides a flavor. Just as in the United States, government involvement in Canadian crop insurance markets has been both extensive and of questionable success. One of the precursors to crop insurance in Canada was the 1939 (federal) Prairie Farm Assistance Act. In the words of the Minister of Agriculture at the time, and referring to a long-standing federal policy of encouraging the settlement of the Prairie provinces, the act "... is intended to take care of people who were put on land that they should never have been put on. That is our reason for being in this at all, and it is our reason for paying two-thirds or three-quarters of the costs out of the treasury of Canada (Standing Committee on Agriculture and Colonization)". Sigurdson and Sin (1994) provide a description of the political history of Canadian crop insurance policy, and Gardner (1994) gives an overview of the United States crop insurance policy in relation to other agricultural policies.

In the United States, one of the more important political aspects of crop insurance is the unwillingness of the federal government to ignore the pleas for monetary disaster assistance when a crop failure is widespread. Given that farm-level crop failures tend to be strongly positively correlated, this undermines the incentive to purchase crop insurance. Disaster assistance is an example of one economic problem – moral hazard – that afflicts crop insurance markets.

When considering a risk, insurance companies may observe certain parameters of the decision environment such as geographic location, soil type, and yield history. They may also observe certain actions such as input use. It is often infeasible to observe all relevant facts, however, and even if observable it may be impossible to write an insurance contract based upon these observations. When it is impossible or excessively costly to write a contract based upon relevant actions, then moral hazard problems may arise. Similarly, when contracts based upon relevant environmental parameters are infeasible, then adverse selection problems may arise. In the remainder of this section, we delineate the nature of the two major economic incentive problems that impede well-functioning crop insurance contracts, and we discuss possible remedies to these problems.

5.2.1. Moral hazard

A risk-neutral insurer who is contemplating the business of a risk-averse producer will seek to specify a contract payout schedule, net of premium, such that a profit is made on the average and also that the producer finds the contract to be sufficiently attractive to sign. Using a standard principal-agent model, as in Chambers (1989), let R be gross revenue and let $I(R)$ be the net contract payoff schedule (premium minus payout), with $C[I(R)]$ as the cost of administering that payoff schedule. Then, assuming symmetric information, i.e., that the insurer can contract upon observable input choices, the insurer's problem is

$$\operatorname*{Max}_{x, I(R)} \int_a^b \{I(R) - C[I(R)]\}\, \mathrm{d}F(R \mid x) \quad \text{such that}$$

$$\int_a^b U\bigl[R - I(R) - rx\bigr]\, \mathrm{d}F(R \mid x) \geqslant \bar{u}, \tag{5.12}$$

where R is supported on $[a, b]$, $\bar{u}$ is the minimum level of expected utility that must be maintained to entice the producer to insure, $F(R \mid x)$ is the revenue distribution function conditional on the input vector x, and r is the input price vector.

Standard analysis, due to Borch (1962), yields the requirement that $I(R)$ satisfy the point-wise condition

$$\frac{1 - C_{I(R)}[I(R)]}{U_\pi[\pi]} = \mu, \tag{5.13}$$

where μ is the Lagrange multiplier for the EU constraint in problem (5.12). Now, if the insurer's cost is invariant to the nature of the schedule, then optimality requires $U_\pi[\pi]$

to be constant, and so for risk-averse producers $I(R)$ must be such that $R - I(R) - rx$ is constant. This is the classical risk-sharing result, namely that the risk-neutral insurer should accept all risk from the risk-averse producer. Under general conditions, this result continues to hold if the insurer is risk averse but contracts upon a large number of independent risks.[40] Because the insurer here assumes all the risk, and given the participation constraint, then $I(R) = R - rx - U^{-1}[\bar{u}]$, and the optimal x is that which maximizes the producer's expected profit.[41]

This set-up is drastically changed, and moral hazard problems arise, when the insurer contracts on a risk-averse producer whose inputs are unobservable (i.e., there is asymmetric information). This is because the insurer has but one instrument, the payoff schedule, to address two goals. To be attractive a contract must mitigate the uncertainty facing insurers, but to make a profit the contract must ensure that producers do not take advantage of the limited control over insurance payoffs that arise from the insurer's inability to observe input use. The insurer's problem when inputs are not observable, but the stochastic technology $F(R \mid x)$ is known, can be stated as

$$\begin{aligned} &\operatorname*{Max}_{I(R)} \int_a^b \big\{I(R) - C[I(R)]\big\}\, \mathrm{d}F(R \mid x) \quad \text{such that} \\ &\int_a^b U\big[R - I(R) - rx\big]\mathrm{d}F(R \mid x) \geqslant \bar{u}, \\ &x = \arg\max \int_a^b U\big[R - I(R) - rx\big]\mathrm{d}F(R \mid x). \end{aligned} \tag{5.14}$$

The additional incentive compatibility constraint ensures that the rational insurer endogenizes the input consequences of the payoff schedule posed. For both problems (5.12) and (5.14), in general the participation constraint is binding and the producer achieves utility level $\bar{u}$. Under moral hazard, however, it is not optimal for the risk-neutral principal to assume all risk. Some residual risk must be borne by the (risk-averse) producer and hence, to achieve a given $\bar{u}$, the expected payouts to the producer have to be larger than under symmetric information. Chambers (1989) discusses the welfare loss associated with the incentive constraint as well as the possibility that it might cause crop insurance markets to fail.

The implications of the moral hazard problem are not as clear-cut as intuition might suggest. Being relieved of some of the consequences of low input use, the producer may reduce input intensity. On the other hand, as previously shown, if input use is risk

---

[40] Unfortunately, risks across crop production units usually tend to be more systematic than idiosyncratic in nature.

[41] In the trivial case where inputs are unobservable but the producer is risk neutral, this expected profit-maximizing result may also be achieved by setting the schedule $I(R)$ equal to a constant. In this way, the producer faces all the consequences of the actions taken. But then, of course, the insurance company serves no purpose and will never be able to cover any administrative costs.

increasing then a high-risk environment may cause the producer to use fewer inputs than a lower-risk environment. Thus the existence of insurance may, in mitigating risk, encourage input use. That is, risk sharing and moral hazard effects may oppose each other.

To model econometrically the moral hazard problem, the crop producer contemplating whether to insure may be viewed as having to make two decisions: whether or not to insure, and the choice of input vector. In one of the first econometric analyses of the effects of crop insurance, Horowitz and Lichtenberg (1993) assumed that the decision to insure affects input use but not the other way around. Modeling the insurance decision by Probit analysis and modeling input choice as a linear regression on the insurance decision, among other regressors, they studied corn production decisions in ten Corn Belt states and concluded that the decision to insure increased significantly the use of nitrogen and pesticides. These results are somewhat surprising, so other researchers sought to confirm the conclusions on different data sets and using other methodologies. Smith and Goodwin (1996) estimated a simultaneous equations model of input use and crop insurance purchases for Kansas dryland wheat farmers, and concluded that insurance and input decisions are likely simultaneously determined. Further, their results suggest that insurance reduces the use of agricultural chemicals. Estimating an input-conditioned beta distribution for farm-level Iowa corn production, Babcock and Hennessy (1996) simulated optimal input use under different types and levels of insurance for risk-averse producers and also concluded that insurance would likely decrease input use. Although more empirical investigations are warranted, it would appear that risk sharing through crop insurance reduces input use.

The moral hazard problem was also studied in the West African Sahel region, which is at risk to drought. Following on work by Hazell (1992), among others, Sakurai and Reardon (1997) identified quite strong potential demand for area-level rainfall insurance. Their analysis also raises the concern that moral hazard arising from food aid could undermine the viability of such contracts.

In identifying two types of risk, production risk and land value risk arising from soil depletion, Innes and Ardila (1994) suggest an intertemporal environmental aspect to the incentive problem. For fragile land, a contract tailored to insure against production risk may exacerbate land value deterioration, and so one might not be able to ignore dynamic aspects of moral hazard. This is especially true if the operator does not own the land. Dynamic issues also arise in work by Coble et al. (1997) who find evidence that input reduction by insured producers occurs mainly when a crop loss is most likely, thus exacerbating the magnitude of the loss.

Moral hazard problems may not be confined to input intensity issues. If output is difficult to verify, then false yields may be reported. Such illegal acts raise questions concerning contract design, the structure of legal sanctions, and the nature of detection technologies. Hyde and Vercammen (1997) argue that, whereas it is difficult to motivate the structure of insurance contracts actually offered (i.e., the attributes of monotonicity, convexity, deductibility, and co-insurance) as a response to moral hazard on input use alone, actual contracts can plausibly be an optimal response to moral hazard on both input use and yield verification together.

5.2.2. Adverse selection

When, unlike the producer, the insurer is not completely informed about the nature of the risk being insured, then the insurer faces the problem of adverse selection. Ignoring input choices, let a risk-neutral insurer have categorized three production units owned by different operators and of equal size (say, one acre without loss of generality), A, B, and C, into the same risk cohort. From the information available to it, say common average yield ($\bar{y}$), the insurer can observe no difference among these three acres. In fact, the associated yield distributions differ; suppose all acres realize two outcomes, each with probability $1/2$, but the realizations for acre A are $\{\bar{y} - 10, \bar{y} + 10\}$, those for B are $\{\bar{y} - 20, \bar{y} + 20\}$, and those for C are $\{\bar{y} - 30, \bar{y} + 30\}$. With unit price, if the insurance payout equaled $\mathrm{Max}[\bar{y} - y, 0]$, then the expected payouts for acres A, B, and C would be 5, 10, and 15, respectively. In such a case, assuming full participation, the actuarially fair premium for a contract covering all three risks would be 10/acre. However, if the acre A producer is insufficiently risk averse, then she may conclude that the loss ratio for acre A, at $5/10 = ½$, is too low and may not insure the acre. If the insurer continues to charge 10/acre when covering only acres B and C, then an average loss of 22½/acre is incurred. On the other hand, if the premium is raised to 122½/acre so that a loss is avoided, then acre B may not be insured. Thus, the market may unravel in stages.

Avoiding adverse selection may require the successful crop insurance program to identify, acquire, and skillfully use data that discriminate among different risks. Although perhaps costly to implement, such data management procedures may be crucial because, unless rates are perceived as being acceptable, the market may collapse. The phenomenon of unravelling suggests that identifying a sufficiently large number of relatively homogeneous risks is a prerequisite for a successful contract. Useful discriminators would appear to include mean yield. Skees and Reed (1986) and Just and Calvin (1993) have found evidence suggesting that yield variance may decrease with increased mean yield, and so, even if the trigger insurance yield increases with mean yield, rates should probably be lower for more productive acres. Goodwin (1994), studying Kansas crops (1981–90), finds the relationship between yield variability and mean yield to be tenuous and suggests that farm yield histories be used to calculate yield variability rather than impute variability from historical mean yield. He also concludes that other factors, such as enterprise size, could be informative in setting premium rates.

The degree of homogeneity required to sustain the contract depends upon, among other things, the degree of risk aversion expressed by producers. The more risk averse the producers, the more tolerant they will be of actuarially unfair rates. In an investigation of adverse selection in contracts on corn production, Goodwin (1993) studied county-level data for the ninety-nine Iowa counties over the period 1985 to 1990 and found the elasticities of acreage insured to expected payoff to be in the range of 0.3–0.7. At the farm level, these elasticities may be higher. Further, he found that counties where the risk of payout is low are quite sensitive to the premium charged, so that an across-the-state (of Iowa) premium increase might not make corn yield insurance more profitable because substantial cancellations by the better risk prospects may occur. He concluded that the best approach to loss ratio reduction may involve fine-tuning the rate setting at the county or farm level.

Adverse selection may be either spatial or temporal in nature. The problem type discussed thus far may be categorized as being spatial in the sense that the factors differentiating risks occur at a given point in time. An alternative form of adverse selection, identified by Luo, Skees and Marchant (1994), may arise when attributes of a given risk vary temporally.[42] Coble et al. (1996) consider the case of adverse selection in crop insurance contracts for Kansas dryland wheat farmers over the years 1987 to 1990. Preseason rainfall was used as an indicator for intertemporal adverse selection whereby an unseasonably low (high) level of rainfall occurring before contract signing would entice marginal risks into (out of) signing, thus increasing the loss ratio if rates do not reflect the implications of the water deficit prevailing at signing. Although finding some evidence of adverse selection, they did not identify any of an intertemporal nature.

There are, of course, many factors other than adverse selection that determine the decision for, and the magnitude of, crop insurance participation. To understand adverse selection it is necessary to isolate its impact by accounting for other determinants of participation. In addition to the aforementioned research, econometric analyses of the determinants of insurance participation have been conducted by Gardner and Kramer (1986), Just and Calvin (1990), and Smith and Baquet (1996), among others. Although the conclusions are somewhat mixed, an overview of results suggests that participation tends to increase with farm size. This may be because of the negative correlation between farm size and the importance of off-farm income, or because of increased borrowing. Also, enterprise specialization tends to increase participation, presumably because of increased risk exposure. Further, and suggestive of adverse selection, higher yield variability land is more likely to be insured. However, estimates by Coble et al. (1996) infer that this is true even if rates account for the increased riskiness.

5.2.3. *Further discussion*

Though conceptually distinct, the differences between the moral hazard and adverse selection problems often disappear in practice. Noting that both moral hazard and adverse selection are problems of information asymmetry, Quiggin, Karagiannis and Stanton (1993) posed the situation in which a wheat and corn producer contemplating crop insurance has one acre of good land and one acre of bad land. Given the decision to insure wheat but not corn, the planting of wheat on poor quality land might be viewed as moral

---

[42] If the producer is better informed about the temporal evolution of risk, then adverse selection may occur. However, as discussed in [Knight and Coble (1997)], the insurer may be just as informed about the temporal risk as the producer, but may be either unable or unwilling to adjust rates. In such a situation, the problem is not one of adverse selection.

hazard. However, given the decision to plant poor land to wheat, the decision to insure wheat only may be viewed as adverse selection. Thus, it should be no surprise that the potential remedies to each problem are similar.

Due to the informational nature of the main barriers to successful crop insurance markets, the obvious solution is, where feasible, to acquire and use as much information as marginal cost and profit considerations allow. To improve performance by reducing adverse selection, the FCIC changed its approach to rate setting in 1985 to accommodate additional information. Subsequent contracts changed the determination of the insurable yield from an average of past yields observed in a locality to an average of past yields observed on the farm in question. Even so sensible a reform, however, may give rise to incentive problems. As pointed out by Vercammen and van Kooten (1994), producers might manipulate input use in a cyclical manner to build up insurable yield levels before cashing in (in a probabilistic sense) by reducing input use for a few years.

On the other hand, area yield insurance [Halcrow (1949), Miranda (1991), Mahul (1999)], where indemnities are based upon the average yield of a suitably wide area (say, a county), eliminates the moral hazard problem and may reduce or eliminate adverse selection. In addition, just as futures markets permit hedge ratios in excess of one, a producer may take out an arbitrary level of area yield insurance coverage without giving rise to concerns about increased moral hazard. Area yield insurance rates are likely to be lower than farm-specific rates because an area yield index will usually be less variable than yield on a given farm. However, because farm-specific risks are not insured, producers may continue to be subjected to some (possibly substantial) production risk.

Revenue insurance is a recurrently popular concept because it directly addresses the income risk problem facing producers. A further possible advantage is that, in combining price and yield insurance, the approach may mitigate somewhat the incidence of moral hazard and adverse selection. Miranda and Glauber (1991), as well as Babcock and Hennessy (1996), conducted simulation analyses for U.S. crop production, and Turvey (1992a, 1992b) studied the costs and benefits of such a program in Canada. The potential for revenue insurance arises from the fact that, even together, price contingent markets (for a fixed quantity) and yield contingent markets (for a fixed price) are not likely to fully stabilize income. Hennessy, Babcock and Hayes (1997) have shown that this targeting attribute of revenue insurance means that it can increase the welfare impact of a given expenditure on income support relative to various alternatives of price and yield support.

Compulsory insurance has often been proposed to eliminate the political need for continual ex-post interventions. If adverse selection is a major problem in competitive insurance markets, however, then compulsory insurance is unlikely to gain the political support necessary for a long-term solution. More effective re-insurance on the part of crop insurers may facilitate the reduction of market rates, and thus reduce adverse selection, because systemic risk is pervasive in the insurance of crop risks and so pooling is largely ineffective for the insurer [Miranda and Glauber (1997), Duncan and Myers (1997)]. Given the diminishing importance of agriculture in developed economies, the introduction of crop loss risks into a well-diversified portfolio of risks would reduce the high level of systematic risk in crop insurance markets, and so may reduce the risk premia required by crop insurers. But crop insurance differs in many ways from other forms of insurance, and it may prove difficult to entice reinsurers into accepting these contracts. If a permanent solution exists that is politically more acceptable than a laissez-faire market approach, it may involve a package of reforms that is balanced to mitigate the incentive impacts but incurs low budgetary costs. Such a package should also take care not to undermine existing or potentially viable risk markets. Finally, the policy mix must be flexible because the technology and organization of crop production may undergo fundamental changes in the coming years.

6. Conclusion

It is abundantly clear that considerations of uncertainty and risk cannot be escaped when addressing most agricultural economics problems. The demands imposed on economic analysis are complex and wide-ranging, with issues that extend from the pure theory of rational behavior to the practicality of developing risk-management advice. The economics profession at large, and its agricultural economics subset, has responded to this challenge with a wealth of contributions. In this chapter we have emphasized theoretical and applied analyses as they pertain to production decisions at the farm level. The EU model provides the most common approach to characterizing rational decisions under risk, and it has been the framework of choice for most applied work in agricultural economics. Whereas our review has provided only a nutshell exposition of the framework's main features, the careful student will dig deeper into its axiomatic underpinning as a crucial step to appreciating what modeling decisions under risk means. More generally, we can note that a satisfactory model of decision making under risk requires assuming an extended notion of rationality. Agents need to know the entire distribution of risky variables, and need to take into account how this randomness affects the distribution of outcomes over alternative courses of action. Thus, the decision maker's problem is inherently more difficult under uncertainty than under certainty.

Because the notion of rational behavior under risk arguably requires agents to solve a complex problem, it is perhaps useful to distinguish between whether models are meant to provide a *positive theory* (aiming to describe how agents actually make decisions under risk) or a *normative theory* (the purpose of which is to prescribe a rational course of action for the particular risky situation). This distinction is admittedly somewhat artificial, and most models are suitable to either interpretation. Yet being more explicit about whether one's analysis is pursuing a positive or normative exercise is possibly quite important in applied contexts such as those covered in this chapter. Much agricultural risk management work is meant as a normative activity, and this may have implications for the choice of models. For instance, the EU model has been criticized, on positive grounds, for failing to describe accurately how agents actually behave under risk in some situations; such a critique, of course, says nothing about the suitability of the EU model for normative (prescriptive) purposes.

Models of decision making under risk bring to the forefront the fact that decisions will be affected in a crucial way by the agent's preferences, i.e., her attitudes towards risk. Consequently, it is quite important to quantify the degree of agricultural producers' risk aversion, and a number of studies have endeavored to do just that. The conclusions may be summarized as follows: within the EU framework, producers typically display some aversion to risk, and risk preferences probably conform to DARA. But evidence on the magnitude of risk aversion is less conclusive and falls short of providing useful parameters that are critical for normative statements (whether in terms of risk management advice to farmers or in terms of suggesting desirable government policies).

Considerations of risk aversion also raise concerns about a very common attribute of applied studies that have a positive orientation. Namely, whereas theoretical models are meant for individual decision making, empirical models are often implemented with aggregate data. The danger of ignoring the implicit aggregation problem is obviously a general concern that applies to economic models of certainty as well. But the fact that risk attitudes play an important role in models with risk, and given that such preferences are inherently an individual attribute, suggests that agents' heterogeneity is bound to be more important when risk matters. It seems that more can and should be done to tackle aggregation considerations in a satisfactory manner.

The complexities of the decision maker's problem under risk raise additional issues for the applied researcher. Agents' beliefs about the characteristics of uncertainty are obviously crucial in this context. The EU model, by relying on the notion of subjective probabilities, neatly solves the theoretical modeling question. But the applied researcher may need to model explicitly how the agent makes probability assessments (i.e., to model her expectations). Whereas the rational expectation hypothesis provides perhaps the most ambitious answer to this question, it is informationally very demanding when (as is typically the case in risky situations) the entire distribution of the random variables matters. This raises the question of whether rational expectations are legitimate from a theoretical point of view, but also implies that empirical models that wish to implement rational expectations can be computationally quite demanding, even for the simplest model under risk. Indeed, many empirical models reviewed in this chapter appear somewhat oversimplified. The modus operandi seems to be to allow theoretical modeling to be as sophisticated as desired but to keep empirical models as simple as possible. Such oversimplifications naturally beg the question of the relationship of empirical models to the theoretical constructs that are used to interpret results, and raise some concerns about what exactly we can learn from this body of empirical studies.

Notwithstanding the remaining criticisms and concerns that one may have, the studies surveyed in this chapter have addressed an important set of problems. Uncertainty and risk are essential features of many agricultural activities, and have important consequences for the agents involved and for society at large. Although welfare and policy considerations related to risk are discussed elsewhere in this *Handbook*, we should note that the economic implications of the existence of risk and uncertainty are related to the particular institutional setting in which agents operate. Insofar as the set of relevant markets is not complete, then this market incompleteness has the potential of adversely affecting resource allocation, as well as resulting in less than optimal allocation of risk-bearing. Indeed, the incompleteness of risk markets for agricultural producers has often been cited as a motivation for agricultural policies in many developed countries. But arguably neither existing markets nor government policies have solved the farmers' risk exposure problems. Risk continues to have the potential of adversely affecting farmers' welfare, as well as carrying implications for the long-run organization of agricultural production and for the structure of resource ownership in the agricultural sector.

Acknowledgements

We appreciate the comments of many colleagues including Jock Anderson, Keith Coble, Sergio Lence, Mario Miranda, Bob Myers, and the Handbook's editors. This chapter was largely written in 1997, and slightly revised in 1998. We have not attempted to cover work published more recently. Other unhedged errors and omissions are the responsibility of the authors. Journal paper No. J-17995 of the Iowa Agriculture and Home Economics Experiment Station, Ames Iowa. Project No. 3463, supported by Hatch Act and State of Iowa funds.

References

Allais, M. (1953), "Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école américaine", Econometrica 21:503–546.

Anderson, G. (1996), "Nonparametric tests of stochastic dominance in income distributions", Econometrica 64:1183–1193.

Anderson, J.R., J.L. Dillon and B. Hardaker (1977), Agricultural Decision Analysis (Iowa State University Press, Ames, IA).

Anderson, R., and J.-P. Danthine (1981), "Cross-hedging", Journal of Political Economy 89:1182–1196.

Anderson, R., and J.-P. Danthine (1983), "The time pattern of hedging and the volatility of futures prices", Review of Economic Studies 50:249–266.

Anscombe, F., and R. Aumann (1963), "A definition of subjective probability", Annals of Mathematical Statistics 34:199–205.

Antle, J.M. (1983), "Testing the stochastic structure of production: A flexible moment-based approach", Journal of Business and Economic Statistics 1 (3):192–201.

Antle, J.M. (1987), "Econometric estimation of producers' risk attitudes", American Journal of Agricultural Economics 69:509–522.

Antle, J.M. (1989), "Nonstructural risk attitude estimation", American Journal of Agricultural Economics 71:774–784.

Antle, J.M., and C.C. Crissman (1990), "Risk, efficiency, and the adoption of modern crop varieties: Evidence from the Philippines", Economic Development and Cultural Change 38:517–537.

Antle, J.M., and W.J. Goodger (1984), "Measuring stochastic technology: The case of Tulare milk production", American Journal of Agricultural Economics 66:342–350.

Antonovitz, F., and R. Green (1990), "Alternative estimates of fed beef supply response to risk", American Journal of Agricultural Economics 72:475–487.

Aradhyula, S.V., and M.T. Holt (1989), "Risk behavior and rational expectations in the U.S. broiler market", American Journal of Agricultural Economics 71:892–902.

Aradhyula, S.V., and M.T. Holt (1990), "Price risk in supply equations: An application of GARCH time-series models to the U.S. broiler market", Southern Economic Journal 57:230–242.

Arrow, K. (1964), "The role of securities in the optimal allocation of risk bearing", Review of Economic Studies 31:91–96.

Arrow, K. (1965), Aspects of the Theory of Risk-bearing (Yrio Jahnsson Foundation, Helsinki).

Azam, J.-P. (1996), "The impact of floods on the adoption rate of high-yielding rice varieties in Bangladesh", Agricultural Economics 13:179–189.

Babcock, B.A., and A.M. Blackmer (1992), "The value of reducing temporal input nonuniformities", Journal of Agricultural and Resource Economics 17:335–347.

Babcock, B.A., and D.A. Hennessy (1996), "Input demand under yield and revenue insurance", American Journal of Agricultural Economics 78:416–427.

Baron, D.P. (1970), "Price uncertainty, utility, and industry equilibrium in pure competition", International Economic Review 11:463–480.

Barrett, C.B. (1996), "On price risk and the inverse farm size-productivity relationship", Journal of Development Economics 51:193–215.

Basu, K. (1992), "Limited liability and the existence of share tenancy", Journal of Development Economics 38:203–220.

Batlin, C.A. (1983), "Production under price uncertainty with imperfect time hedging opportunities in futures markets", Southern Economic Journal 49:681–692.

Batra, R.N., and A. Ullah (1974), "Competitive firm and the theory of input demand under price uncertainty", Journal of Political Economy 82:537–548.

Behrman, J.R. (1968), Supply Response in Underdeveloped Agriculture: A Case Study of Four Major Crops in Thailand, 1937–63 (North Holland, Amsterdam).

Benninga, S., R. Eldor and I. Zilcha (1983), "Optimal hedging in the futures market under price uncertainty", Economics Letters 13:141–145.

Bera, A.K. and T.G. Kelley (1990), "Adoption of high-yielding rice varieties in Bangladesh", Journal of Development Economics 33:263–285.

Bernoulli, D. (1738), "Specimen theoriae novae de mensura sortis", Commentarii Academiae Scientiarum Imperialis Petropolitanae 5:175–192 (Translation by Sommer, L., 1954, and published in Econometrica 22:23–36).

Bigman, D. (1996), "Safety-first criteria and their measures of risk", American Journal of Agricultural Economics 78:225–235.

Binswanger, H.P. (1980), "Attitudes toward risk: Experimental measurement in rural India", American Journal of Agricultural Economics 62:395–407.

Binswanger, H.P. (1981), "Attitudes toward risk: Theoretical implications of an experiment in rural India", Economic Journal 91:867–890.

Bollerslev, T.P. (1986), "Generalized autoregressive conditional heteroskedasticity", Journal of Econometrics 31:307–327.

Borch, K. (1962), "Equilibrium in a reinsurance market", Econometrica 30:424–444.

Bottomley, A. (1975), "Interest rate development in underdeveloped rural areas", American Journal of Agricultural Economics 57:279–291.

Bray, M. (1981), "Futures trading, rational expectations, and the efficient market hypothesis", Econometrica 49:575–596.

Brink, L., and B.A. McCarl (1978), "The tradeoff between expected return and risk among cornbelt farmers", American Journal of Agricultural Economics 60:259–263.

Brorsen, B.W., J.-P. Chavas and W.R. Grant (1987), "A market equilibrium analysis of the impact of risk on the U.S. rice industry", American Journal of Agricultural Economics 69:731–739.

Burt, O.R. (1965), "Optimal replacement under risk", Journal of Farm Economics 47:324–346.

Chambers, R.G. (1989), "Insurability and moral hazard in agricultural insurance markets", American Journal of Agricultural Economics 71:604–616.

Chambers, R.G., and J. Quiggin (1998), "Cost functions and duality for stochastic technologies", American Journal of Agricultural Economics 80:288–295.

Chavas, J.-P. (1994), "Production and investment decisions under sunk cost and temporal uncertainty", American Journal of Agricultural Economics 76:114–127.

Chavas, J.-P., and M.T. Holt (1990), "Acreage decisions under risk: The case of corn and soybeans", American Journal of Agricultural Economics 72:529–538.

Chavas, J.-P., and M.T. Holt (1996), "Economic behavior under uncertainty: A joint analysis of risk preferences and technology", Review of Economics and Statistics 78:329–335.

Chavas, J.-P., and R. Pope (1985), "Price uncertainty and competitive firm behavior: Testable hypotheses from expected utility maximization", Journal of Economics and Business 37:223–235.

Coble, K.H., T.O. Knight, R.D. Pope and J.R. Williams (1996), "Modeling farm-level crop insurance demand with panel data", American Journal of Agricultural Economics 78:439–447.

Coble, K.H., T.O. Knight, R.D. Pope and J.R. Williams (1997), "An expected indemnity approach to the measurement of moral hazard in crop insurance", American Journal of Agricultural Economics 79:216–226.

Collins, R.A., and P.J. Barry (1986), "Risk analysis with single-index portfolio models: An application to farm planning", American Journal of Agricultural Economics 68:152–161.

Cox, J.C., and M. Rubinstein (1985), Options Markets (Prentice-Hall, Englewood Cliffs, N.J.).

Coyle, B.T. (1992), "Risk aversion and price risk in duality models of production: A linear mean-variance approach", American Journal of Agricultural Economics 74:849–859.

Dalal, A.J. (1990), "Symmetry restrictions in the analysis of the competitive firm under price uncertainty", International Economic Review 31:207–211.

Danthine, J.-P. (1978), "Information, futures prices, and stabilizing speculation", Journal of Economic Theory 17:79–98.

Day, R. (1965), "Probability distributions of field crop yields", Journal of Farm Economics 47:713–741.

Debreu, G. (1959), Theory of Value: An Axiomatic Analysis of General Equilibrium (Yale University Press, New Haven).

de Finetti, B. (1931), "Sul significato soggettivo della probabilità", Fundamenta Mathematicae 17:298–229.

Dillon, J.L. (1971), "An expository review of Bernouillian decision theory in agriculture: Is utility futility?", Review of Marketing and Agricultural Economics 39:3–80.

Dillon, J.L., and P.L. Scandizzo (1978), "Risk attitudes of subsistence farmers in Northeast Brazll: A sampling approach", American Journal of Agricultural Economics 60:425–435.

DiPietre, D.D., and M.L. Hayenga (1982), "Cross-hedging wholesale products using live hogs futures", American Journal of Agricultural Economics 64:747–751.

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

Duncan, J., and R.J. Myers (1997), Crop insurance under catastrophic risk, Staff Paper 97-10 (Department of Agricultural Economics, Michigan State University).

Eeckhoudt, L., and P. Hansen (1992), "Mean-preserving changes in risk with tail-dominance", Theory and Decision 33:23–39.

Ehrlich, I., and G. Becker (1972), "Market insurance, self-insurance, and self-protection", Journal of Political Economy 80:623–648.

Ellsberg, D. (1961), "Risk, ambiguity, and the savage axioms", Quarterly Journal of Economics 75:643–669.

Engle, R.F. (1982), "Autoregressive conditional heteroskedasticity with estimates of the variance of the United Kingdom inflation", Econometrica 50:987–1007.

Engle, R.F., and T. Bollerslev (1986), "Modelling the persistence of conditional variances", Econometric Reviews 5:1–50.

Epstein, L.G. (1978), "Production flexibility and the behavior of the competitive firm under price uncertainty", Review of Economic Studies 45:251–261.

Feder, G., R.E. Just and A. Schmitz (1980), "Futures markets and the theory of the firm under price uncertainty", Quarterly Journal of Economics 94:317–328.

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

Federal Crop Insurance Corporation (1997), Crop insurance actuarial information public access server, World Wide Web site http://www.act.fcic.usda.gov/./actuarial/

Feinerman, E., E.K. Choi and S.R. Johnson (1990), "Uncertainty and split nitrogen application in corn production", American Journal of Agricultural Economics 72:975–984.

Finkelshtain, I., and J.A. Chalfant (1991), "Marketed surplus under risk: Do peasants agree with Sandmo?", American Journal of Agricultural Economics 73:557–567.

Foster, W.E., and G.C. Rausser (1991), "Farm behavior under risk of failure", American Journal of Agricultural Economics 73:276–288.

Freund, R.J. (1956), "Introduction of risk into a programming model", Econometrica 24:253–263.

Froot, K.A., D.S. Scharfstein and J.C. Stein (1993), "Risk management: Coordinating corporate investment and financing policies", Journal of Finance 48:1629–1658.

Gallagher, P.W. (1987), "U.S. soybean yields: Estimation and forecasting with nonsymmetric disturbances", American Journal of Agricultural Economics 69:796–803.

Gardner, B.L. (1989), "Rollover hedging and missing long-term futures markets", American Journal of Agricultural Economics 71:311–325.

Gardner, B.L. (1994), "Crop insurance in U.S. farm policy", in: D.L. Hueth, and W.H. Furtan, eds., Economics of Agricultural Crop Insurance: Theory and Evidence (Kluwer Academic Publishers, Boston) 17–44.

Gardner, B.L., and R.A. Kramer (1986), "Experience with crop insurance programs in the United States", in: P. Hazell, C. Pomareda and A. Valdés, eds., Crop Insurance for Agricultural Development: Issues and Experience (Johns Hopkins Press, Baltimore) 195–222.

Geweke, J. (1986), "Modelling the persistence of conditional variances: Comment", Econometric Reviews 5:57–61.

Gollier, C. (1995), "The comparative statics of changes in risk revisited", Journal of Economic Theory 66:522–535.

Goodwin, B.K. (1993), "An empirical analysis of the demand for crop insurance", American Journal of Agricultural Economics 75:425–434.

Goodwin, B.K. (1994), "Premium rate determination in the Federal crop insurance program: What do averages have to say about risk?", Journal of Agricultural and Resource Economics 19:382–395.

Hadar, J., and W. Russell (1969), "Rules for ordering uncertain prospects", American Economic Review 59:25–34.

Halcrow, H.G. (1949), "Actuarial structures for crop insurance", Journal of Farm Economics 21:418–443.

Hamal, K.B., and J.R. Anderson (1982), "A note on decreasing absolute risk aversion among farmers in Nepal", Australian Journal of Agricultural Economics 26:220–225.

Hanoch, G., and H. Levy (1969), "The efficiency analysis of choices involving risk", Review of Economic Studies 36:335–346.

Hartman, R. (1976), "Factor demand with output price uncertainty", American Economic Review 66:675–681.

Hazell, P.B.R. (1971), "A linear alternative to quadratic and semivariance programming for farm planning under uncertainty", American Journal of Agricultural Economics 53:53–62.

Hazell, P.B.R. (1982), "Applications of risk preference estimates in firm-household and agricultural sector models", American Journal of Agricultural Economics 64:384–390.

Hazell, P.B.R. (1992), "The appropriate role of agricultural insurance in developing countries", Journal of International Development 4:567–582.

Hennessy, D.A. (1998), "The production effects of income support under price uncertainty", American Journal of Agricultural Economics 80:46–57.

Hennessy, D.A., B.A. Babcock and D.J. Hayes (1997), "Budgetary and producer welfare effects of revenue insurance", American Journal of Agricultural Economics 79:1024–1034.

Hirshleifer, J. (1966), "Investment decisions under uncertainty: Applications of the state-preference approach", Quarterly Journal of Economics 80:252–277.

Hirshleifer, J., and J.G. Riley (1992), The Analytics of Uncertainty and Information (Cambridge University Press, Cambridge, U.K.).

Holt, M.T. (1993), "Risk response in the beef marketing channel: A multivariate generalized ARCH-M approach", American Journal of Agricultural Economics 75:559–571.

Holt, M.T., and G. Moschini (1992), "Alternative measures of risk in commodity supply models: An analysis of sow farrowing decisions in the United States", Journal of Agricultural and Resource Economics 17:1–12.

Holthausen, D.M. (1979), "Hedging and the competitive firm under price uncertainty", American Economic Review 69:989–995.

Horowitz, J.K., and E. Lichtenberg (1993), "Insurance, moral hazard, and chemical use in agriculture", American Journal of Agricultural Economics 75:926–935.

Hurt, C.A., and P. Garcia (1982), "The impact of price risk on sow farrowings, 1967–78", American Journal of Agricultural Economics 64:563–568.

Hyde, C.E., and J.A. Vercammen (1997), "Costly yield verification, moral hazard, and crop insurance contract form", Journal of Agricultural Economics 48:393–407.

Innes, R. (1990), "Government target price interventions in economies with incomplete markets", Quarterly Journal of Economics 106:1035–1052.

Innes, R., and S. Ardila (1994), "Agricultural insurance and soil depletion in a simple dynamic model", American Journal of Agricultural Economics 76:371–384.

Innes, R., and G.C. Rausser (1989), "Incomplete markets and government agricultural policy", American Journal of Agricultural Economics 71:915–931.

Ishii, Y. (1977), "On the theory of the competitive firm under price uncertainty: Note", American Economic Review 67:768–769.

Jewitt, I. (1987), "Risk aversion and the choice between risky prospects: The preservation of comparative statics results", Review of Economic Studies 56:14–25.

Just, R.E. (1974), "An investigation of the importance of risk in farmer's decisions", American Journal of Agricultural Economics 56:14–25.

Just, R.E. (1976), "Estimation of a risk response model with some degree of flexibility", Southern Economic Journal 42:675–684.

Just, R.E., and L. Calvin (1990), "An empirical analysis of U.S. participation in crop insurance", Unpublished report to the Federal Crop Insurance Corporation.

Just, R.E., and L. Calvin (1993), "Adverse selection in U.S. crop insurance: The relationship of farm characteristics to premiums", Unpublished manuscript, University of Maryland.

Just, R.E., and R.D. Pope (1978), "Stochastic specifications of production functions and economic implications", Journal of Econometrics 7:67–86.

Just, R.E., and R.D. Pope (1979), "Production function and estimation and related risk considerations", American Journal of Agricultural Economics 61:276–284.

Just, R.E., and Q. Weninger (1999), "Are crop yields normally distributed?", American Journal of Agricultural Economics 81:287–304.

Just, R.E., and D. Zilberman (1983), "Stochastic structure, farm size and technology adoption in developing agriculture", Oxford Economic Papers 35:307–328.

Kahneman, D., and A. Tversky (1979), "Prospect theory: An analysis of decision under risk", Econometrica 47:263–291.

Karp, L.S. (1988), "Dynamic hedging with uncertain production", International Economic Review 29:621–637.

Karp, L.S., and A. Pope (1984), "Range management under uncertainty", American Journal of Agricultural Economics 66:437–446.

Katz, E. (1981), "A note on a comparative statics theorem for choice under risk", Journal of Economic Theory 25:318–319.

Kimball, M.S. (1990). "Precautionary saving in the small and in large", Econometrica 58:53–73.

Knight, F. (1921), Risk, Uncertainty and Profit (Houghton Mifflin, Boston).

Knight, T.O., and K.H. Coble (1997), "Survey of U.S. multiple peril crop insurance literature since 1980", Review of Agricultural Economics 19:128–156.

Krause, M. (1979), "A comparative static theorem for choice under risk", Journal of Economic Theory 21:510–517.

Lapan, H., and G. Moschini (1994), "Futures hedging under price, basis, and production risk", American Journal of Agricultural Economics 76:465–477.

Lapan, H., and G. Moschini (1996), "Optimal price policies and the futures markets", Economics Letters 53:175–182.

Lapan, H., G. Moschini and S. Hanson (1991), "Production, hedging and speculative decisions with options and futures markets", American Journal of Agricultural Economics 72:66–74.

Lapan, H., G. Moschini and S. Hanson (1993), "Production, hedging and speculative decisions with options and futures markets: Reply", American Journal of Agricultural Economics 75:748–750.

Leahy, J.V. (1993), "Investment in competitive equilibrium: The optimality of myopic behavior", Quarterly Journal of Economics 108:1105–1133.

Lence, S.H. (1995), "On the optimal hedge under unbiased futures prices", Economics Letters 47:385–388.

Lence, S.H. (1996), "Relaxing the assumptions of minimum-variance hedge", Journal of Agricultural and Resource Economics 21:39–55.

Lin, W. (1977), "Measuring aggregate supply response under instability", American Journal of Agricultural Economics 59:903–907.

Lin, W., G. Dean and C. Moore (1974), "An empirical test of utility vs. profit maximization in agricultural production", American Journal of Agricultural Economics 56:497–508.

Losq, E. (1982), "Hedging with price and output uncertainty", Economic Letters 10:65–70.

Love, H.A., and S.T. Buccola (1991), "Joint risk preference-technology estimation with a primal system", American Journal of Agricultural Economics 73:765–774.

Luo, H., J.R. Skees and M.A. Marchant (1994), "Weather information and the potential for inter-temporal adverse selection", Review of Agricultural Economics 16:441–451.

Machina, M.J. (1987), "Choice under uncertainty: Problems solved and unsolved", Journal of Economic Perspectives 1:121–154.

MacMinn, R.D., and A.G. Holtmann (1983), "Technological uncertainty and the theory of the firm", Southern Economic Journal 50:120–136.

Mahul, O. (1999), "Optimum area yield crop insurance", American Journal of Agricultural Economics 81.

Marcus, A.J., and D.M. Modest (1984), "Futures markets and production decisions", Journal of Political Economy 92:409–426.

Markowitz, H. (1952), "Portfolio selection", Journal of Finance 7:77–91.

Mas-Colell, A., M.D. Whinston and J.R. Green (1995), Microeconomic Theory (Oxford University Press, Oxford).

McKinnon, R.I. (1967), "Futures markets, buffer stocks, and income instability for primary producers", Journal of Political Economy 75:844–861.

Menger, K. (1934), "Das unsicherheitsmoment in der wertlehre", Zeitschrift für Nationalökonomie. Translation in: M. Shubik, ed., 1967, Essays in Mathematical Economics in Honor of Oskar Morgenstern (Princeton University Press, Princeton).

Meyer, J. (1987), "Two-moment decision models and expected utility maximization", American Economic Review 77:421–430.

Meyer, J., and M.B. Ormiston (1983), "The comparative statics of cumulative distribution function changes for the class of risk averse agents", Journal of Economic Theory 31:153–169.

Meyer, J., and M.B. Ormiston (1985), "Strong increases in risk and their comparative statics", International Economic Review 26:425–437.

Meyer, J., and M.B. Ormiston (1989), "Deterministic transformations of random variables and the comparative statics of uncertainty", Journal of Risk and Uncertainty 2:179–188.

Miranda, M.J. (1991), "Area-yield crop insurance reconsidered", American Journal of Agricultural Economics 73:233–242.

Miranda, M.J., and J.W. Glauber (1991), "Providing crop disaster assistance through a modified deficiency payment program", American Journal of Agricultural Economics 73:1233–1243.

Miranda, M.J., and J.W. Glauber (1993), "Estimation of dynamic nonlinear rational expectations models of primary commodity markets with private and government stockholding", Review of Economics and Statistics 75:463–470.

Miranda, M.J., and J.W. Glauber (1997), "Systemic risk, reinsurance, and the failure of crop insurance markets", American Journal of Agricultural Economics 79:206–215.

Moscardi, E., and A. de Janvry (1977), "Attitudes toward risk among peasants: An econometric approach", American Journal of Agricultural Economics 59:710–716.

Moschini, G. (1999), Production Risk and the Estimation of Ex-ante Cost Functions, Staff Paper # 315 (Department of Economics, Iowa State University). Forthcoming, Journal of Econometrics, 2001.

Moschini, G., and H. Lapan (1992), "Hedging price risk with options and futures for the competitive firm with production flexibility", International Economic Review 33:607–618.

Moschini, G., and H. Lapan (1995), "The hedging role of options and futures under joint price, basis, and production risk", International Economic Review 36:1025–1049.

Moss, C.B., and J.S. Shonkwiler (1993), "Estimating yield distributions with a stochastic trend and nonnormal errors", American Journal of Agricultural Economics 75:1056–1062.

Myers, R.J. (1988), "The value of ideal contingency markets in agriculture", American Journal of Agricultural Economics 70:255–267.

Myers, R.J. (1989), "Econometric testing for risk averse behavior in agriculture", Applied Economics 21:542–552.

Myers, R.J., and S.D. Hanson (1996), "Optimal dynamic hedging in unbiased futures markets", American Journal of Agricultural Economics 78:13–20.

Myers, R.J., and S.R. Thompson (1989), "Generalized optimal hedge ratio estimation", American Journal of Agricultural Economics 71:858–868.

Nelson, D.B. (1990), "Stationarity and persistence in the GARCH(1,1) model", Econometric Theory 6:318–334.

Nelson, C.H., and P.V. Preckel (1989), "The conditional beta distribution as a stochastic production function", American Journal of Agricultural Economics 71:370–378.

Nerlove, M. (1958), The Dynamics of Supply: Estimation of Farmers Response to Price (Johns Hopkins University Press, Baltimore).

Newbery, D.M. (1988), "On the accuracy of the mean-variance approximation for futures markets", Economics Letters 28:63–68.

Newbery, D.M.G., and J.E. Stiglitz (1981), The Theory of Commodity Price Stabilization – A Study in the Economics of Risk (Clarendon Press, Oxford).

O'Donnell, C.J., and A.D. Woodland (1995), "Estimation of Australian wool and lamb production technologies under uncertainty: An error components approach", American Journal of Agricultural Economics 77:552–565.

Ormiston, M.B. (1992), "First and second degree transformations and comparative statics under uncertainty", International Economic Review 33:33–44.

Pagan, A.R., and G.W. Schwert (1990), "Alternative models for conditional stock volatility", Journal of Econometrics 45:267–290.

Paris, Q. (1989), "Broken symmetry, symmetry, and price uncertainty", European Economic Review 33:1227–1239.

Pesaran, M.H. (1987), Limits to Rational Expectations (Blackwell Press, Oxford).

Pitt, M.M., and G. Sumodiningrat (1991), "Risk, schooling and the choice of seed technology in developing countries: A meta-profit function approach", International Economic Review 32:457–473.

Pope, R.D. (1978), "The expected utility hypothesis and demand-supply restrictions", American Journal of Agricultural Economics 60:619–627.

Pope, R.D. (1980), "The generalized envelope theorem and price uncertainty", International Economic Review 21:75–86.

Pope, R.D. (1988), "A new parametric test for the structure of risk preferences", Economics Letters 27:117–121.

Pope, R.D., and J.-P. Chavas (1994), "Cost functions under production uncertainty", American Journal of Agricultural Economics 76:196–204.

Pope, R.D., and R.E. Just (1991), "On testing the structure of risk preferences in agricultural supply analysis", American Journal of Agricultural Economics 73:743–748.

Pope, R.D., and R.E. Just (1996), "Empirical implementation of ex-ante cost functions", Journal of Econometrics 72:231–249.

Pope, R.D., and R.A. Kramer (1979), "Production uncertainty and factor demands for the competitive firm", Southern Economic Journal 46:489–501.

Pratt, J. (1964), "Risk aversion in the small and in the large", Econometrica 32:122–236.

Purvis, A., W.G. Boggess, C.B. Moss and J. Holt (1995), "Technology adoption decisions under irreversibility and uncertainty: An *ex ante* approach", American Journal of Agricultural Economics 77:541–551.

Pyle, D.H., and S.J. Turnovsky (1970), "Safety first and expected utility maximization in mean-standard deviation portfolio analysis", Review of Economics and Statistics 52:75–81.

Quiggin, J. (1993), Generalized Expected Utility Theory – The Rank-Dependent Model (Kluwer Academic Publishers, Boston).

Quiggin, J., G. Karagiannis and J. Stanton (1993), "Crop insurance and crop production: An empirical study of moral hazard and adverse selection", Australian Journal of Agricultural Economics 37:95–113.

Ramaswami, B. (1992), "Production risk and optimal input decisions", American Journal of Agricultural Economics 74:860–869.

Ramírez, O.A. (1997), "Estimation and use of a multivariate parametric model for simulating heteroskedastic, correlated, nonnormal random variables: The case of corn belt corn, soybean, and wheat yields", American Journal of Agricultural Economics 79:191–205.

Rausser, G.C., and E. Hochman (1979), Dynamic Agricultural Systems: Economic Prediction and Control (North-Holland, New York).

Ray, D. (1998), Development Economics (Princeton University Press, Princeton NJ).

Robison, L.J., and P.J. Barry (1987), The Competitive Firm's Response to Risk (MacMillan, New York).

Rolfo, J. (1980), "Optimal hedging under price and quantity uncertainty: The case of a cocoa producer", Journal Political Economy 88:100–116.

Rosenzweig, M.R., and H.P. Binswanger (1993), "Wealth, weather risk and the composition and profitability of agricultural investments", Economic Journal 103:56–78.

Rothschild, M., and J.E. Stiglitz (1970), "Increasing risk: I. A definition", Journal of Economic Theory 2:225–243.

Rothschild, M., and J.E. Stiglitz (1971), "Increasing risk: II. Its economic consequences", Journal of Economic Theory 3:66–84.

Roumasset, J.A., M.W. Rosegrant, U.N. Chakravorty and J.R. Anderson (1989), "Fertilizer and crop yield variability: A review", in: J.R. Anderson, and P.B.R. Hazell, eds., Variability in Grain Yields (Johns Hopkins University Press, Baltimore) 223–233.

Saha, A. (1993), "Expo-power utility: A 'flexible' form for absolute and relative risk aversion", American Journal of Agricultural Economics 75:905–913.

Saha, A., C.R. Shumway and H. Talpaz (1994), "Joint estimation of risk preference structure and technology using expo-power utility", American Journal of Agricultural Economics 76:173–184.

Sakurai, T., and T. Reardon (1997), "Potential demand for drought insurance in Burkina Faso and its determinants", American Journal of Agricultural Economics 79:1193–1207.

Samuelson, P. (1952), "Probability, utility, and the independence axiom", Econometrica 20:670–678.

Sandmo, A. (1971), "On the theory of the competitive firm under price uncertainty", American Economic Review 61:65–73.

Savage, L.J. (1954), The Foundations of Statistics (Wiley, New York).

Sengupta, J.I., and R.E. Sfeir (1982), "Risk in supply response: An econometric application", Applied Economics 14:249–268.

Sharpe, W.F. (1963), "A simplified model for portfolio analysis", Management Science 2:277–293.

Sigurdson, D., and R. Sin (1994), "An aggregate analysis of Canadian crop insurance policy", in: D.L. Hueth, and W.H. Furtan, eds., Economics of Agricultural Crop Insurance: Theory and Evidence (Kluwer Academic Publishers, Boston) 45–72.

Sinn, H.-W. (1989), "Two-moment decision models and expected utility maximization: Comment", American Economic Review 79:601–602.

Skees, J.R., and M.R. Reed (1986), "Rate-making for farm-level crop insurance: Implications for adverse selection", American Journal of Agricultural Economics 68:653–659.

Smale, M., R.E. Just and H.D. Leathers (1994), "Land allocation in HYV adoption models: An investigation of alternative explanations", American Journal of Agricultural Economics 76:535–546.

Smith, V.H., and A. Baquet (1996), "The demand for multiple peril crop insurance: Evidence from Montana wheat farms", American Journal of Agricultural Economics 78:189–201.

Smith, V.H., and B.K. Goodwin (1996), "Crop insurance, moral hazard, and agricultural chemical use", American Journal of Agricultural Economics 78:428–438.

Standing Committee on Agriculture and Colonization (1947), Minutes of the Proceedings and Evidence, (Canadian House of Commons, Ottawa, Canada) June 10.

Tauer, L.W. (1983), "Target MOTAD", American Journal of Agricultural Economics 65:606–610.

Taylor, C.R. (1990), "Two practical procedures for estimating multivariate nonnormal probability density functions", American Journal of Agricultural Economics 72:210–218.

Teague, M.L., D.J. Bernardo and H.P. Mapp (1995), "Farm-level economic analysis incorporating stochastic environmental risk assessment", American Journal of Agricultural Economics 77:8–19.

Tobin, J. (1958), "Liquidity preference as behavior towards risk", Review of Economic Studies 25:65–86.

Tolley, H.D., and R.D. Pope (1988), "Testing for stochastic dominance", American Journal of Agricultural Economics 70:693–700.

Traill, B. (1978), "Risk variables in econometric supply models", Journal of Agricultural Economics 29:53–61.

Traxler, G., J. Falck-Zepeda, R.J.I. Ortiz-Monasterio and K. Sayre (1995), "Production risk and the evolution of varietal technology", American Journal of Agricultural Economics 77:1–7.

Turvey, C.G. (1992a), "An economic analysis of alternative farm revenue insurance policies", Canadian Journal of Agricultural Economics 40:403–426.

Turvey, C.G. (1992b), "Contingent claim pricing models implied by agricultural stabilization and insurance policies", Canadian Journal of Agricultural Economics 40:183–198.

Udry, C. (1994), "Risk and insurance in a rural credit market: An empirical investigation in Northern Nigeria", Review of Economic Studies 61:495–526.

United States Department of Agriculture (1996), Agricultural Statistics (various issues from 1990 to 1996).

United States General Accounting Office (1999), Agriculture in Transition – Farmers' Use of Risk Management Strategies, GAO/RCED-99-90, Washington, D.C.

Vercammen, J., and G.C. van Kooten (1994), "Moral hazard cycles in individual-coverage crop insurance", American Journal of Agricultural Economics 76:250–261.

von Neumann, J., and O. Morgenstern (1944), Theory of Games and Economic Behavior (Princeton University Press, Princeton).

Vukina, T., D. Li and D.M. Holthausen (1996), "Hedging with crop yield futures: a mean-variance analysis", American Journal of Agricultural Economics 78:1015–1025.

Williams, J. (1986), The Economic Functions of Futures Markets (Cambridge University Press, Cambridge, U.K.).

Williams, J.R., G.L. Carriker, G.A. Barnaby and J.K. Harper (1993), "Crop insurance and disaster assistance designs for wheat and grain sorghum", American Journal of Agricultural Economics 75:435–447.

Wright, B.D. (1993), "Public insurance of private risks: Theory and evidence from agriculture", in: M.S. Sniderman, ed., Government Risk Bearing (Kluwer Academic Publishers, Boston).

Wright, B.D., and J.D. Hewitt (1990), All Risk Crop Insurance: Lessons from Theory and Experience (Giannini Foundation, California Agricultural Experiment Station, Berkeley).

Young, D.L. (1979), "Risk preferences of agricultural producers: Their use in extension and research", American Journal of Agricultural Economics 61:1067–1070.

Chapter 3

EXPECTATIONS, INFORMATION AND DYNAMICS

MARC NERLOVE

Department of Agricultural and Resource Economics, University of Maryland, College Park, MD

DAVID A. BESSLER

Department of Agricultural Economics, Texas A & M University, College Station, TX

Contents

Handbook of Agricultural Economics, Volume 1, Edited by B. Gardner and G. Rausser

Abstract

The role of expectations in the empirical analysis of agricultural supply is examined under the assumption of separation of expectations and constraints in dynamic decision making. Extrapolative, adaptive, implicit, rational and quasi-rational, and futures-based models of expectation formation are discussed. Empirical and experimental evidence for and against various models of expectation is summarized.

JEL classification: Q11

"We decide on one particular course of action out of a number of rival courses because this one gives us, as an immediately present experience, the most enjoyment *by anticipation* of its outcome. Future situations and events cannot be experienced and therefore their degree of desirableness cannot be compared: but situations and events can be *imagined*, and the desirableness of these experiences which happen in the imagination can be compared. What gives imagined things a claim to be treated as the equivalents of future things? It is some degree of belief that the imagined things will take actual shape at the dates we assign to them."
G.L.S. Shackle, 1952.

"All production is for the purpose of ultimately satisfying a consumer. Time usually elapses, however – and sometimes much time – between the incurring of costs by the producer (with the consumer in view) and the purchase of the output by the ultimate consumer ... Meanwhile the entrepreneur ... has to form the best expectations he can as to what the consumers will be prepared to pay when he is ready to supply them (directly or indirectly) after the elapse of what may be a lengthy period; and he has no choice but to be guided by these expectations, if he is to produce at all by processes which occupy time ... the behaviour of each individual firm ... will be determined by its *short-term expectations* ... The *actually realised* results ... will only be relevant in so far as they cause a modification of subsequent expectations.
John Maynard Keynes, 1936.

1. Introduction

We consider the role of expectations and new information in agricultural economics, with reference to other work on expectation formation. The chapter is presented in four main parts. In the opening section we describe the structure of the problem of modeling dynamic optimizing behavior under uncertainty. Central to almost all treatments of the subject since the work of Keynes and Hicks in the 1930s is the separation assumption, in which dynamic decision problems are modeled by separating expectation formation from optimizing behavior. Two examples of dynamic models of agricultural supply response are used to illustrate the approach. In the second part, we present the five principal alternative approaches for modeling expectation formation: extrapolative expectations, adaptive expectations, implicit expectations, rational and quasi-rational expectations, and futures markets. In the third part, we consider the evidence on the validity of these five approaches, focusing primarily on rational expectations and the more operational variant, quasi-rational expectations. Evidence from both indirect tests, such as restrictions on parameters in an econometric model, and direct tests, such as tests of unbiasedness and orthogonality of elicited expectations obtained from survey and experiments, is presented. The chapter concludes by offering directions for future research.

2. Expectations and dynamic optimization

2.1. *The structure of the problem*

If current decisions did not constrain future possibilities, opportunities, or costs, expectations of future events would not be relevant to these decisions. It is precisely because what we do today constrains what we can do tomorrow that the future is relevant to the present.[1] (See [Nerlove (1972)].) Current events influence what we do today both directly and indirectly; directly, because present circumstances affect the desirability or profitability of actions now; indirectly, because events in the present influence our expectations of the future. These two effects may be quite different. What is the relation between dynamic optimization under uncertainty with respect to future opportunities and constraints and how economic agents form their expectations of the future and make decisions and plans?[2]

Hicks (1946) found a solution to the problem of formulating a dynamic theory of the firm under certainty by dating all variables and applying static theory to the expanded set of variables and constraints, although, in the end, he was clearly not happy with this solution [Hicks (1977)]. The Hicksian solution essentially converts the dynamic decision-making problem into a static problem. It fails to reveal the dynamic structure of decisions and constraints and to deal explicitly with uncertainty, the costs of information, or the costs of formulating plans and decisions. In principle, we know how to set the problem up as a dynamic programming problem under uncertainty, in which conditional distributions of future unknown exogenous variables are estimated by using all available information up to the present [Nerlove (1972)]. The problem of costly information is more difficult to incorporate since its value is usually not known until it is acquired, but this problem can be resolved within a Bayesian framework. (See, *inter alia*, [Horvath and Nerlove (1996), Kiefer (1988–89)].) In such a "theoretically correct" formulation, decisions and expectations are not separable; the explanation of behavior proceeds directly from assumptions about agents' priors and the dynamic constraints of their optimization problem to the decisions they take now and in the future in response to future events.[3]

---

[1] This is also true with respect to future events over which we have no control, such as events after one's death. The imminent end of the world, if known, would certainly change behavior today because constraints current behavior would impose on future options would no longer hold after the end of the world. In this sense, the future matters because of the constraints it would impose on current behavior if there were a future.

[2] The problem of what constitutes rational behavior in a dynamic context is not so simple; see [McClennen (1990)] for a careful analysis from a philosopher's point of view. Nor is it a trivial matter to make the concepts of information and uncertainty precise. There is a very extensive literature in economics on these matters which has been artfully summarized and integrated in [Hirshleifer and Riley (1992)].

[3] Notwithstanding, Mundlak (1966, 1967) has suggested that a dynamic theory should be formulated in a manner which takes explicit account of the restrictions implied by the Hicksian extension of static theory. This is an extreme form of the separation assumption, to which we would not subscribe. As Treadway (1967) has shown, the propositions of usual comparative static theory do not generally hold in a dynamic context. But

The theoretically correct formulation of the problem of dynamic decision making under uncertainty does not lend itself to empirical application, nor has it generally been adopted in studies of agricultural supply or other topics investigated by agricultural economists and, more generally, in empirical studies of expectations and plans (see [Nerlove (1983)], and the references cited therein). Instead, a separation between expectations and decisions is made and the effects of changing expectations on behavior is analyzed. "The Hicksian model of dynamic planning under certainty is the basis for a more empirically relevant framework for the analysis of plans and expectations .... The Hicksian assumption of certainty means that information about the future value of a variable is single valued and costless. We continue to regard expectations and plans as single valued but recognize that the economic agent knows that they may turn out to be wrong. As of a particular date, information about the future can be acquired only at a cost, albeit a cost which decreases for a particular future date as that date draws near. Planning and decision making are themselves costly activities. Therefore only what is necessary to plan will be planned, only decisions which cannot be postponed will be made, and only the information about the future necessary to those plans and decisions and only to the accuracy warranted by the cost of error will be gathered. Plans will not always be fulfilled, single-valued expectations will often turn out to be wrong, and both will be continually revised" [Nerlove (1983, p. 1252)]. We refer to the assumption that dynamic decision problems can be analyzed in terms of expectations and the impact of expectations on decisions as the *separation assumption*. It is clearly only an approximation, albeit an empirically and theoretically useful one.

Even when the separation assumption is adopted, there is another serious problem which models of expectation formation and dynamic behavior share with most other models on which econometric analyses are based: they typically assume a representative economic agent whose optimizing decisions are the basis for the analysis. Not only does such an assumption raise the question so ably and concisely discussed by Kirman (1992), but another branch of the literature has emphasized the role of heterogeneity

---

this does not mean that separation of expectations and optimizing behavior is impossible within the context of an appropriately formulated dynamic model [Nerlove (1972)].

The econometric modeling of dynamic decision making processes has recently enjoyed a resurgence of interest; see, for example Kapteyn, Kiefer, and Rust (1995), especially the paper by Miranda and Schnitkey (1995). It is, however, not clear to us whether such econometric "fine-tuning" is really desirable, notwithstanding Nerlove (1972) and more recently Nerlove and Fornari (1997). Carrying forward the research of more than two decades, Hildenbrand (1994), for example, shows that the specification of behavioral relationships at the individual level does not play a dominant role in determining the sort of relationship commonly estimated econometrically. He argues, in the context of cross-section expenditure studies, that certain invariant features of the distribution of household characteristics and attributes are much more important in determining the relationships of interest, and that these can be derived without any need to specify a precise model of microeconomic behavior. We believe that Hildenbrand's conclusions are valid generally and beyond the context of econometric analysis of household expenditure surveys. Many restrictions imposed by microeconomic theory, whether static or dynamic, are of very limited value in improving econometric estimation. Other aspects of the data-generating process are much more important. To attempt to fine-tune the econometrics by imposing such restrictions can lead to results which may be highly misleading.

of expectations in the determination of aggregate outcomes [Nerlove (1983), Frydman and Phelps (1983)]. Such heterogeneity is inconsistent with the representative agent assumption.[4]

2.2. *Examples of the separation of expectations and constraints in dynamic decision making*

The device of separation of expectations from plans and decisions and the utility of such separation in both theory and empirical analysis may be illustrated by two models of agricultural supply: The first of these examples is the well-known model of agricultural supply response developed by Nerlove (1956a, 1956b, 1958c) for corn, cotton, and wheat in the U.S. The second is a more elaborate model of small ruminant production and supply in Indonesia developed by Nerlove and Soedjana (1996).[5] The importance of the second example is to show that a comparative static analysis is possible in models involving both dynamic optimization and uncertainty, *even though the process of expectation formation is not specified*, as long as the separation assumption is maintained. The representative agent assumption is also common to these examples.

2.2.1. *The Nerlove supply model*[6]

Stripped to its essentials, this model for an annual crop consists of three equations:

$$A_t - A_{t-1} = \gamma(A_t^* - A_{t-1}), \tag{1}$$
$$P_t^* - P_{t-1}^* = \beta(P_{t-1} - P_{t-1}^*), \tag{2}$$
$$A_t^* = a_0 + a_1 P_t^* + a_2 Z_t + U_t, \tag{3}$$

where A_t is actual area under cultivation in t; P_t, actual price of the crop per unit in t; A_t^*, "desired" or equilibrium area to be under cultivation in t; P_t^*, "expected normal" price in t for subsequent future periods; Z_t, other observed, presumably exogenous, factors; U_t, unobserved, "latent" factors affecting area under cultivation in t; and β and γ are "coefficients of expectation and adjustment", reflecting the responses of expectations to observed prices and observed areas under cultivation to changes in equilibrium areas.

---

[4] See also the discussion of heterogeneity in the determination of aggregate outcomes in the preceding footnote.

[5] The interesting study of Miranda and Schnitkey (1995) does not employ this separation. They assume that the two relevant stochastic variables, revenue less variable cost of milk and the market price of a heifer less the slaughter value of a replace cow, follow a first-order vector autoregression (VAR), known to the dairy farmer, the parameters of which are to be estimated along with the rest of their model. However, such a model could be interpreted in terms of rational or quasi-rational expectations under the separation assumption; see below, Section 3.4.

[6] This discussion is taken from [Nerlove (1979)].

The statistical problems of estimating a model such as (1)–(3), particularly of identifying relevant observed exogenous variables, not subject to expectational lags, and problems due to serially correlated disturbances, are well known. In addition, the use of area cultivated, one input in the production process to represent planned output, the problem of choosing the relevant price or prices, and other issues of specification, such as the inclusion of expected yields, weather conditions, and price and yield variances to take account of elements of risk, have been widely discussed in the literature (see, for example, *inter alia* [Just (1974), Askari and Cummings (1976, 1977)]).

The Nerlove supply response model incorporates dynamic elements in two different ways: First, a distinction is made between a long-run equilibrium position, toward which producers are assumed to be moving, and their current position. The former is determined on the basis of a static theory of optimization, in this case the standard microeconomic theory of the firm and the assumption that the exogenous variables of the problem, in this case mainly prices, are given once and for all. Nerlove (1972, p. 225) called this the assumption of static, or stationary, expectations. The important point is that whatever these expectations are and however they are formed, the concept of a long-run equilibrium solution to the optimization problem is well defined only if it can be assumed that the values of the exogenous variables expected in the future are unchanging; it does not matter if the constant future value of each variable differs from its current value, as indeed it plausibly will. Having a well-defined notion of a long-run equilibrium position then permits us to examine the question of why producers are currently at a position different from that equilibrium. At this point the discussion usually becomes vague; one can argue in various ways (Nerlove, 1972, pp. 228–231), but perhaps the most common approach is through the introduction of adjustment costs. Rarely, however, are models explicitly introducing these costs formulated or the rationale for such costs carefully examined.[7]

The dynamic element in the basic supply response models is introduced at this point without a formal theory by the simple ad hoc assumption that in each period, if we are dealing with discrete time, a fraction of the difference between the current position and the long-run equilibrium is eliminated, i.e., Equation (1) above.

The second way in which dynamic elements are incorporated in the basic supply response model is through a description of expectation formation, e.g., the adaptive expectations generated by Equation (2), in which expected "normal" prices are revised each period in proportion to the difference between last period's observed price and the previous expectation. Above, we argued that static, or stationary, expectations are necessary to make the concept of a long-run equilibrium meaningful; the adaptive expectations model does not violate this principle, since it is not solely next period's price to which P_t^* refers but "normal" price, i.e., an average price expected to prevail in all future periods. Nerlove (1956a, 1956b, 1958c) makes the argument that farmers rationally

---

[7] The literature up to about 1970 is surveyed, and two models of investment behavior incorporating both separable and non-separable adjustment costs are discussed, in [Nerlove (1972, pp. 231–241)]; see also [Nerlove et al. (1979 and 1995, pp. 317–320)].

should respond, not to the best forecast they can make of next period's price, but rather to some average or "normal" level; the argument rests intuitively on the idea that there are costs of adjustment. However, virtually any plausible model one can construct, with costs of rapid adjustment of, say, a durable factor of production, will generally involve response to prices in many future periods, although the weights which attach to the more distant future will usually be less than to the near future. Moreover, unless the optimization problem has a specific form, it will generally be non-optimal to behave as if one were responding to a point estimate of each future value. When the optimization problem is of this specific form, however, we say that there exist certainty equivalents to the uncertain future values of the variables to which response is occurring [Theil (1957), Malinvaud (1969)]. Such certainty equivalents are the conditional expectations of the variables to which they refer; they are minimum-mean-square-error forecasts based on the information available up to the time the forecast is made and taking into account the structure of the system generating the data. Muth (1961) has termed such forecasts "rational expectations". We will discuss rational expectations models of agricultural supply at some length below.

2.2.2. A model of small ruminant production and supply

The dynamics of annual crop supply are particularly simple; their very simplicity may obscure the relation between expectations and dynamic optimizing behavior. Better examples of greater dynamic complexity may be found in the study of perennial crops, such as rubber, coffee, cocoa, palm oil or asparagus, or of livestock. The following model shows that a comparative static analysis is possible in models involving both dynamic optimization and uncertainty, *even though the process of expectation formation is not specified*, as long as the separation assumption is maintained. Nonetheless, it also illustrates the importance of expectations in determining dynamic optimizing behavior.

Small ruminant production and supply presents an ideal case to illustrate the points made above, being neither too simple nor, because of the short gestation and maturation period of the animals, as complex as cattle and many perennial crops. The following development is based on Nerlove and Soedjana (1996), hereinafter N&S, whose primary purpose is to elucidate the role which small ruminants play as a store of value in the context of traditional Indonesian society. In their paper, details of which are not elsewhere published, they make considerable use of neoclassical monetary theory, an aspect of the analysis which we neglect here. Small ruminants in general are referred to as "sheep".

N&S assume that sheep live for two periods. In the first period, they are gestating or prepubescent. In the second period, all the time that they remain in the herd, they reproduce at a rate $\alpha > 1$. At the end of the first period, which is the same as the beginning of the second period, some are sold and do not survive to reproduce. Let

S_t = the stock of sheep at the beginning of period t;

s_{t+1} = sales at the end of period t or the beginning of period $t + 1$.

Then, the stock at the beginning of period $t+1$ is

$$S_{t+1} = \alpha[S_t - s_{t+1}]. \tag{4}$$

Let

$C(S)$ = the costs of maintaining a herd of size S for one period;
p_t = the price per sheep sold expected in period t;
p_0 = the actual price in the current period, $t = 0$, at the end of which s_0 sheep are sold.

Assume that these expectations are held with certainty, or alternatively, that the structure of the problem is such as to admit of certainty equivalents. Let S_0 be the initial herd size. The costs of maintaining this herd during the first period are sunk costs and must be borne out of revenues generated previously. Current gross revenue at the end of the initial period is $p_0 s_1$, but the costs of maintaining the herd in the following period $C(S_1)$ must be paid from these revenues, so that net revenue in the current period is $R_0 = p_0 s_1 - C(S_1)$. In general,

$$R_t = p_t s_{t+1} - C(S_{t+1}), \quad t = 0, 1, \ldots. \tag{5}$$

Along the lines of neoclassical monetary theory, N&S assume that the utility function of the representative farmer is additively separable over time and a homothetically weakly separable function of the stock of sheep and current revenue (which can be taken as a Hicks-composite commodity if the prices of real commodities consumed by the farmer are assumed not to change). That is, we assume that the farmer's consumption decisions are determined by maximizing a "branch" utility function in real commodities given the revenues realized from the sale of sheep at the beginning of each period. Thus, the utility of the farmer in each period is given by

$$u_t = U\big[\varphi(R_t), S_{t+1}\big] = U\left[\varphi\left(p_t\left\{S_t - \frac{S_{t+1}}{\alpha}\right\} - C(S_{t+1})\right), S_{t+1}\right]. \tag{6}$$

Given the additive temporal separability of total utility, as is well known [Barro (1974), Barro and Becker (1989), Nerlove and Raut (1997)], total utility can be expressed as

$$TU = \sum_{t=0}^{\infty} \beta^t u_t, \quad \text{where } 0 < \beta < 1. \tag{7}$$

In a perfectly functioning capital market, β would equal the rate of interest at which the farmer could borrow, but in the absence of such a market, as we assume here, β

expresses the farmer's rate of time discount. Assume that the farmer has an infinite horizon as if he expected to live forever.

Assume that φ is chosen so that U_1 (where the subscript denotes a derivative with respect to the argument in question) is normalized to 1, i.e., $\varphi' U_1 = 1$, and that $U_2 \geqslant 0$ and $U_{22} \leqslant 0$ and that the farmer maximizes TU by choosing the sequence of herd sizes $S_1, S_2, \ldots$, given the initial herd size S_0.

Maximizing TU with respect to the sequence S_t, $t = 1, \ldots, \infty$, given S_0, is now in the form solved by Stokey and Lucas (1989, pp. 68–84), who show that it is equivalent to maximizing

$$U\bigl[\varphi(R_0), S_1\bigr] + \beta v(S_1, p_1, p_2, \ldots) \tag{8}$$

with respect to S_1, where

$$R_0 = p_0\left[S_0 - \frac{S_1}{\alpha}\right] - C(S_1)$$

and where v is the maximized value of TU in the next period given the value S_1 of the initial stock in that period, chosen in the initial period, and price expectations in all future periods.

The first-order condition for this problem, recalling that U_1 is normalized to 1, is

$$-\frac{p_0}{\alpha} - C' + U_2 + \beta v' = 0. \tag{9}$$

Define

$$\mu(S_1) = \beta v' + U_2,$$

which is the value of a sheep saved in the current period in terms of future breeding capacity, and therefore addition to future revenues and utility plus the utility of having her in stock next period as a store of value. Rearranging terms, we have

$$\alpha\mu(S_1) - p_0 = \alpha C'. \tag{10}$$

Equation (10) is quite intuitive. It says that at an optimum of the producer, the marginal cost of maintaining an animal in the herd next period must be equal to the value of a sheep saved minus the opportunity cost of not selling her. The coefficient $\alpha > 1$ multiplies both μ and C' to account for the fact that a sheep saved today will become α sheep tomorrow.

The left-hand side of (10) is proportional to marginal cost. Average cost may decline initially for very small herd sizes because of certain fixed costs such as barns, but must rise after a certain size of herd (rather small in semi-subsistence Indonesian agriculture),

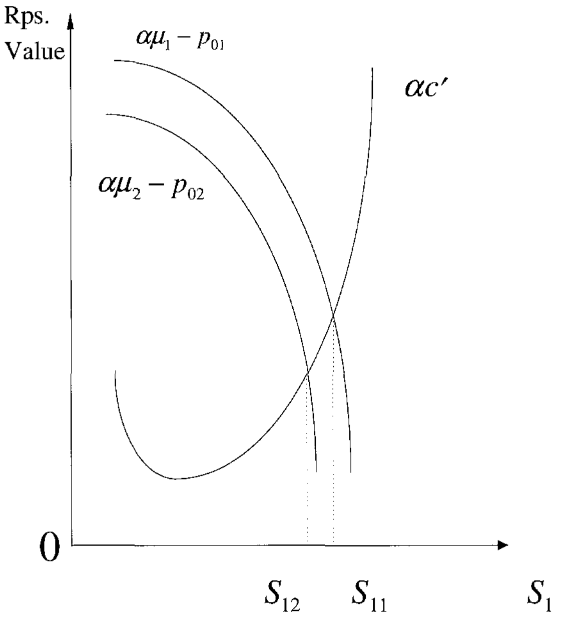

Figure 1. Relationship of the optimal stock of livestock to prices and costs.

and, at some point, begin to rise steeply because of the labor and other resource constraints which the farmer faces. The behavior of the right-hand side is more problematic: $U_{22} \leqslant 0$, so that the second term of μ must be declining with S_1, but if expected future prices of sheep are rising fast enough, v' may not decline with S_1, even if marginal future costs of increasing herd size are rising rapidly. N&S assume that this is not the case.[8] This provides the first illustration of the power of the *separation assumption.*

The solution is graphically depicted in Figure 1.

It is apparent both from Equation (10) and the figure that an increase in the current price of sheep, expectations of future prices unchanged, will lead, *ceteris paribus,* to a decline in the herd size next period, and thus to an increase in sales. But if an increase in the current price is accompanied by an increase in expectations of future prices, causing a rise in μ sufficient to offset the increase in p_0, the current supply of sheep to the market may actually decline. (Of course, this is true irrespective of whether the stock of sheep enters the utility function directly.)

---

[8] This corresponds to the well-known transversality condition, which is generally assumed in dynamic optimization problems (see [Stokey and Lucas (1989, p. 98)]. If this condition does not hold and if expectations of rapidly rising future prices are not offset by rapidly rising costs of future herd size, μ may rise with S_1, in which case an equilibrium of the producer would still exist in the rising part of the $\alpha C'$ curve, but not in the region of increasing returns to herd size in which $\alpha C'$ is falling.

The rationale for such perverse supply response to price, in general, was cogently argued by Jarvis (1986) and Rosen (1986); Jarvis' empirical verification for Argentine beef cattle relied on ad hoc assumptions about expectation formation; Nerlove and Fornari (1997) provide evidence for the U.S. beef cattle industry of a positive response to current price holding expectations of future prices constant, but a negative response to increases in expected future prices holding current prices constant, using a rational expectations model of price expectation formation. The N&S result, as is the case with Rosen's result, is free of any significant restriction on the nature of expectation formation. If the $\alpha\mu - p_0$ curve cuts the $\alpha C'$ curve in the segment of the latter that is rising extremely rapidly, we would expect to observe hardly any supply response either to current price or to expected future prices.[9]

As the foregoing model illustrates, it is unnecessary to make any specific assumption about the formation of expectations to derive useful results concerning the role of expectations in the determination of dynamic optimizing behavior. Nonetheless, in order to study such behavior econometrically, it is necessary to specify a model of the way in which expectations are formed. To this we now turn.

3. Alternative models of expectation formation[10]

In this section we examine the leading models of expectation formation used in empirical time series analysis of agricultural supply and in other areas of applied economics. The justification for considering models of expectation formation in the context of a model of economic (optimizing) behavior rests in large part on the separation assumption discussed above, to which must be added the assumptions that (1) group behavior can be adequately explained by treating it as the behavior of a single representative and hypothetical decision maker (the representative agent assumption); and (2) the representative decision maker behaves as if responding to single-valued certainty equivalents (the certainty equivalent assumption). The expectations, to which economic agents are assumed to respond, are both subjective and aggregative. They are not necessarily, or even generally, directly observable. The problem in empirical analysis discussed in this

---

[9] Many other results follow from this model. For example, N&S deduce the effects of improved access to financial institutions: Changes in the effectiveness with which local financial institutions serve the semi-subsistence Indonesian farmer will, in the first instance, primarily affect U_2, the direct marginal utility of holding an additional sheep in the herd. Less directly, changes in U_2 will also affect μ, the value of future maximized net revenues from sheep raising. If U_2 is set to zero v', the future value of having a sheep in the herd will fall even if expectations of future prices and costs are unchanged; the term U_2 in μ will be eliminated entirely. Consequently, the entire curve $\alpha\mu - p_0$ will shift downwards relative to the $\alpha C'$ curve. Unless, before the assumed change the curves crossed in the very nearly vertical portion of the $\alpha C'$ curve and still cross there, the optimal herd size will be reduced by better access to financial institutions. This situation is depicted in Figure 1 by the vertical portion of the $\alpha C'$ curve.

[10] This section is adapted ftom [Nerlove (1958c, Chapter 2), Nerlove (1961), and Nerlove and Fornari (1997)].

section is to construct a hypothesis which relates these expectations to observable variables. In this section we will consider five types of models or approaches to the study of expectation formation within the context of a simple model of supply response: (a) Extrapolative; (b) Adaptive; (c) Implicit; (d) Rational and Quasi-Rational; and (e) Futures Market Based. In the next section, we consider research when data related to expectations are more directly observable, for example from surveys or experiments. One can argue that futures prices, when available, are intermediate between direct and indirect observation of expectations.[11]

3.1. Extrapolative

The classical approach in agricultural supply analysis (at least prior to [Nerlove (1956a, 1956b)] was to suppose that expectational variables could be directly identified with some past actual value of the variable to which the expectation refers. For example, the supply of an agricultural commodity at a future time depends on its price expected at that time. It might be assumed that this expectation is the current value of price, so that supply is simply related to lagged price. An extension of this approach, due to [Goodwin (1947)], is to suppose that expected price in period t is actual price in $t-1$ plus (or minus) a fraction of the change in price from period $t-2$ to $t-1$:

$$p_t^* = p_{t-1} + \alpha(p_{t-1} - p_{t-2}), \tag{11}$$

where p_t^* is the price expected in period t. Muth (1961) calls the expectations generated by (11) "extrapolative".

3.2. Adaptive

The origins of adaptive expectations are somewhat obscure. Nerlove (1956a) attributes the idea to Phillip Cagan in his 1956 Ph.D. dissertation on hyper-inflations; but later (1956b, 1958c) says that the idea is essentially Hicks'. Milton Friedman claims he got the idea from Bill Phillips of Phillips Curve fame. After an exhaustive look at empirical studies of expectations that existed before 1956, here's what Nerlove (1958c, pp. 50–53) writes:

> ... the main results of the ... studies examined indicate that there is widespread underestimation of actual changes and that forecasters could generally do a better job at predicting the levels of actual outcomes if they used some simple mechanical device such as a projection of the current value of the variable to be predicted. The question immediately arises as to whether entrepreneurs are really trying to forecast a particular value of an economic variable, or whether, as suggested above

---

[11] However, for storable commodities, cash prices also reflect the same information, so that a futures price is no more and no less an expectation than the current price.

> they try to forecast the "normal" *level* of future values of the variable. As indicated above, entrepreneurs' response to a change which they consider only temporary may be very limited. True, entrepreneurs could make greater profits the more accurate their knowledge of the future; but these profits might not be much greater than those they might make if they altered their plans only in response to changes in the expected level of future values of the economic variable under consideration. ... Hence, any model of expectation formation should take account of the fact that these expectations probably do not refer to the immediate and temporary future.
>
> We may take ... a concept of the normal as a starting point in our development of a model of expectation formation. The discussion at this point may most easily be couched in terms of prices and price expectations. If more specific information is not available, it seems reasonable to assume that the "normal" price expected for some future date depends in some way on what prices have been in the past. Expectations of "normal" price are, of course, shaped by a multitude of influences, so that a representation of expected price as a function of past prices may merely be a convenient way to summarize the effects of these many and diverse influences. ...
>
> How should we use past prices to represent expected "normal" price? Each past price represents only a very short-run market phenomenon, an equilibrium of those forces present in the market at the time. ... We observe, however, that entrepreneurs' expectations, if taken as forecasts of the immediate future, predict the levels of actual outcomes in the immediate future less well than would a simple naive model forecast of no change. This fact suggests that entrepreneurs do not regard any particular past price or actual outcome as overwhelmingly indicative of long-run normal conditions. If they did their expectations might do better when considered as forecasts.

Continuing, Nerlove relates the idea to Hicks' definition of the elasticity of expectations:

> Hicks may very well have had this notion in mind when he defined "the elasticity of a particular person's expectations of the price of a commodity x as the ratio of the proportional rise in expected future prices of x to the proportional rise in its current price" (1946, p. 205). Hicks, it will be remembered, distinguished two limiting cases: an elasticity of zero, implying no effect of a change in current price upon expected future prices; and an elasticity of one implying that if prices were previously expected to remain constant, i.e., were at their long-run equilibrium level, they will now be expected to remain constant at the level of current price. By allowing for a range of elasticities between the two extremes, Hicks implicitly recognized that a particular past price or outcome may have something, but not everything, to do with people's notion of the "normal".

And then the key concept of *expected normal price*:

Past values of prices, then, affect people's notions of the "normal" level of prices; individual past prices do not exert their influence equally, however: more recent prices are a partial result of forces expected to continue to operate in the future; the more recent the price, the more it is likely to express the operation of forces relevant to "normal" price. An obvious extension of this point of view would be the representation of people's notion of "normal" price by a weighted moving average of past prices in which the weights decline as one goes back in time. Using Hicks' concept of an elasticity of expectation we can go beyond this formulation; indeed, we can derive it.

Hicks' definition of the elasticity of expectation implies that prices have actually been "normal" up until the time when some change occurred. But, of course, we know that conditions are seldom if ever "normal" in the real world; and "normality" itself is a subjective matter. Let P_t^* stand for people's expectation at time t of long-run "normal" price, and let P_t stand for actual price. Hicks' notion may then be expressed by saying that P_t^* is last period's expected "normal" price plus some factor depending on the elasticity of expectation and last year's actual price. We will go further than this and say that the adjustment factor is proportional to the difference between actual and expected "normal" price. Intuitively this seems quite reasonable. Mathematically we may write

$$P_t^* = P_{t-1}^* + \beta[P_{t-1} - P_{t-1}^*], \quad 0 < \beta \leqslant 1, \tag{12}$$

where β is a constant. If β were equal to zero, it is clear that actual prices would have no effect whatsoever on expected "normal" price. On the other hand, if β were equal to one, expected "normal" price would be equal to last year's actual price. The case of $\beta = 1$ thus corresponds to the type of forecasts generated by the naive model discussed above. In what follows we call β the coefficient of expectation [to distinguish it from an *elasticity*]. The hypothesis proposed may be stated in words: each period people revise their notion of "normal" price in proportion to the difference between the then current price and their previous idea of "normal" price.

At this point, Nerlove (1958c, p. 54) shows that the adaptive expectation hypothesis implies a representation of "expected normal price" as a weighted average of past prices with weights which decline geometrically as one goes back in time:

$$P_t^* = H(1-\beta)^t + \sum_{\lambda=0}^{t} \beta(1-\beta)^{t-\lambda} P_{\lambda-1}, \tag{13}$$

where H is a constant the value of which depends upon the initial conditions. Let us assume that an equilibrium situation existed at and prior to time $t = 0$. Let us further assume, without essential loss of generality, that all prices are expressed as deviations from the equilibrium price existing at time $t = 0$. Then H may be

> taken to be equal to zero and (13) becomes
>
> $$P_t^* = \sum_{\lambda=0}^{t} \beta(1-\beta)^{t-\lambda} P_{\lambda-1}.$$
>
> We have thus expressed people's notion of the normal price as a weighted average of past prices. The weights of past prices are functions of β and they decline as one goes back in time, since β is between zero and one.

Because expected normal prices at $t = 0$ and before are not observable, the geometrically weighted average can represent only an approximation valid for $t = 0$ in the distant past. And, in practice, because annual agricultural prices can be obtained only for short periods, Nerlove (1956a; 1956b, Chapter 8) proposed to approximate these expectations in terms of farmers' past observed behavior: In effect, if last year's supply depends on last year's expectation of normal price, then last year's supply can be used as a "stand-in" for the unobserved variable.[12]

3.3. *Implicit expectations*

In a remarkable dissertation [Mills (1955)], which was later largely incorporated in [Mills (1962)], Mills develops the idea of what he calls *implicit expectations*. Here is what he later wrote (1962, pp. 37–39):

> The approach ... starts with a recognition that an expectation, in addition to being a function of observable variables, is also the decision maker's estimate or prediction of a variable. As with any other estimate, an expectation has certain statistical properties which, in principle, are discoverable. This is perfectly obvious. What appears to be an innovation is the argument that, on certain assumptions about the statistical properties of the estimate, the economist can estimate both the behavior relation and the expectation itself in an indirect or implicit way. ... the expectational error [is defined] by
>
> $$x = x^e + u.$$
>
> In words, u is the decision maker's error in predicting x. Substituting [x^e in the behavioral relation to be inferred, $Y(x^e)$]:
>
> $$y = Y(x^e) = Y(x-u)$$

---

[12] Eckstein (1985) presents a model of agricultural supply which, under the assumptions made, is observationally equivalent to the Nerlove supply model with adaptive expectations. Further details are given below in connection with our discussion of rational and quasi rational expectations, Section 3.4.

a relation of the error in the variable type, since the observed variable x differs from the true variable x^e by an error of observation u. Virtually all that is known about statistical properties of estimates of such relations is concerned with the case in which the Y function is linear. ... further discussion will be restricted to the case in which the decision rule is linear in x^e, that is,

$$y = Y(x^e) = \alpha + \gamma x^e.$$

Then, ... we obtain

$$y = \alpha a + \gamma x - \gamma u = \alpha + \gamma x - \varepsilon,$$

where $\varepsilon = -\gamma u$. Now [this equation] is a standard statistical specification of a linear structural equation connecting the observable variables y and x, and on certain assumptions concerning the statistical properties of standard least squares techniques will yield good estimates of α and γ.[13] Assume for the moment that these properties ... are present and that we have least squares estimates a and c of a and g from a sample of observations of x and y. We then obtain an estimate $\hat{y}$ of y from the regression equation,

$$\hat{y} = a + cx,$$

[where a and c are supposed to be the OLS estimates of α and γ]. This estimate is subject to a regression error e defined by

$$e = y - \hat{y}$$

the difference between the observed and predicted values of y. Now the regression error is an estimate of the true residual e:

$$e = \text{est}\, e + \text{est}(-\gamma u).$$

Therefore,

$$e/c = \text{est}(-u) = -\hat{u}.$$

From this we obtain an estimate $\hat{x}^e$ of x^e

$$\hat{x}^e = x - e/c = x - \hat{u} = (y - a)/c.$$

---

[13] Note by MN and DB: The problem is that the standard assumptions cannot hold because u is correlated with the *observed value of x by definition*. This problem is resolved by "rational" expectations, discussed in the next subsection.

> We refer to $\hat{x}^e$ as the implicit expectation. The basic idea behind this calculation is very simple. The implicit expectations approach makes possible an estimate of the behavior equation without first obtaining an estimate of the independent variable x^e. Once the behavior equation has been estimated, however, the inverse function provides an estimate of the expectation as a function of the observed decision. We refer to $\hat{x}^e$ as the implicit expectation since it is an estimate of the value such that, if this were the true expectation, it would lead to the behavior actually observed.

The bottom line is that implicit expectations amounts to substitution of the observed future value of a variable, the expectations to which economic agents are assumed to react, by its actual value. The approach runs aground because the expectational errors, which now comprise part of the disturbance in the relation to be estimated, are, by their very definition, correlated with those same observed variables. This problem is resolved by the rational expectations hypothesis, in which the expectational variable is assumed to be the conditional expectation of the future value of the variable conditional on all the information available up to the point at which the expectation is formed.

3.4. *Rational expectations and quasi-rational expectations*

Since the introduction by Muth (1961), the rational expectations hypothesis (REH) has occupied a central position in discussions of what ought to be done that is however incommensurate with its limited application in econometric practice. It is difficult to disagree with the basic tenet of REH that economic agents make purposeful and efficient use of information, just as they do of other scarce resources, in optimizing their decisions.[14] Yet in actual implementation, the general form of the REH is replaced by the implication that anticipated future values of relevant variables are equal to their expectations conditional on all past data and the model itself, which describes the behavior based on those expectations. (Hereinafter, we refer to this form of the REH exclusively.) There are many reasons why this form of the REH may fail: (1) The objective functions being maximized by agents are not quadratic subject to linear stochastic constraints. (2) Agents are learning about both the processes generating exogenous variables and/or about the model characterizing their behavior in aggregate (see [Horvath and Nerlove (1996)]). (3) The econometrician may fail to specify the behavioral model, especially

---

[14] A devastating indictment of self-fulfilling expectations, the theoretical form of rational expectations, from a strictly theoretical point of view is given in a recent paper by Grandmont (1998). Essentially Grandmont argues that the informational requirements of RE lead to the defense that they are the convergent outcome of a fast learning process. In turn, such an argument requires us to consider the question of stability. His analysis shows that when expectations matter a lot and agents are uncertain about the local dynamics of the system of which they are a part, learning generates locally unstable equilibria. That is, RE are incompatible with stability of equilibrium! Since econometric analysis generally presupposes that we observe a sequence of attained equilibria, such instability implies that such observations do not exist. Needless to say, we ignore this point in the remainder of this chapter, but it is something to ponder.

its dynamics, and/or the information available to agents correctly. (For an extensive and general discussion of the limits of RE, see [Pesaran (1987)].)

Quasi-rational expectations (QRE) are a form of rational expectations obtained by neglecting some of the restrictions implied by the REH.[15] Because of their close relation, we deal with both RE and QRE in the present section. Treating them together, rather than RE first and then QRE, makes for a briefer exposition.

To illustrate the ideas involved, consider a model with a single structural equation relating one endogenous variable, y_t, to one exogenous variable, z^*_{t+1}, with a random white noise disturbance, w_t:

$$y_t = a + bz^*_{t+1} + w_t, \tag{14}$$

w_t i.i.d. $\mathrm{WN}(0, \sigma^2_w)$. Suppose that z_t follows a simple ARMA model, say AR(1), for simplicity:

$$z_t = \alpha z_{t-1} + v_t, \tag{15}$$

where the v_t are i.i.d. $\mathrm{WN}(0, \sigma^2_v)$ independently of w_t. Then if observations on past values of y_t and z_t are the only information available at t, the RE are

$$z^*_{t+1} = \mathrm{E}(z_t \mid \Omega_t) = \alpha z_t, \tag{16}$$

where Ω_t is the relevant information set, consisting of past observations on y_t and z_t and other variables, which are, however, according to this model, irrelevant. Thus one should estimate a, b, α, σ^2_w, and σ^2_v jointly:

$$\begin{aligned} y_t &= a + b\alpha z_t + w_t, \\ z_t &= \alpha z_{t-1} + v_t, \end{aligned} \tag{17}$$

subject to the constraint $b\alpha/\alpha = b$ and $\mathrm{cov}(w_t, v_t) = 0$. The resulting estimate $\hat{\alpha}$ provides the basis for calculating the RE z^*_{t+1} from $\hat{\alpha} z_t$. The QRE are obtained by estimating the second equation of (17) and then in the second stage substituting the calculated values of z_{t+1} as $\hat{z}_{t+1}$ from this estimated equation. Since w_t and v_t are assumed to be independent there is no failure of consistency. Moreover, the QRE are not less efficient, because (17) is a recursive system. In a general QRE model we would not restrict z_t to be AR(1), and this would lead only to a loss of efficiency if the model were really correctly specified, not to inconsistent estimates.

---

[15] Nerlove (1967) contains essentially the idea behind quasi-rational expectations, which are further developed in [Carvalho (1972)].

Next, let

q_t = quantity demanded = quantity supplied;
p_t = market price;
p_t^* = price expected to prevail in t on the basis of information in $t-1$ when production decisions are made;
z_{1t} = exogenous variable, e.g., income;
z_{2t} = exogenous variable, e.g., weather;
$u_{1t}, u_{2t}, v_{1t}, v_{2t}$ = latent disturbances, white noise, possibly contemporaneously correlated with each other;
w_t = latent disturbances, not necessarily white noise, but current value of which is not correlated with any variable in Ω_{t-1}.

Assume the following model:

Demand: $$q_t = \beta_1 p_t + \gamma_1 z_{1t} + u_{1t}; \tag{18}$$
Supply: $$q_t = \beta_2 p_t^* + \gamma_2 z_{2t} + u_{2t}; \tag{19}$$
Expectations: $$p_t^* = \mathrm{E}(p_t \mid \Omega_{t-1}) = p_t + w_t; \tag{20}$$
Exogenous variables: $$z_{1t} = \alpha_1 z_{1t-1} + v_{1t}, \quad z_{2t} = \alpha_2 z_{2t-1} + v_{2t}. \tag{21}$$

To obtain the fully rational expectations (FRE) estimates, equate supply and demand, replace p_t^* by $\mathrm{E}(p_t \mid \Omega_{t-1})$, and solve for

$$p_t^* = \frac{\gamma_1}{\beta_2 - \beta_1}\mathrm{E}(z_{1t} \mid \Omega_{t-1}) - \frac{\gamma_2}{\beta_2 - \beta_1}\mathrm{E}(z_{2t} \mid \Omega_{t-1}). \tag{20$'$}$$

Substitute in (19) and replace $\mathrm{E}(z_{it} \mid \Omega_{t-1})$ by $\alpha_i z_{it-1}$:

$$q_t = \frac{\beta_2 \gamma_1 \alpha_1}{\beta_2 - \beta_1} z_{1t-1} - \frac{\beta_2 \gamma_2 \alpha_2}{\beta_2 - \beta_1} z_{2t-1} + \gamma_2 z_{2t} + u_{2t}. \tag{22}$$

Estimate (18), (21), and (22) by FIML, taking into account all the cross-equation restrictions resulting from the fact that the coefficients in these equations are combinations of a smaller number of underlying parameters. The FRE of p_t, given information up to $t-1$, may be calculated from

$$p_t^* = \frac{\hat{\gamma}_1 \hat{\alpha}_1}{\hat{\beta}_2 - \hat{\beta}_1} z_{1t-1} - \frac{\hat{\gamma}_2 \hat{\alpha}_2}{\hat{\beta}_2 - \hat{\beta}_1} z_{2t-1}, \tag{23}$$

where the “hatted” values are the FIML estimates.

To obtain the strict QRE estimates of (18) and (19) we would simply replace p_t^* by $\mathrm{E}(p_t \mid p_{t-1}, p_{t-2}, \ldots)$, as calculated from the best-fitting ARIMA model. If this seems excessively simple, various intermediate possibilities are open as we shall see. How can one justify QRE in this case? The system (18)–(21) determines current values of q_t, p_t, p_t^*, z_{1t}, and z_{2t} as linear combinations of their own past values and of u_{1t}, u_{2t}, w_t, v_{1t}, v_{2t}, and their past values. For example, the result for p_t, where L is the lag operator, can be written

$$\begin{aligned}(1-\alpha_1 L)(1-\alpha_2 L)p_t& \\ &= \frac{(1-\alpha_2 L)\gamma_1 v_{1t}}{\beta_2-\beta_1} - \frac{(1-\alpha_1 L)\gamma_2 v_{2t}}{\beta_2-\beta_1} + (1-\alpha_1 L)(1-\alpha_2 L)\frac{u_{1t}}{\beta_2-\beta_1} \\ &\quad - (1-\alpha_1 L)(1-\alpha_2 L)\frac{u_{2t}}{\beta_2-\beta_1} - (1-\alpha_1 L)(1-\alpha_2 L)\frac{\beta_2 w_t}{\beta_2-\beta_1}. \end{aligned} \tag{24}$$

There is a similar equation for each of the other variables. If the latent variable w_t were uncorrelated with u_{1t}, u_{2t}, v_{1t}, and v_{2t}, Equation (24) and each of the corresponding equations for the other variables, including the unobserved variable p_t^*, would be a classical unobserved-components (UC) model [Nerlove (1967), Nerlove et al. (1979)], which has a canonical form that is an ARMA or ARIMA model. The UC formulation places additional within-equation restrictions on the coefficients which appear in each. The REH assures us that w_t is uncorrelated with any past values of u_1, u_2, v_1, and v_2, but in general it is not so with respect to contemporaneous values. This means that these UC representations contain additional parameters reflecting these correlations. While these additional parameters generally result in a failure of identification for the usual univariate UC model, they do not do so in this multivariate context because of the strong cross-equation restrictions implied by the REH.

Writing the canonical forms of (24) and the equations corresponding to it for q_t, z_{1t}, and z_{2t}, we arrive at the VARIMA model, which Sargent (1981) has suggested might be an appropriate basis for estimation, suitably restricted, for the FRE model (18)–(21). If one really did want to obtain the FIML estimates of the FRE model, however, it would be better to work within the framework of the structural equations themselves. We would estimate (18), (19), and

$$p_t = p_t^* - w_t = \frac{\gamma_1\alpha_1}{\beta_2-\beta_1} z_{1t-1} + \frac{\gamma_2\alpha_2}{\beta_2-\beta_1} z_{2t-1} - w_t, \tag{25}$$

subject to all cross-equation restrictions, assuming v_{1t} and v_{2t} to be contemporaneously uncorrelated with u_{2t} and w_t, but allowing the latter pair to be correlated.

An alternative approach to the application of rational expectations models of agricultural supply is developed in Eckstein (1985). Building on earlier work of Muth (1960), in which conditions for the optimality of adaptive expectations were derived (generalized in [Nerlove (1967)], Sargent and Wallace (1973) and Sargent (1976a, 1976b), who present models in which adaptive and rational expectations models are observationally equivalent, Eckstein presents a dynamic rational expectations agricultural supply model which leads to an acreage response equation identical to the one formulated in [Nerlove (1956a, 1956b, 1958c)]. His model "... explicitly specifies the market conditions and costs of production for a given crop. The dynamic supply equation is derived from the farmer optimization problem and the equilibrium movements of the commodity price, production, and land allocation. ... It is shown that this simple rational expectations equilibrium model, which considers dynamic constraints on land allocations through the cost function, can justify the Nerlovian supply equation.... Further, the two models have the same reduced-form equations such that they are observationally equivalent" [Eckstein (1985, p. 204)]. Of course, as might be expected, the assumptions and model specification required to arrive at this conclusion are stringent and specific: (1) production is proportional to acreage with an additive economy-wide shock; (2) cost of production per acre is a linear function of initial and harvest-time costs, and current and lagged acreage; (3) aggregate demand for the crop is a function of its price (at harvest) and income, which is assumed to evolve over time in accordance to a stationary second-order autoregressive process; and (4) the market for the crop clears. Changing the assumptions to yield a more realistic supply model would result in one not observationally equivalent – which might not be a bad thing.

3.5. *Futures price based models of expectation formation*

Rational expectations models are based on the idea that all information up to the moment at which the expectation is formed is used in the process. In practice, only observations on the past values of variables, either exogenous or endogenous entering the model, are used. A likely candidate for other information, however, in models of agricultural supply, is provided by the futures price, if one exists, for the commodities in question. The problem is that, in the case of storable commodities, it is arguable that futures prices contain no information about the aggregate of market expectations other than the current spot price. Writing in 1947, Johnson put the matter as follows:

> In commodities in which stocks are held in important volume ... the cash price is a futures price to the same extent as the price in the futures market. Because of the existence of stocks (except for a situation noted below), the present price is a consequence of a combination of forces representing the present value of the product and anticipations relative to prospective values. If anticipations are that the price of the product will be higher six months hence, this will be reflected in both cash and futures prices, since the commodity can be stored and held forward. ... In one case it might be assumed that the futures market represents a better estimate of the future prices than the cash market. This case occurs when there are no stocks, other than working stocks to be carried from one production period to the next. In such a case the cash price could go above the futures price for the future closing after the new harvest. Even here the difference is less marked than it

> might be supposed because of the length of the production process in utilizing the product and the necessity of holding stocks for this purpose.... *For the reasons given above, the variation in the cash prices of the storable commodities from year to year, in most cases, represents the whole bundle of anticipations that goes to make the market*... ([Johnson (1947, p. 83)], italics added)

This is essentially the position for storable commodities articulated earlier by Working (1942):

> For the most part, relations between futures prices, or between spot and futures prices, indicate merely the market appraisal of price changes that are likely to occur in consequence of activated marginal net cost of carrying the commodity, these marginal net costs being potentially either positive or negative.

But this was not the universally accepted view prior to Working. Working quotes a 1924 report from the Federal Trade Commission to make the point:

> ... there is no definite commercial connection between the two prices [spot and future] tending to hold them together, but instead merely a comparison of the present with a future of which the surrounding and determining conditions are not so related to the present conditions that merchants in general have objective data on the basis of which to calculate a connection between them. The future price set becomes a matter of prediction in a sense involving guesswork instead of commercial calculation of probabilities. [Working (1942, p. 40)]

Tomek and Gray (1970) reiterate Working's position:

> The element of expectations is imparted to the whole temporal constellation of price quotations, and the futures prices reflect essentially no prophecy that is not reflected in the cash price and is in that sense already fulfilled. [Tomek and Gray (1970, p. 373)]

Thus there are strong theoretical reasons to suppose that, empirically, futures prices should offer little improvement over the use of current cash prices in the case of storable commodities. Such markets, moreover, frequently fail to exist, particularly in the case of nonstorable commodities. In such cases, even when they do exist they generally depend on a variety of factors extraneous to producers' behavior. (See [Williams (1986)], especially Chapter 5.)

Notwithstanding these arguments, there is a theoretical literature suggesting that futures prices should be an essential driving variable in understanding agricultural supply. Holthausen (1979) and Feder, Just, and Schmitz (1980) show that a producer's utility-maximizing planting decision will equate marginal cost with the futures price for harvest time delivery, even if his own subjective price expectation differs from the futures price.

The empirical evidence regarding the efficacy of including futures prices in empirical analyses of agricultural supply is summarized in the next section.

4. Empirical studies of expectation formation

In this part of the chapter we deal with the evidence for or against various theories of expectation formation. Our emphasis, almost exclusive in the section on indirect evidence, is on the rational expectations hypothesis and its more operational variant, quasi-rational expectations. There are two reasons for such emphasis: First, since the publication of Muth's paper in 1961, the REH has virtually "swept the field". Except for a few experimental studies, most empirical investigations attempt to test or to exploit the REH. Second, as one colleague put it, "what's the alternative?" There is no generally theoretically acceptable hypothesis other than the REH on which one can base the expectational part of an aggregative behavioral model.

4.1. Direct versus indirect tests

Indirect tests of expectations, through restrictions on parameters from an econometric model, will necessarily be joint tests of both the expectation process used by agents and the underlying economic theory. Rejection of rational expectations, for example, when tested within the confines of a model, is a joint test of the underlying behavioral theory and the agent's use of that theory in forming his expectations on future endogenous variables. Pesaran (1987) summarizes this point:

> In the absence of direct observations on expectations, empirical analysis of the expectations formation process can be carried out only indirectly, and conditional on the behaviourial model which embodies the expectational variables. This means that conclusions concerning the expectations formation process will not be invariant to the choice of the underlying behavioral model ... Only when direct observations on expectations are available is it possible to satisfactorily compare and contrast alternative models of expectations formation. [Pesaran (1987, p. 207)]

Direct study of the expectations of individual agents elicits the response from critics that such study is assumption testing and not consistent with the positive economics precepts of Friedman (1953). One might argue, as well, that direct study of expectations is not consistent with Muth's original purposes:

> The only real test [of the REH], however, is whether theories involving rationality explain observed phenomena any better than alternative theories. [Muth (1961, p. 330)]

While this last statement clearly puts Muth (1961) in Friedman's instrumentalist camp, his more recent work suggests that he has broken camp and moved on to direct testing; see [Muth (1985)].

One might counter such objections to direct testing in several ways:

The difference between an assumption and a theorem is arbitrary. In mathematics it is largely a matter of esthetics which is which. In economics, the distinction is based on other considerations, but still basically a matter of esthetics: The central paradigm consists of a number of assumptions on maximizing behavior and equilibrium. The theorems are the result of adding other assumptions to this core. But evidence is evidence where'er we find it. The hierarchical structure of assumptions in economics and our reluctance to modify the core leads to the wrong-headed idea that one can't test assumptions at all.

A better argument is as follows: Indirect tests are clearly joint tests, so if we reject a particular hypothesis it is not clear whether we are rejecting the underlying behavioral theory or the expectational hypothesis in question. Agents may well form expectations rationally or some other specified way, but in an explicit test may misrepresent their behavior in response to those expectations. Direct observations on agent's expectations may allow us to break the joint hypothesis into two parts, one dealing with the formation of expectations, the other with consequent behavior.

Below, we summarize direct tests of expectational hypotheses. Here we continue with a brief summary of the indirect evidence, primarily related to testing models of rational or quasi-rational expectations.

4.2. *Indirect tests*[16]

In this section, we will focus on indirect testing of rational expectations (RE) and quasi-rational expectations (QRE) models of expectation formation.

There are several complementary approaches for testing RE and QRE models which have been widely discussed in the literature. These may be grouped into four categories: (a) Minimalist or general tests for whether elements of dynamic structure originate in the process of expectation formation. Such tests depend crucially on the correctness of the nonexpectational dynamics. (b) Tests based on solving the model for its so-called "final form" and checking whether the restrictions implied by various models of expectation formation, including RE, are satisfied. Such tests are difficult to carry out in multivariate cases involving more than one expectational variable. We know of only two instances, both univariate. (c) Tests of RE based on the restrictions imposed on the structural form of the model. Finally, in (d), we consider direct tests based on comparing observations of reported expectations with subsequent realizations, either from survey data or from experimental data.

4.2.1. *Minimalist tests*

The idea of a "minimalist" test is simple: Generally, a model of dynamic decision making under uncertainty will give rise to several related behavioral relationships. If we specify a model of expectation formation independently of the behavior optimized, such expectations will depend on observed information, frequently lagged values of variables appearing in the model. Inserting the expectations in an equation thus gives rise to a

---

[16] This section draws on [Nerlove and Fornari (1997)].

distributed lag (DL) relationship, or expectational distributed lag (EDL). One minimal characteristic shared by all models involving distributed lags of an expectational nature (EDL) is that the variables subject to EDL enter all related behavioral equations with exactly the same DL distribution [Nerlove (1958b)]. If they do not, we would be led to reject expectations as a source of the lagged behavior. In [Nerlove (1958a)], this criterion was the basis for a test of Friedman's permanent income hypothesis (PIH) against the alternative of DL due to habit persistence in a system of consumer demand functions. The PIH was rejected in this case. Closer to home, Orazem and Miranowski (1986) deal with acreage allocation decisions of Iowa farmers, 1952–77, for four crops – corn, soybeans, hay, and oats – accounting for all but a minute portion of harvested acres in Iowa, and clearly reject EDL. If EDL is rejected, indirect tests of any model of expectation formation cannot be carried out in the context of the behavioral model; thus, EDLs are almost always *assumed.*

4.2.2. Tests based on "final form" VAR or VARMA models

Hansen and Sargent (1981) clearly state the need for multivariate RE models to test the restrictions imposed across behavioral equations, as well as between behavioral equations and the stochastic processes generating the exogenous variables of the model. Multivariate RE models can generally be reduced, at least approximately, to multivariate vector autoregression (VAR) or vector autoregressive moving average (VARMA) models embodying very large numbers of restrictions of an extremely complicated sort. In an earlier paper, Hansen and Sargent (1980) suggest basing a test of REH on the restricted vs. the unrestricted VAR. In his study of land allocation in Egypt between cotton and wheat, Eckstein (1984) formulates the problem as a univariate one and tests the implied restrictions, but we know of no similar attempts in a multivariate context. The reason is not only the complexity of the restrictions resulting from the REH but also the failure of identification when the underlying dynamic structure is not pinned down rather precisely (Wallis). Therefore Nerlove (1972), Nerlove, et al. (1979), and Sargent (1981), among others, stress the need for explicit dynamic optimizing models to lay the basis for specifying which anticipated future values matter, and what other leads and lags are involved in the structural behavioral relationships to be estimated.[17]

4.2.3. Tests in a structural context

It is easier to impose a priori structural restrictions in structural than reduced-form estimation. Restrictions which are relatively simple and transparent in structural terms

---

[17] Despairing of being able to do this satisfactorily led Sargent and Sims (1977) to recommend the use of unrestricted VAR models to check consistency with several possible structural models. The model of beef cattle supply formulated in [Nerlove and Fornari (1997)] is based on an explicit model of dynamic optimization which does permit very precise specification of the future values entering each of several behavioral relationships as well as of the other lags involved, and thus, in principle, leads to a satisfactory test of the REH. The problem of complex restriction in the VARMA form leads the authors to formulate a partial test in a structural context.

are extremely complex in terms of the final form equations (approximating VAR or VARMA). Wallis (1980), Hansen and Sargent (1981), and Fair and Taylor (1983, 1990), among others, clearly recognize the need for structural specification and estimation in this context, but only Eckstein (1984) and Goodwin and Sheffrin (1982), the latter only partially, seem to have carried it out, and then only in a univariate context. The results are inconclusive. The problem is that even in the case of univariate structure, restrictions across the behavioral equation and the stochastic equations generating the exogenous variables, anticipated future values of which affect behavior, are extremely difficult to impose, and, indeed, we would argue, highly problematic in any case, due to the difficulty of correctly specifying the structure of the processes generating these variables (see [Nerlove et al. (1979, pp. 201–290)], for general methods of specifying appropriate univariate and multivariate time-series models). QRE circumvents this difficulty by separating the stochastic relations generating the exogenous variables from the structural behavioral relationships, thus permitting cross-behavioral restrictions to be taken into account more easily, and minimizing contamination due to errors in specifying that part of the model determining the exogenous variables, albeit at the cost of full efficiency. But full efficiency is predicated on correct specification and, for this reason, is rarely a high priority in econometric practice. Nerlove and Fornari (1997) carry out such a test. Some of the structural restrictions are accepted, some not; the results are again inconclusive.

4.2.4. *Futures-based models of expectation formation*

In a pioneering paper, Gardner (1976) suggests using futures prices for all expectations in agricultural markets. Gardner presents evidence that "Futures prices can be valuable as an adjunct to and as a vehicle for evaluating lagged-price, lagged dependent variable models".

Gardner (1976, p. 81) writes:

> ... an alternative approach to estimating supply elasticity ... [is] to exploit the theoretically well-grounded hypothesis that the price of a futures contract for next year's crop reflects the market's estimate of next year's cash prices. Since the appropriate price for supply analysis is the price expected by producers at the time when production decisions are being made, a futures price at this time is a good candidate for a directly observable measure of product price in supply analysis.
>
> In the context of crop supply, there are several problems to be faced in the use of futures prices. First, "the market's" estimate as given by a futures price reflects the expectations of nonfarm speculators as well as crop producers, and it reflects directly the expectations only of those crop producers who themselves make futures transactions. Second, there is the issue of which futures contract is most appropriate. Third, at what date should the futures price be observed?
>
> With respect to the first issue, the use of a futures price can be justified by the hypothesis of rational expectations as developed by John Muth. Under rational

> expectations, there is no reason for farmers to have different price expectations from futures speculators, nor for farmers who make no futures transactions to have expectations different from those who do. If the price expectations of those out of the futures market differ from the futures price, there is great incentive for them to enter. Thus, those out of the market likely have price expectations similar to the market price of futures.
>
> The second issue should cause no serious problem so long as the futures contract pertains to the new crop. Of course, even old-crop cash prices are influenced by expectations concerning the new crop. But the cash-futures basis changes from year to year and secularly as the cost of storage (which includes interest) changes. The present analysis uses the first futures price after the crop is in.
>
> The third problem is most difficult because it is not clear exactly when the production decision is made. There may not be any preharvest date at which a farmer can be said to have made irrevocably his decision about planned output. Even after the crop is planted, planned output can be revised and actions taken accordingly in fertilization, pest control, and other practices, such as plowing under a crop or using it for forage. However, the main production decisions are the choices of acreage and techniques to follow in planting. This suggests taking as the expected price the futures price in the period immediately preceding the planting season.

The evidence supporting Gardner's suggestion is mixed. Eales et al. (1990) study price expectations for Illinois corn and soybeans market participants (farmers and grain merchandisers). Surveys of individuals are aggregated into groups. The mean response of the groups shows no significant difference from the nearby futures price; whereas some difference was noted in price variance. Tronstad and McNeil (1989) show that sow farrowing response models using futures price as the expected price compare favorably with similar response models using cash prices. Just and Rausser (1981) provide evidence that futures markets offer superior forecasts of subsequent cash prices when compared to forecasts from large-scale econometric models for several U.S. crops (their results on cattle prices, however, offer much weaker, if not conflicting, evidence on the superiority of futures prices relative to the large-scale model's forecasts).

Other evidence conflicts with Gardner's suggestion [Leuthold (1974), Martin and Garcia (1981), Goodwin and Sheffrin (1982)]. Chavas et al. (1983) offer an interesting discussion of their results on corn and soybean acreage supply response, which is not unlike the suggestion offered earlier by Working (1942), Johnson (1947), and Tomek and Gray (1970) (see the discussion above): "... as argued by Gardner, the futures price appears to be a good substitute for the cash price lagged one year in supply analysis. This is the case for corn and soybeans because the two prices are highly correlated and appear to reflect similar market information. As a result, using both futures and cash prices in supply equations may lead to multicollinearity problems, while deleting one of the two appears to make little difference in estimates of supply elasticities" [Chavas et al. (1983, p. 32)].

Antonovitz and Green (1990) study supply response for fed beef, focusing on naive expectations, ARIMA-type expectations, futures price expectations, and fully rational price expectations. They summarize their results: "... empirical evidence does not support any one model in particular, suggesting that expectations are heterogeneous rather than homogeneous" [Antonovitz and Green (1990, p. 473)]. They conclude with the interesting note, based on root-mean-squared error (rmse) measures on fit models: "the highest value (of rmse) was observed for the rational expectations model perhaps adding additional evidence to reject the hypothesis that expectations are formed rationally" [Antonovitz and Green (1990, p. 485)].

The question of how good an indicator the futures price is of the future cash price was addressed for live cattle (a nonstorable commodity) in [Leuthold (1974)]. There he finds that for distant contracts, "The cash price is a more accurate indicator of future cash price conditions than is the futures price". He concludes, "The producer who looks at the futures prices routinely to establish a feeding margin so that he can decide whether or not to purchase and feed cattle may receive false signals and be misled into a costly decision, either a money loss or foregone profits" [Leuthold (1974, p. 276)].

Covey and Bessler (1995) consider the question of predictability of cash prices using the information in current and past futures prices as a problem of cointegration.[18] They find that daily cash and futures prices for a storable commodity, corn, are cointegrated; however, this long-run relation does not offer any out-of-sample forecast improvement relative to that contained in past cash corn prices, a result which agrees with Working's conclusion (1942). On the other hand, they find no cointegrating relation between cash and futures prices for live cattle prices. Further, short-run forecasts (on data not used to test for cointegration) of cash prices of cattle are improved by conditioning on past futures prices of cattle – suggesting that there is (short-run) information relevant to future cash prices in the current futures price, which is not in the current cash price.

There have been several studies contrasting the forecasting ability of the futures market relative to a publicly available alternative forecast. Just and Rausser (1981) find that forecasts of cash prices made by several commercial forecasting companies were generally not superior, in a mean squared error sense, to corresponding futures market prices. Futures on soybean meal and oil are superior forecasts of subsequent cash prices relative to the econometric forecasts, indicating that the futures market did capture some information which was not captured by 1970s-style econometrics. However, this general superiority (of futures market forecasts) does not extend to forecasts of livestock prices: "some of the econometric forecasts seem to be preferable for livestock commodities ..." [Just and Rausser (1981, p. 207)]. Martin and Garcia (1981, p. 214) find "... the performance of the cattle and hog futures as a rational price formation agency is suspect," a result which supports the finding of Bessler and Brandt (1992), who find that

---

[18] Two variables are said to be cointegrated if they each have one or more unit roots (stochastic trend) and if one can find a linear combination of the two which has fewer unit roots (e.g., each has one unit root, and the regression of one on the other has residuals which are stationary). It is usual to interpret such a result in causal terms. See [Stock and Watson (1988) and Hamilton (1994)].

a commodity expert is root mean squared error superior in forecasting cash cattle prices relative to live cattle futures prices, over a fifteen year period – this forecaster's advantage being reduced forecast error variance, with no advantage found due to reduced bias.

The evidence from futures markets is mixed, but two points emerge. First, futures on livestock prices (apparently) do not capture important long-run information for subsequent cash prices. This finding is supported by empirical studies using econometric models [Just and Rausser (1981), Covey and Bessler (1995)] and by studies of expert opinion [Bessler and Brandt (1992)] and of actual lagged cash prices [Leuthold (1974), Martin and Garcia (1981)]. Second, while futures on storable commodities are able to outperform econometric models, it is not clear that they can outperform optimal univariate ARMA or ARIMA model predictors of cash price. Working's initial thoughts on the subject and Johnson's summary appear not to have been seriously challenged by subsequent analysis. The econometric results of Just and Rausser may be more a criticism of 1970s-style econometrics on data characterized by unit roots than an empirical endorsement of futures markets as forecasts of subsequent cash prices, and is indeed supported by some [Covey and Bessler (1995)]. Of course, these results also suggest the possible superiority of futures prices over poorly specified commodity models.

Gardner's results remain to be explained. He finds that supply response models on two storable commodities using futures market prices as proxies for expectations perform as well as a lagged price expectations models. One argument is that subjects do indeed look at the current cash market and make planting decisions based on that variable (as suggested by Schultz and Brownlee (1941–42); see our discussion below); however, Gardner may have misrepresented such expectations by using the average price from the previous year to represent the most recent cash price. This explanation is not inconsistent with the position taken by Working (1942), Johnson (1947), and Tomek and Gray (1970), and is supported by the empirical result of Covey and Bessler (1995) – that no long-run forecast information, in addition to that contained in current cash price, is in the current futures price. The lagged average price over the previous year may not have been optimal, and the futures market (April or May quote for January delivery) may be closer than the lagged average price to an optimal statistical predictor, which farmers may have been using. Of course, the difficulty in assessing expectations indirectly makes all explanations tentative.

4.3. *Tests based on direct observation*

4.3.1. *What can we learn by asking people what they expect?*

Below we summarize attempts to study agents' expectations by directly observing them. These works have been both experimental and nonexperimental or observational, for example in surveys or informal interview studies, where no experimental control is imposed in the collection of the data. Here, we offer a brief discussion on the question of what one can hope to learn by asking people what they expect.

Merely asking a person what he expects a variable to be at some future date has met with considerable skepticism from both decision theorists and psychologists. Savage is an early critic of direct interrogation:

> Attempts to define the relative probability of a pair of events in terms of the answers people give to direct interrogation has justifiably met with antipathy from most statistical theorists. In the first place, many doubt that the concept "more probable to me than" is an intuitive one, open to no ambiguity and yet admitting no further analysis. Even if the concept were so completely intuitive, which might justify direct interrogation as a subject worthy of some psychological study, what could such interrogation have to do with the behavior of a person in the face of uncertainty, except of course for his verbal behavior under interrogation? If the state of the mind in question is not capable of manifesting itself in some sort of extra verbal behavior, it is extraneous to our main interest. [Savage (1954, p. 27)]

Evidence supporting criticism of direct interrogation (interrogation without explicit motivation) comes from the work of the experimental psychologists, Siegel and Goldstein (1959). They hypothesized that observed behavior of participants in an experiment in which no financial incentives were provided, which was inconsistent with assumptions of an underlying theory, may have been due to boredom or game-playing by the experimental subjects. Further, such boredom might be overcome by providing subjects financial rewards, which were directly related to the "goodness" of their responses in experimental tests of an underlying theory. Their experiments show differences between subjects' responses who received financial rewards relative to subjects' who received none. The former are closer than the latter to the a priori predicted response. Davis and Holt (1993) summarize their assessment of the Siegel–Goldstein experiment:

> What we may conclude from this experiment is that financial incentives can sometimes eliminate subtle and unintended biases. For this reason, the payment of financial incentives is a critical element in the administration of economics experiments. [Davis and Holt (1993, pp. 88–89)]

Just how one provides an incentive has been a subject of considerable study. One doesn't want the incentive mechanism itself to induce strategic behavior to mask or misrepresent a subject's beliefs. Scoring rules are measures of goodness used to encourage the assessor to be honest in reporting his true beliefs. Since these beliefs exist solely in the assessor's mind, there is no way to determine whether or not this requirement is satisfied. However, by rewarding or penalizing the assessor according to certain scoring rules, one can encourage an assessor to make his stated beliefs correspond to his true beliefs. Scoring rules have been developed for elicitation of probabilistic beliefs and as such provide a natural application to testing rational expectations. In Muth's words, "... the subjective probabilities of the agent are distributed around the objective probability of the data".

Consider a task where an individual must make a probability assessment [$r' = (r_1, r_2, \ldots, r_n)$] for an event E which consists of n mutually exclusive and collectively exhaustive outcomes $E_1, E_2, \ldots, E_n$. The quadratic rule, $Q(d,r)$, is defined as

$$Q(d,r) = \left[1 - \sum_{l=1}^{n}(r_i - d_i)^2\right],$$

where $d_j = 1$ if event E_j occurs, and zero if not. $Q(d,r)$ encourages the assessor to set r (his revealed beliefs) equal to his true probabilities (p). So that if a risk-neutral assessor is rewarded according to a quadratic rule, his or her optimal response is to set r (his or her vector of stated or revealed beliefs on events 1 through n) equal to p (his or her true beliefs on outcomes 1 through n). As the range of the quadratic rule, as given above, is $[-1, +1]$ and one may not wish to entertain negative payoffs, the quadratic rule is used in the form

$$Q(d,r)^* = \left[1 - (1/2)\sum_{l=1}^{n}(r_i - d_i)^2\right]$$

which has the range [0, 1]. The quadratic rule has been applied in weather forecasting, where it is labeled the "Brier score" (see [Murphy and Winkler (1970), and Brier (1950)]). De Finetti has suggested the quadratic rule as a motivational device for testing responses in educational psychology [de Finetti (1965)].[19]

One immediate consequence of the scoring rule literature is that rewards should be designed relative to the utility function of the individual subject. This point is given

---

[19] Other proper scoring rules exist, some of which are reviewed in Murphy and Winkler (1970). Perhaps the most interesting of these is the logarithmic (log) rule, defined as

$$L(d,r) = \left[\ln\left(\sum_{l=1}^{n} d_i r_i\right)\right],$$

where d and r are defined as above and ln is the natural logarithmic operation. Shuford, Albert, and Massengill (1966) show that the log scoring rule is the only proper scoring rule which gives payoffs just in terms of the stated probability of the event which actually occurs (notice that the quadratic rule defines payoffs in terms of probabilities assessed to both the event which obtains as well as all other events).

The log rule presents analysts with an interesting "problem" as its range is $[-\infty, 1]$. When the assessed response is $r_j = 0$ and $d_j = 1$, the negative infinity reward is difficult to work with in applied settings. Following Shuford, Albert, and Massengill (1966, p. 137), a truncated log scoring rule is

$$L(r_k, d_k) = \begin{cases} 1 + \ln r_k, & 0.01 < r_k \leqslant 1, \\ -1, & 0 \leqslant r_k \leqslant 0.01. \end{cases}$$

Bessler and Moore (1979) suggest a version of $L(r_k)$, where all payoffs are positive. Of course truncation will potentially induce responses for which $r' \neq p'$, especially in the neighborhood of truncation. Shuford et al. (1966) study the effects of truncation and conclude, "... for extreme values of p_i, some information about [the subject's] degree-of-belief is lost, but from the point of view of applications, the loss of accuracy is insignificant" (1966, p. 137).

consideration in [Murphy and Winkler (1970) and Holt (1986)]. They suggest the direct elicitation of each subject's utility function and design of a scoring rule (motivational device) which yields optimal responses for that particular utility function.[20]

In a more pragmatic vein, Nelson and Bessler (1989), following work first reported in [Nelson (1987)], pre-screen individuals for linear utility (in the range of rewards offered in subsequent probability elicitation experiments). This "pre-screening" allows them to use the familiar quadratic rule (discussed above), but requires them to drop nearly 60 percent of their original subject pool, as the dropped subjects exhibited significant nonlinear utility. They provide an empirical test of the quadratic scoring rule when used in comparison to a linear rule: $H(r_k, d_k) = a_0 r_k d_k$; where $d_k = 1$ if event k obtains, and zero otherwise. The rule defined by H is improper and should induce subjects not to reveal probabilities $r' = p'$. In fact, the optimal response, for a risk-neutral subject facing H, is to find a corner point solution, which disguises the single highest probability as a one and all lower probabilities as zeros (see [Nelson and Bessler (1989, pp. 364–365)]. They find that for risk-neutral subjects, "... the scoring rule used (linear or quadratic) had a significant effect (p-value < 0.0001) on the number of zeros used in a forecast when the observations from eight subjects in each treatment over all forty forecast periods were used". The Nelson and Bessler work provides evidence to suggest that the way subjects are paid is an important consideration in assessment studies.[21]

4.3.2. Experimental data

In the concluding section, we discuss the mixed findings from surveys of expectations from both individual decision makers and commodity experts. Here we deal with laboratory studies of expectation formation, in which monetary payoffs can be tied directly to the assessment and subsequent realization and rewards made according to the loss or utility function of the respondent with control for what the subject saw prior to the assessment. Nelson (1987), Nelson and Bessler (1992), Hey (1994), Dwyer, Williams,

---

[20] In this context, scoring rules are viewed as *ex ante* motivational devices, which aid in helping a subject make his stated beliefs correspond with his "true" beliefs. There is a rich parallel literature in which these same rules are used to evaluate the ex post "goodness" of probabilistic forecasts. For discussion of such see [Kling and Bessler (1989)]. Applications can be found in [Diebold and Rudebusch (1989), and Zellner et al. (1991)].

[21] The issue of payoffs is important as one wants the motivational device to be sensitive to responses which deviate from optimal subject response. That is to say, if the payoff rule is flat in the neighborhood of the optimal response, the subject may have little incentive to respond with precision. We may erroneously conclude from an experimental study that agents respond in a suboptimal manner because the reward mechanism is for all practical purposes flat in a sizable neighborhood of the optimal response. Murphy and Winkler (1970) explore this issue with respect to the log scoring rule (which is particularly flat in a neighborhood of the optimal response). Essentially the same issue has been revisited by the experimental economist Harrison (1989, 1992). Harrison (1989, p. 759) argues: "... anomalies observed in the experiments in question may simply reflect the failure of the experiment to meet widely accepted sufficient conditions for a valid controlled experiment ... the result of this failure is simply that the opportunity cost of 'misbehavior' in these experiments is, by any reasonable standard, minuscule".

Battalio, and Mason (1993), Plott and Sunder (1988), and Williams (1987) are included in the list of experimental studies which attempt to isolate expectation formation from the myriad of other confounding influences that plague reliable inference on expectations formation.

Perhaps the most elaborate experimental study of expectations in market data has been that carried out and reported by Williams (1987). Here agents are provided information on their own (induced) limit price, and they have information on all previous prices generated in the marketplace and all accepted and unaccepted price quotations from these previous periods. Further, they know the total number of market participants and can infer the number of buyers and sellers from price signals sent in the market. Subjects were asked to forecast the mean price they expected to occur in recursive trading periods ($t = 2, 3, 4, 5$); actual trading of units at negotiated prices followed in the manner carried out by subjects in period $t = 1$. The induced supply and demand arrays were constructed to be stationary across periods for the explicit purpose of reproducing in the laboratory a "theoretical steady-state" [Williams (1987, p. 4)], thus, providing opportunity for subjects to learn across time. The inducement of limit prices (costs and values) is important as it provides subjects with structural information (their own cost or value structure) about the underlying market, not dissimilar to the market information behind a Muthian rational agent's behavior. Other experimental studies (see below) of rationality provide the subject with historical prices (realizations). This additional bit of information (individual valuations) and the fact that agents act in a market make Williams' study particularly interesting. A $1.00 forecasting-accuracy payment was paid to the subject with the lowest summed absolute forecast error at the end of the experiment. While this inducement is probably not proper (as we are provided no information on the underlying preference structure of the subjects, e.g., we do not know if they were risk neutral or risk averse), Williams suggests that his reward scheme was sufficient for subjects to take the assessment task seriously and he makes no mention that the reward motivated strategic responses that masked "true" expectations [Williams (1987, p. 7)]. Forecasts from 532 observations from this experiment were found to be biased, based on ordinary least squares regression on pooled time series cross section data in the actual price in period t on the individual forecast of that price made at the end of period $t - 1$. Tests on subsamples trading on individual periods (e.g., $t = 2$; $t = 3$; $t = 4$; $t = 5$) showed biased forecasts as well; although trading for experienced subjects (subjects who had participated in double auctions before this round of experiments) in the last period ($t = 5$) were not found to be biased, suggesting perhaps that experience and learning (in a stable environment) may result in rational expectations. Williams concludes his work as follows: " Using price forecast observations obtained from 146 participants in twelve separate experimental double-auction markets, little empirical support is found for strict Muthian rational expectations assumptions. The forecasts are biased estimates of the realized mean price and forecast errors display significant first-order serial correlation" [Williams (1987, p. 16)]. Williams' results are replicated in double auction asset markets in [Smith et al. (1988)].

Earlier, Schmalensee (1976) presented a total of twenty-three subjects with price observations from a nineteenth century British wheat market, and had them submit both point and interval forecasts of five-year averages of the price series. An adaptive expectations model was found to outperform an extrapolative expectations model, with the response speed in the adaptive model tending to fall at turning points. Schmalensee did not study rational expectations.

More recently, Nelson and Bessler (1992) tested quasi-rationality of laboratory subjects, who were pre-screened for linear utility and motivated through payments from a quadratic scoring rule. Subjects were shown forty to sixty earlier realizations from Monte Carlo generated data on one of five univariate processes: an autoregression of order one (AR1), an autoregression of order two (AR2), a random walk (RW), an integrated moving average process of order one (IMA1), and a subset autoregression of order four (AR4). They were then asked to provide recursive probabilistic one-step-ahead forecasts of the next 40 to 60 data points – with actual outcomes and payoff numbers being revealed on a computer screen sequentially throughout the forecasting exercise. The expected value for each forecasted distribution from each subject was taken to be his expectation. Both individual and aggregate performances were judged by three criteria: Are forecast errors significantly different from zero? Are the forecast errors correlated through time? And do forecasts result in significantly larger mean squared errors relative to forecasts from a minimum mean squared error predictor applied to historical realizations (can the human forecasters perform as well as a model, in a mean square error sense, when both see the same historical observations?)? Table 1 is a summary of the performance tests for aggregates for each of the five time series processes.

Interestingly, the AR1 and RW processes are forecasted well under all three tests. If we apply a standard .05 significance level, the aggregate forecasts from these "simple" series appear to be quasi-rational. However, the more complicated series, the AR4 (a subset AR4) and the IMA1 (which is of course a nonlinear model), are not forecasted as well. The forecasts of the AR4 process fail all three tests, and the forecasts of the IMA1 process fail two of the three tests.

Results for individuals offer less support for quasi-rationality. Over the entire 41 sets of forecasts (8 individuals forecasted the AR1 process; 10 forecasted the AR2 process; 8 forecasted the RW process; 9 forecasted the IMA1 process; and 6 forecasted the AR4 process), 30 individuals passed the bias test, 26 passed the white noise residuals test, and

Table 1
Performance of aggregate forecasts in five experiments [Nelson and Bessler (1992)]

	AR1	AR2	RW	IM1	AR4
Bias test p-value	.133	.891	.494	.264	.020
White noise test p-value	.532	.012	.317	.053	.017
MSE test p-value	.128	.145	.092	.004	.010

8 passed the mean squared error test. The Nelson and Bessler (1992) results appear to be consistent with the results from the survey literature [Zarnowitz (1983)]: aggregates are more likely to generate forecasts which pass tests of rationality (quasi-rationality) than are forecasts of individual agents. This finding appears to be consistent with a line of research on forecasting in general, which finds that aggregates or composites of forecasts outperform individual forecasts (models or people) in out-of-sample forecast evaluations; see [Granger (1989)].

Additional experimental work has followed Nelson (1987) in testing expectations formation on univariate processes. Dwyer et al. (1993) ask subjects to provide one-step-ahead forecasts of a univariate random walk: $x_t = x_{t-1} + \varepsilon_t$, where ε_t is distributed normally with mean zero and variance 1.0. Subjects were motivated with financial rewards, which depended upon both their reported forecast and its ultimate realization. There is no mention that the incentive structure was matched to the utility function of respondents, so we are not able to comment on the possibility of subjects reporting expectations that masked their underlying beliefs. They find that expectations are well described as rational expectations, a random walk embedded in an additive error. This study was extended in [Beckman and Downs (1997)] under four alternative levels of noise. Here the error term was uniformly distributed under four alternative treatments: treatment I, $\varepsilon_t \sim u[-5, +5]$; treatment II, $\varepsilon_t \sim u[-10, +10]$; treatment III, $\varepsilon_t \sim u[-15, +15]$; and treatment IV, $\varepsilon_t \sim u[-20, +20]$. Here each subject received all four treatments; each was randomly assigned to one of the 24 different possible orders in which the four treatments could have been presented (order I, II, II, IV was different from order I, III, II, IV, etc.). Payments were not (apparently) matched to the utility function of the subjects, and followed the rule:

$$\text{Payment} = \$11.00 - \sum_{t=1}^{100} |A_t - F_t|,$$

where A_t is the actual value of x_t and F_t is the subject's forecast of x_t in period t. The hypothesis of interest is that under rational expectations, there should be no difference in the rational expectations forecasts across treatments. What they find is (1993, p. 598): "... increasing the variance of a random walk does create a more diffuse set of deviations from theoretically correct behavior. A one percent increase in the standard deviation of the random error generates a 0.9% increase in the standard deviation of the forecast about the rational expectation."

Hey (1994) considers 48 subjects' assessments on three univariate time series. Each assessor is provided monetary motivation, which is a function of the actual realization of the variable to be forecasted and each subject's forecast, such that risk-neutral agents will be motivated to set their forecast equal to their expected value of the random variable. The first two series are univariate autoregressions; while the third series is an autoregression, which exhibits a structural break partway through the forecast period. Tests of rationality are rejected; however, post hoc data analysis seems to suggest a type of extrapolative expectation scheme, which was sensitive to the particular characteristics of the underlying series. Hey concludes:

> So our statistical tests reject the detailed rational expectations hypothesis, though the general flavor of the data support the general notion that the subjects were trying to be rational in a broader sense. It would appear that the subjects had some general model of the Data Generating Process in their minds which they were using in a broadly sensible fashion. More importantly this general model appeared to be series specific; so that subjects had a different "model of the world" for Series 2 than they had for Series 3. [Hey (1994, p. 20)]

Additional work on experimental tests of rationality exists, some of which is reviewed in [Swenson (1997)]. He concludes: " It is safe to conclude from these (and many other) studies that individuals' forecasts of prices almost never satisfy RE (rational expectations) ... " [Swenson (1997, p. 434)].

4.3.3. Survey and semi-survey data

Most work on expectations data in economics has been nonexperimental, without the use of explicit benefit or motivational devises. A vast literature exists on the analysis of surveys. Several surveys in agriculture were conducted in the pre-Muthian period and, in fact, were cited by Muth as providing support for rational expectations. Heady and Kaldor (1954) studied over one hundred Iowa farmers over a three-year period, 1947–1949. They tested no particular models of price expectations, but their impressions from their surveys are suggestive:

> No attempt was made in this study to test alternative models used by farmers. Nevertheless, certain impressions were gained while interviewing farmers. No single procedure was employed by all farmers. Moreover, the same farmer often used more than one procedure, depending upon the amount of information possessed and upon the degree of confidence attached to it. In December 1947, some producers were using a simple "parallel" model for their long-range forecasts which implied that prices following World War II would decline as they did after World War I. Other farmers were using a model giving explicit recognition to the supply of corn as a price-making variable. For their 1948 and 1949 forecasts the majority was not using simple mechanical models such as the projection of the current price or recent price trend into the next year but was attempting to analyze and predict the more complex price-making forces. A rather common procedure appeared to start the process of devising expected prices from current prices. The current price then was adjusted for the expected effects of important supply-and-demand forces. [Heady and Kaldor (1954, p. 35)]

Earlier, Schultz and Brownlee (1941–42) considered expectations of Iowa farmers on 1940 corn yields and hog prices. In a sample of 200 farmers, Schultz and Brownlee find that expectations on yield were as follows:

> ... expectations are not marked up by farmers to the level of recent experience. Instead, recent increases in yields in corn in Iowa are discounted about one-half. The other half is looked upon as a real gain, one which farmers anticipate will continue to be forthcoming, a gain which farmers ascribe to improvements in management practices, hybrid corn, and to the reduction in corn acreage which was occasioned by the AAA, and which resulted in the less productive land being taken out of corn. [Schultz and Brownlee (1941–42, p. 496)]

Later, Bessler (1982) argued that similar expectations behavior would characterize optimal expectations for California and Indiana crop yields. He finds that such yields follow a (0, 1, 1) ARIMA process and expectations of such might be described by the "permanent yield hypothesis":

> ... we can say that for these yield series a notion of permanent yield might be a useful concept ... that is, farmers forming optimal expectations on these yield series might view yield as composed of both permanent and transitory components ... such behavior might be justified if one notes that specific changes in yield might be viewed as permanent in that they reflect basic changes in technology (new crop varieties, pesticides, and herbicides), whereas other changes might reflect year-to-year variability in weather. [Bessler (1982, p. 22)]

Schultz and Brownlee's survey of 97 hog farmers (1941–42) reaches a different conclusion on the process generating hog price expectations. They note that fluctuation in hog prices "are both numerous and irregular. This behavior of prices probably accounts for the strong preference which Iowa farmers show for current prices in formulating their price expectations". They continue:

> Iowa farmers in March 1940 were operating on the assumption that hog prices would continue at about the exceedingly low levels which then prevailed. Changes in supplies, the outbreak of the war, and the two and one-half year decline in hog prices apparently had not been instrumental either in further depressing or lifting prices which farmers anticipated for the hogs which were being farrowed at that time. [Schultz and Brownlee (1941–42, pp. 495–496)]

Surveys conducted post-Muth (since 1961) have generally confirmed much of the qualitative findings of Schultz and Brownlee. These efforts included analysis at both the aggregate level and at the micro- (individual agent) level. Here analysts have been interested in whether agents' expectations are unbiased forecasts of the random variable of interest and whether errors from such forecasts are uncorrelated with information available to the forecaster at the time of the forecast. That is to say, interest has focused on (a) whether $\mathrm{E}\{\varepsilon_t = (p_t - {}_{t-k}p_t)\} = 0$, where here ε_t is the forecast error based on forecast ${}_{t-k}p_t$ of an individual agent or the aggregate of forecasts of a group of agents on endogenous variable p made at period $t-k$ for realization at period t, p_t is the actual realization of that same endogenous variable at period t, and E is the expectation operator; and (b) $\mathrm{E}\{\varepsilon_t \mid \Omega_{t-k}\} = 0$, where Ω_{t-k} is the set of all available information available to agents at time $t-k$. While Schultz and Brownlee did not formulate such a general set of hypotheses (they focused attention on three particular scenarios, one having price (yield) falling, one having price (yield) remaining constant, and one having price (yield) increasing over the next year), they did show quite clearly that for their sample of farmers, and for their particular year, differences were present in the behavior of agents in forming expectations of yields versus prices. In words not used by Schultz and Brownlee we might say expectations on corn yields followed a nonrandom walk process; whereas expectations on hog prices appeared to follow a random walk.

Studies of aggregate expectations include Carlson (1977), Turnovsky (1970), Jacobs and Jones (1980), Zarnowitz (1983), and many others (see [Pesaran (1987)], for a survey through the mid 1980s).

Zarnowitz (1983) found that in forecasting numerous aggregate economic time series, individual experts participating in the quarterly National Bureau of Economic Research and American Statistical Association survey of business conditions perform worse than group average forecast. He finds:

> ... it is difficult for individuals to predict consistently better than the group ... for most people, most of the time, the predictive record is spotty ... a series of group averages has the advantage that it is helped by cancellation of individual errors of opposite sign. [Zarnowitz (1983, p. 17)]

Studies of aggregate expectations in agriculture include Bessler (1980), Ravallion (1985, 1987), Runkle (1991), Garcia and Leuthold (1992), and Colling, Irwin, and Zulauf (1992). Runkle (1991) finds that farmers' reported expectations of sow farrowings are not rational forecasts of sow farrowings, and suggests that such a result may be less due to underlying irrationality and more due to motivation (or lack thereof) in the assessment survey. Runkle (1991, pp. 599–600) writes:

> Although it may be somewhat surprising that farmers announce irrational forecasts of their own future actions, it would be considerably more surprising if market analysts were to announce irrational forecasts of farmers' actions. Because the market analysts, unlike farmers, are paid for the accuracy of the forecasts they report, they have a strong economic incentive to report accurately.

This suggestion follows the earlier suggestion of Keane and Runkle (1990):

> The survey data include only forecasts from professional forecasters, who have an economic incentive to be accurate. Because these professionals report to the survey the same forecasts that they sell on the market, their survey responses provide a reasonably accurate measure of their expectations. [Keane and Runkle (1990, p. 715)]

Keane and Runkle (1990) and Runkle (1991) provide no evidence that "their survey responses provide a reasonably accurate measure of their expectations". Faith in the market to induce reasonably accurate forecasts might lead one to conclude that astrologers are reasonably accurate because they sell their forecasts in the market!

There is related evidence that professionals are no better in assessing the future than nonprofessionals. Stael von Holstein (1970) found meteorologists' assistants outperformed the meteorologists (who were paid for their expertise) in simple probability forecasting – the latter tended to give tight forecasts, while the former gave diffuse distributions.

Earlier in this century H.A. Wallace gave a spirited summary of a not unrelated point of using experts in judging corn yields:

> That the corn judges did not know so very much about the factors which make for yields is indicated by the fact that their scores were correlated with yield to the extent of.2. The difficulty seems to be that they placed too much emphasis on length of ear and possibly also some fancy points, which caused them to neglect placing as much emphasis on sound, healthy kernel characteristics as they should. [Wallace (1923, p. 304)]

Wallace goes on to suggest that "the things which really are in their [the judges'] minds are considerably different from... [those which they] professed" [Wallace (1923, p. 304)].

Following Nelson and Bessler (1989), discussed above, it is not just a matter of payment – how one is paid is not unrelated to what one says.

Bessler (1980) finds that the means of aggregate subjective probability distributions of farmers on yields of California field crops are not significantly different from the one-step-ahead forecasts of yields from ARIMA representations of historical county-level yield data; however, higher moments of the aggregate subjective distributions do not match their time series representations. These farmers were not paid for their responses.

Ravallion (1987) finds that daily rice price expectations from a sample of twenty-eight Bangladesh traders (Aratdars) fail both tests of unbiasedness and orthogonality. Ravallion offers possible reasons for these rejections:

> All interviews were done in Bangla by a single interpreter under reasonably close supervision, particularly in the early stages.... Although a good deal of care was exercised in collecting these data, it seems likely that the results overstate the level of agreement amongst the traders. The interview process can act to transmit information between traders. This is produced by the tendency of an interviewer to form expectations of the answers on the basis of previous interviews which are then used as prompts. [Ravallion (1987, p. 132)]

It would appear that this same criticism offered of Ravallion's survey would apply to the results found in [Bessler (1980)] as well. (Much recent experimental work has adopted computer technology to help in collecting expectations and reporting them without introducing the potential bias associated with the use of human interviewers; see [Davis and Holt (1993, p. 23)] for a general discussion of computers in experimental economics.)

Nerlove (1983) suggests that analysis of aggregate expectations is but a first step in a more elaborate program of analysis of expectations:

> While the use of aggregates derived from surveys is an important first step in the analysis of expectations and plans, such analysis should be supplemented by studies based on the micro-data themselves for several reasons: First, the micro-data should be consistent with hypotheses regarding the behavior of the aggregates; for example, expectational aggregates could provide unbiased forecasts of realized aggregates, as asserted by the theory of rational expectations, yet forecasts of individual agents could be systematically and persistently biased. Second, some factors affecting deviations between expectations or plans and subsequent realizations may affect all individuals simultaneously yet vary from period to period and some factors may affect individuals; only through analysis of the micro-data can we disentangle those two groups of effects. Finally, individual variation in variables related to expectations, plans, and realizations may be reduced or obscured in aggregate data. [Nerlove (1983, p. 1256)]

The experience from analysis of non-agricultural micro-data supports Nerlove's recommendation to study micro-data directly. Lovell (1986) summarizes his studies with Hirsch:

> For 30 percent of the sampled firms, the mean of anticipated sales, two-months horizon, differed from the mean of actual realizations at the 5 percent level of significance. However, the overestimates of the optimistic firms roughly canceled the underestimates of pessimistic firms so that for industry aggregates there is no bias; this offsetting of systematic error partially explains why the aggregates of anticipation data appear to be more accurate than the predictions of individual firms. [Lovell (1986, p. 115)]

Muth (1985) studied expectations and anticipations data from five Pittsburgh-based firms. He finds that "the standard deviation of the forecast of at least three firms is inconsistent with the rational expectations hypothesis: ... [these firms] have standard deviations greater than the standard deviation of the actual. Since the rational forecast specifies $A = F + e$, where $\mathrm{E}(Fe) = 0$, the variance of A must clearly exceed that of F" [Muth (1985, p.13)].[22] Muth's study raises the issue of costs and benefits of rationality directly; in particular he concludes his study as follows: "... that some of the most significant deviations from rationality occur with firms having a small forecast error ... this suggests that the operating benefits from improved forecasts of the type analyzed here are not worth the extra cost" [Muth (1985, p. 28)].

Much of the non-agricultural expectations survey data are based on categorical responses to surveys where, in particular, respondents are asked to respond as follows: increase (+), no change (=), or decrease (−). Early efforts using such data created aggregate balances, where the number or proportion of respondents reporting a "−" are subtracted from the number or proportion reporting a "+". Nerlove (1983) suggests that

---

[22] Here Muth's idea is as follows: A is the actual realization of the variable of interest, F is the firm's forecast of A based on information held at a previous period, and e is an error term.

such categorical data be analyzed as conditional log-linear probability models; however, such models do not recognize the ordering behind typical categorical responses. Accordingly, Nerlove (1988) suggests treating such data as categorical responses from continuous latent variable models, with categories defined as thresholds. These methods are applied, *inter alia,* by Horvath, Nerlove, and Wilson (1992), and Nerlove and Weeks (1992). Nerlove and Schuermann (1995a, 1995b) estimate such a model by simulation maximum likelihood methods.[23] For quarterly surveys of British manufacturing firms and for Swiss firms, they reject both rational expectations and adaptive expectations.

In studies related to agriculture, we also see rejections of rational expectations with micro data. Irwin and Thraen (1994) summarize ten studies in agriculture which test rationality of individual expectations. In seven of these they find rejections of bias or orthogonality conditions (conditions (a) and (b) given above) which are basic to rational expectations. The authors of the survey attempt to explain differences in results by whether survey respondents had "direct monetary incentives to accurately report their expectations." Unfortunately, the evidence is scanty, if it exists at all, that "direct monetary payoffs" were present in any of the assessment tasks described, and further, the linking of the reward or incentive to the actual survey response is at best unclear in any of the cases considered.

4.3.4. Summary of the evidence

The evidence from both surveys and experimental studies can be summarized as follows:

(1) Aggregates of individual expectations are more likely to pass rationality tests than are individual expectations; Zarnowitz (1983), Nelson and Bessler (1992), Williams (1987), Nerlove and Schuermann (1995a, 1995b).
(2) There is considerable heterogeneity in individual expectations; Schultz and Brownlee (1941–42), Nelson and Bessler (1992).
(3) Subjects are able to recognize difference in underlying stochastic processes and adapt their forecasts to accommodate these differences, but not necessarily in an optimal manner; Schultz and Brownlee (1941–42), Hey (1994).

That agents in experimental markets look as if they are trying to build rational components into their forecasts [Hey (1994), Swenson (1997)], but do not do so adequately to pass a rationality test accords with the qualitative findings from Heady and Kaldor's (1954, p. 39) survey: "The current price then was adjusted for the expected effects of important supply-and-demand forces". That agents are not able to pass more stringent tests of rationality was recognized long ago in the psychological literature. Starting with the work of Meehl (1954), psychologists have (almost always) found clinical judgments of numerical variables to be inferior to mechanical (statistical) predictions. That is, when both a clinical judgment and a statistical predication of a criterion variable are available, such as academic success or prisoner parole recidivism, the statistical prediction

---

[23] See [McFadden and Ruud (1994)].

is rarely inferior to the clinical judgment. The twenty cases studied in Meehl's seminal book generated a plethora of additional studies, all reaching similar conclusions – "an apparent superiority for mechanical models ..." [Sawyer (1966, p. 178)].

Our survey of experimental work finds that only for the most simple univariate process (the random walk studied in [Dwyer et al. (1993)] do we find clear evidence of rational expectations. Introduction of complexity, in terms of more complex univariate structures [Nelson and Bessler (1992), and Beckman and Downs (1997)] or market equilibria [Williams (1987)] results in clear rejections of rationality. Perhaps Simon's assessment of human behavior captures what the experimental results are telling us:

> Human behavior ... is not to be accounted for by a handful of invariants. It is certainly not to be accounted for by assuming perfect adaptation to the environment. Its basic mechanisms may be relatively simple, and I believe they are, but that simplicity operates in interaction with extremely complex boundary conditions imposed by the environment and by the very facts of human long-term memory and the capacity of human beings, individually and collectively, to learn. [Simon (1979, p. 510)]

Environmental conditions related to the costs and benefits of responding in a "rational manner" to laboratory questions ought to be a prime point of focus in future laboratory work.

5. Conclusions and directions for further research

We began our discourse on expectations and their role in dynamic optimizing behavior with a statement of the central simplifying assumption which runs through both theoretical and empirical work in this area and one which is adopted in the remainder of our essay. This is the assumption of separation of expectations and optimizing behavior, which goes back at least to the work of Keynes and Hicks in the 1930s. Such separation is a powerful simplification both theoretically and empirically, but we know that it is not theoretically correct. In a "theoretically correct" but essentially useless formulation, decisions and expectations are not separable; the explanation of behavior proceeds directly from assumptions about agents' priors and the dynamic constraints of their optimization problem to the decisions they take now and in the future in response to future events. We do not see any viable alternative over most of the range of problems in dynamic optimizing behavior under uncertainty studied by agricultural and general economists. Yet the state of the results of recent experimental studies of expectations, discussed further below, suggests the need to relax or modify this assumption and to provide a clearer framework of analysis for understanding the relation between how expectations are formed and reported and the uses to which such expectations are put and the rewards of optimizing behavior. The importance of incentives in experimental design suggests that experimental subjects may be better able to say what they will do than what they expect on the basis of the information presented to them.

Two models showing how the separation assumption works in practice were discussed in some detail. The first, oriented toward empirical application to historical time series data, was the old Nerlove supply model. The second, designed not for direct empirical application but rather for the derivation of comparative statics results, was a model of small ruminant production and supply. In this connection, we showed that, provided the separation assumption can be maintained, it is generally unnecessary to know anything about the mechanism by which expectations are formed in order to draw interesting and useful theoretical conclusions, and thus illustrated the power of the assumption. Were the authors of this chapter primarily economic theorists, we might, in view of the many difficulties discussed above in the main body of this chapter, conclude that further research ought to focus on questions of a purely theoretical nature. Unfortunately, most serious, real world, empirical questions do require a component of the model designed to deal with people's responses that includes some specification of the way in which expectations are formed and how they influence behavior.

Next, in Section 2, we turned to the five principal models of expectation formation used in analyses of aggregate time series data: extrapolative; adaptive; implicit; two variants of rational, fully rational and quasi-rational; and futures price based models. Extrapolative and adaptive expectations were used extensively in early studies of agricultural supply and related phenomena. Adaptive expectations models held up well in the sense that they generally yielded intuitively plausible conclusions with respect to the other parameters being estimated, but these models had the unfortunate tendency to confound expectation formation with other dynamic aspects of behavior and, moreover, in practice generally produce highly variable results for the same product supply in different periods or circumstances, which suggests, as argued in [Nerlove (1979)], that we are leaving out far too much in the nonexpectational part of our models. Implicit expectations, which were introduced prior to the formulation of the rational expectations hypothesis, share many of the latter's attractive features but suffer from a fatal flaw, corrected in Muth's 1961 formulation. Futures price based expectations for storable commodities are simply inconsistent with the basic paradigms of economics (utility maximizing agents and equilibrium),[24] and the evidence supporting them is mixed, all the more so for nonstorable commodities.[25] In any case, such models are of limited significance for aggregate time series studies since futures markets do not exist, or have not existed for considerable periods, for those commodities we would like to study. Rational expectations (RE) are the most theoretically attractive model of expectation formation. The model is, however, difficult to apply in practice, and, as shown in Section 3, generally fails to be supported empirically in those few attempts to test the REH. The difficulties of applying the rational expectations model in practice are corrected by

[24] Which doesn't, of course, mean that they're wrong, only that in accepting them we'd be forced to discard too much else which has proved useful and valid in the discipline.

[25] Gardner's (1976) results give weak support to the theory that the futures price is an *indicator* of expected future prices, but the evidence he presents is also consistent with misspecification of the underlying supply models.

the simplification of quasi-rational expectations (QRE). QRE are extremely easy to apply to time series data, are less subject to problems related to the specification of the underlying behavioral model, and are asymptotically equivalent to the RE under correct specification. For agricultural economists who continue to analyze aggregate time series data, adoption of RE as a maintained hypothesis and application in the form of QRE would allow a highly desirable concentration on the substance of the behavioral part of the model, and strikes us as the way to go. But, as a tool for research on expectation formation itself, we believe that such studies are a dead end.

The final section of this chapter, Section 3, considers the evidence, both direct and indirect, principally for rational expectations, since this hypothesis is now the leading, if not the sole, contender for our hearts and minds. Apart from the minor difficulty that all expectational models of distributed lags (EDL) fail what we call "minimalist tests", models based on QRE appear to work fairly well for aggregate time series data, in the sense that assuming them gives behavioral results consistent with theory. As we point out, however, the analyses so far undertaken are not really tests of rational expectations but rather of the dynamic optimization model in which they are imbedded. We conclude that further attempts to test RE in an aggregate time series context, while not exactly futile, are not worth the effort.

The unsatisfactory state of affairs with respect to conventional econometric analysis in this area has led to considerable recent research, building on the earlier work of the Iowa State group, which emphasizes direct observation of the expectations themselves, as reported by respondents to survey questionnaires or as predictions in an experimental context. The goal of this research is not only to test models of expectation formation freed from the constraints imposed by the behavioral model in aggregate time series analysis, but also to understand better the way in which expectations are actually formed and how they might influence subsequent behavior, and to refine models of expectation formation in the light of this evidence. It is in this connection, as our discussion of motivation – particularly of the payoff structure to participants in experiments – suggests, that the separation assumption begins to break down. When asked in a survey what they expect with respect to such-and-such, about what and how do respondents answer? On the whole we remain ignorant of respondents' state of mind, and really carefully designed surveys directed to elucidating these matters remain to be carried out. The problem is that most economic surveys are designed for purposes other than understanding people's behavior and, particularly, how they form expectations of the future and respond to those expectations.

Experimental studies are carried out on a far smaller scale than surveys. For this reason, experimental studies of how people predict, which may perhaps be assumed to be indicative of how expectations are formed, have recently been undertaken. Unfortunately, insufficient attention has been paid to the conditions set in the experiments related to the costs and benefits of responding in a "rational manner" to the laboratory questions. Following Simon's dictum on the complexity of "boundary conditions" imposed by the environment, complexity has been added in a haphazard manner. Subjects have not been able to respond in ways consistent with the experimenters' theories.

Perhaps this is because the subjects have been brought into the lab without giving consideration to external validity: Do the experimental results have anything to say about real world agents? Experimentalists have focused instead on internal validity: Were the results valid within the scope of this particular experiment? Little or no motivation, with flat payoff functions (in the neighborhood of the rational expectation response), may have given the impression that laboratory subjects were either bored, irrational, or both. But the real world provides large incentives as payoffs, and those who fail to respond in an acceptable manner are dropped from the experiment – the market does not allow subjects who consistently forecast poorly to stay around very long, especially those who do not begin with large initial endowments.

Our view is that motivation in experiments, particularly related to how subjects predict the future and how they behave in the context of such predictions, needs to be taken more seriously if we're going to make the leap from nice, simple, internally valid results to useful, externally valid results. In this context, the separation assumption, which has been central to virtually all theoretical thinking and empirical study of dynamic optimizing behavior, may need to be discarded or, at the very least, relaxed.

Whereas the problems of internal validity are solvable within the limits of the logic of probability, the problems of external validity are not logically solvable in any neat, conclusive way. Generalization always turns out to involve extrapolation into a realm not represented in one's sample. Campbell and Stanley (1963, p. 17–18) write:

> ... there is a general empirical law we are assuming, along with all scientists. This is the modern version of Mill's assumption as to the lawfulness of nature. In its modern, weak version, this can be stated as the assumption of the "stickiness" of nature: we assume that the closer two events are in time, space and measured value on any or all dimensions, the more they tend to follow the same laws. While complex interactions and curvilinear relationships are expected to confuse attempts at generalization, they are more to be expected the more the experimental situation differs from the setting to which one wants to generalize. Our call for greater external validity will thus be a call for that maximum similarity of experiments to the conditions of application which is compatible with internal validity.

In assessing external validity, we will have to come to terms with the incentive structure built into our experiments. Adding complexity [Nelson and Bessler (1992), Williams (1987), Beckman and Downs (1997), Swenson (1997), Hey (1994)] without a matching incentive structure is almost asking for chaotic results. The direction which additional complexity should take in the laboratory should be dictated by the types of behavioral questions asked in other contexts. We ought to add more complexity in studying behavior of relevance to questions related to the formulation of policy. Keep the experiment simple on all other counts and, if we are serious about testing rationality, make the slope of the payoff function match the real world in the neighborhood of the rational expectations response. But the real issue is not what model to use, but rather how we might best proceed to get answers to the substantive questions with which we are concerned.

Acknowledgements

Nerlove's contribution was supported by the Maryland Agricultural Experiment Station. Bessler's contribution was supported by the Texas Agricultural Experiment Station. We are indebted for helpful suggestions to the editors of this volume, Bruce L. Gardner and Gordon C. Rausser, to an anonymous referee, and to Eliseu Alves, Ray Battalio, Zvi Eckstein, David M. Grether, Wallace E. Huffman, Richard E. Just, Anke Meyer, Robert G. Nelson, and Sarah Teichner, who commented on portions of earlier drafts.

References

Amman, H.M., D. Kendrick and S. Achath (1995), "Solving stochastic optimization models with learning and rational expectations", Economics Letters 48:9–13.

Antonovitz, F., and R. Green (1990), "Alternative estimates of fed beef supply response to risk", American Journal of Agricultural Economics 72:475–487.

Askari, H., and J.T. Cummings (1976), Agricultural Supply Response: A Survey of the Econometric Evidence (Praeger Publishers, New York).

Askari, H., and J.T. Cummings (1977), "Estimating agricultural supply response with the Nerlove model: A survey", International Economic Review 18:257–292.

Barro, R.J. (1974), "Are government bonds net wealth?", Journal of Political Economy 82:1095–1117.

Barro, R.J., and G.S. Becker (1989), "Fertility choice in a model of economic growth", Econometrica 57:481–501.

Beckman, S.R., and D. Downs (1997), "Forecasters as imperfect information processors: Experimental and survey evidence", Journal of Economic Behavior and Organization 32:89–101.

Bessler, D.A. (1980), "Aggregated personalistic beliefs on yields of selected crops estimated using ARIMA processes", American Journal of Agricultural Economics 62:666–674.

Bessler, D.A. (1982), "Adaptive expectations, the exponentially weighted forecast, and optimal statistical predictors: A revisit", Agricultural Economics Research 34:16–23.

Bessler, D.A., and J.A. Brandt (1992), "An analysis of forecasts of livestock prices", Journal of Economic Behavior and Organization 18:249–263.

Bessler, D.A., and C.V. Moore (1979), "Use of probability assessments and scoring rules for agricultural forecasts", Agricultural Economics Research 31:44–47.

Brier, G.W. (1950), "Verification of forecasts expressed in terms of probability", Monthly Weather Review 78:1–3.

Campbell, D., and J.C. Stanley (1963), Experimental and Quasi-Experimental Designs for Research (Houghton Mifflin Company).

Carlson, J. (1977), "A study of price forecasts", Annals of Economic and Social Measurement 6:27–56.

Carvalho, J.L. (1972), Production, investment and expectations: A study of the United States cattle industry, Ph.D. dissertation (Department of Economics, University of Chicago).

Chavas, J.-P., R. Pope and R. Kao (1983), "An analysis of the role of futures prices, cash prices and government programs in acreage response", Western Journal of Agricultural Economics 8:27–33.

Colling, P.L., S.H. Irwin and C.R. Zulauf (1992), "Weak-and strong-form rationality tests of market analysts' expectations of USDA hogs and pigs reports", Review of Agricultural Economics 14:263–270.

Covey, T., and D.A. Bessler (1995), "Asset storability and the information content of intertemporal prices", Journal of Empirical Finance 2:103–115.

Davis, D.D., and C.A. Holt (1993), Experimental Economics (Princeton University Press, Princeton, NJ).

De Finetti, B. (1965), "Methods for discriminating levels of partial knowledge concerning a test item", The British Journal of Mathematical and Statistical Psychology 18:87–123.

Diebold, F.X., and G.D. Rudebusch (1989), "Scoring the leading indicators", Journal of Business 62:369–391.

Dwyer, G.P., A.W. Williams, R.C. Battalio and T.I. Mason (1993), "Tests of rational expectations in a stark setting", The Economic Journal 103:586–601.

Eales, J., B. Engle, R. Hauser and S. Thompson (1990), "Grain price expectations of Illinois farmers and grain merchandisers", American Journal of Agricultural Economics 72:701–708.

Eckstein, Z. (1984), "A rational expectations model of agricultural supply", Journal of Political Economy 92:1–19.

Eckstein, Z. (1985), "The dynamics of agricultural supply: A reconsideration", American Journal of Agricultural Economics 67:204–214.

Fair, R.C., and J.B. Taylor (1983), "Solution and maximum-likelihood estimation of dynamic nonlinear rational expectations models", Econometrica 50:1169–1185.

Fair, R.C., and J.B. Taylor (1990), "Full information estimation and stochastic simulation of models with rational expectations", Journal of Applied Econometrics 5:381–392.

Feder, G., R. Just and A. Schmitz (1980), "Futures markets and the theory of the firm under price uncertainty", Quarterly Journal of Economics 96:317–328.

Fornari, I. (1994), U.S. beef cattle supply: Tests of quasi-rational expectations against fully rational expectations, Unpublished Ph.D. Dissertation (University of Pennsylvania).

Foster, K.A. (1990), A Dynamic model of cattle inventories and supply in the U.S. beef cattle industry, Ph.D. Dissertation (University of California, Davis).

Friedman, M. (1953), "The methodology of positive economics", in: Essays in Positive Economics (The University of Chicago Press, Chicago) 3–43.

Friedman, M. (1957), A Theory of the Consumption Function (Princeton University Press, Princeton).

Frydman, R., and E.S. Phelps (1983), Individual Forecasting and Aggregate Outcomes (Cambridge University Press, New York).

Garcia, P., and R.M. Leuthold (1992), "The effect of market information on corn and soybean markets", in: Proceedings of the 1992 NCR-134 Applied Commodity Price Analysis, Forecasting and Market Risk Management Conference, Chicago, IL, April, pp. 59–74.

Gardner, B.L. (1976), "Futures prices in supply analysis", American Journal of Agricultural Economics 58:81–84.

Goodwin, R.M. (1947), "Dynamical coupling with especial reference to markets having production lags", Econometrica 15:181–204.

Goodwin, T.H., and S.M. Sheffrin (1982), "Testing the rational expectations hypothesis in an agricultural market", Review of Economics and Statistics 64:658–667.

Grandmont, J.-M. (1998), "Expectations formation and stability of large socioeconomic systems", Econometrica 66:741–781.

Granger, C.W.J. (1989), "Combining forecasts – twenty years later", Journal of Forecasting 8:167–173.

Hamilton, J.D. (1994), "Cointegration", in: Time Series Analysis, Ch. 19 (Princeton University Press, Princeton).

Hansen, L.P., and T.J. Sargent (1980), "Formulating and estimating dynamic linear rational expectations models", Journal of Economic Dynamics and Control 2:7–46.

Hansen, L.P., and T.J. Sargent (1981), "Linear rational expectations models for dynamically interrelated variables", in: R.E. Lucas and T.J. Sargent, eds., Rational Expectations and Econometric Practice (University of Minnesota Press, Minneapolis) 127–156.

Hansen, L.P., and T.J. Sargent (1991), "Exact linear rational expectations models: Specification and estimation", in: L.P. Hansen and T.J. Sargent, eds., Rational Expectations Econometrics (Westview Press, Boulder, CO) 45–75.

Harrison, G. (1989), "Theory and misbehavior of first-price auctions: A reply", American Economic Review 79:749–762.

Harrison, G. (1992), "Theory and misbehavior of first-price auctions: Reply", American Economic Review 82:1426–1443.

Heady, E.O., and D.R. Kaldor (1954), "Expectations and errors in forecasting agricultural prices", Journal of Political Economy 62:34–47.

Hey, J. (1994), "Expectations formation: Rational or adaptive or ...?" Journal of Economic Behavior and Organization 25:329–349.

Hicks, J.R. (1946), Value and Capital, 2nd edn. (Clarendon Press, Oxford).

Hicks, J.R. (1977), Economic Perspectives Further Essays on Money and Growth (Clarendon Press, Oxford).

Hildenbrand, W. (1994), Market Demand: Theory and Empirical Evidence (Princeton University Press, Princeton).

Hirshleifer, J., and J.G. Riley (1992), The Analytics of Uncertainty and Information (University Press, Cambridge).

Holt, C.A. (1986), "Scoring-rule procedures for eliciting subjective probabilities and utility functions", in: P. Goel and A. Zellner, eds., Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti (Elsevier Science, New York) 279–290.

Holthausen, D. (1979), "Hedging and the competitive firm under price uncertainty", American Economic Review 69:989–995.

Horvath, B., and M. Nerlove (1996), "Some econometric implications of learning", in: E. Helmstädter, G. Poser and H.J. Ramser, eds., Beiträge zur Angewandten Wirtschaftsforschung (Duncker u. Humblot, Berlin) 319–337.

Horvath, B., M. Nerlove and D. Willson (1992), "A reinterpretation of direct tests of forecast rationality using business survey data", Mimeo (University of Pennsylvania).

Irwin, S.H., and C.S. Thraen (1994), "Rational expectations in agriculture? A review of the issues and the evidence", Review of Agricultural Economics 16:133–158.

Jacobs, R.L., and R.A. Jones (1980), "Price expectations in the United States: 1947–1975", American Economic Review 70:269–276.

Jarvis, L.S. (1986), "Supply response in the cattle industry: The Argentine case", Giannini Foundation Special Report (University of California).

Johnson, D.G. (1947), Forward Prices for Agriculture (The University of Chicago Press, Chicago).

Just, R.E. (1974), "Econometric analysis of production decisions with government intervention: The case of the California field crops", Giannini Foundation Monograph 33.

Just, R., and G. Rausser (1981), "Commodity price forecasting with large-scale econometric models and the futures market", American Journal of Agricultural Economics 63:197–209.

Kapteyn, A., N. Kiefer and J. Rust, eds. (1995), "The microeconometrics of dynamic decision making", Special issue, Journal of Applied Econometrics 10 (supplement).

Keane, M.P., and D.E. Runkle (1990), "Testing the rationality of price forecasts: New evidence from panel data", American Economic Review 80:714–735.

Keynes, J.M. (1936), The General Theory of Employment, Interest and Money (Harcourt, Brace & Co, New York).

Kiefer, N.M. (1988–89), "Optimal collection of information by partially informed agents", Econometric Reviews 7:133–148.

Kirman, A.P. (1992), "What or whom does the representative agent represent?", Journal of Economic Perspectives 6:117–136.

Kling, J., and D.A. Bessler (1989), "Calibration-based predictive distributions: An application of prequential analysis to interest rates, money, prices and output", Journal of Business 62:477–499.

Leuthold, R. (1974), "The price performance on the futures market of a nonstorable commodity", American Journal of Agricultural Economics 56:271–279.

Lovell, M.C. (1986), "Tests of the rational expectations hypothesis", American Economic Review 76:110–124.

Malinvaud, E. (1969), "First order certainty equivalence", Econometrica 37:706–718.

Martin, L., and P. Garcia (1981), "The price forecasting performance of futures markets for live cattle and hogs: A disaggregated analysis", American Journal of Agricultural Economics 63:209–215.

McClennen, E.F. (1990), Rationality and Dynamic Choice: Foundational Explorations (University Press, Cambridge).

McFadden, D., and P.A. Ruud (1994), "Estimation by simulation", Review of Economics and Statistics 76:591–608.

Meehl, P. (1954), Clinical versus Statistical Prediction: A Theoretical Analysis and Review of the Evidence (University of Minnesota Press, Minneapolis).

Mills, E.S. (1955), The theory of inventory decisions, Ph.D. Dissertation (University of Birmingham, England).

Mills, E.S. (1962), Price, Output, and Inventory Policy: A Study in the Economics of the Firm and Industry (John Wiley, New York).

Miranda, M.J., and G.D. Schnitkey (1995), "An empirical model of asset replacement in dairy production", in: Kapteyn et al., eds., op. cit.

Mundlak, Y. (1966), "On the microeconomic theory of distributed lags", Review of Economics and Statistics 48:51–60.

Mundlak, Y. (1967), "Long-run coefficients and distributed lag analysis: A reformulation", Econometrica 35:278–293.

Murphy, A., and R. Winkler (1970), "Scoring rules in probability assessment and evaluation", Acta Psychologica 34:273–286.

Muth, J.F. (1960), "Optimal properties of exponentially weighted forecasts", Journal of the American Statistical Association 55:299–305.

Muth, J.F. (1961), "Rational expectations and the theory of price movements", Econometrica 29:315–335.

Muth, J.F. (1985), "Short-run forecasts of business activity", Paper presented at the joint Pittsburgh Meeting of the Eastern Economic Association-International Society for Inventory Research, March.

Nelson, R.G. (1987), Probability elicitation and the formation of expectations: An experimental approach, Ph.D. Thesis (Texas A&M University).

Nelson, R.G., and D.A. Bessler (1989), "Subjective probabilities elicited under proper and improper scoring rules: A laboratory test of predicted responses", American Journal of Agricultural Economics 71:363–369.

Nelson, R.G., and D.A. Bessler (1992), "Quasi-rational expectations: Experimental evidence", Journal of Forecasting 11(2):141–156.

Nerlove, M. (1955), "The predictive test as a tool for inference: The demand for meat in the United States", Unpublished M.A. Dissertation (The Johns Hopkins University, Baltimore, MD).

Nerlove, M. (1956a), "Estimates of elasticities of supply of selected agricultural commodities", Journal of Farm Economics 38:496–509.

Nerlove, M. (1956b), "Estimates of the elasticities of supply of corn, cotton, and wheat", Ph.D. Dissertation (The Johns Hopkins University, Baltimore, MD).

Nerlove, M. (1958a), "The implications of Friedman's permanent income hypothesis for demand analysis", Agricultural Economics Research 10:1–14.

Nerlove, M. (1958b), Distributed Lags and Demand Analysis, Agricultural Handbook No. 141 (Government Printing Office, Washington, D.C.).

Nerlove, M. (1958c), The Dynamics of Supply: Estimation of Farmers' Response to Price (Johns Hopkins University Press, Baltimore, MD).

Nerlove, M. (1961), "Time-series analysis of the supply of agricultural products", in: E.O. Heady et al., eds., Agricultural Supply Functions: Estimating Techniques and Interpretation (Iowa State University Press, Ames, IA).

Nerlove, M. (1967), "Distributed lags and unobserved components in economic time series", in: W. Fellner et al., eds., Ten Economic Essays in the Tradition of Irving Fisher (John Wiley & Sons, New York).

Nerlove, M. (1972), "Lags in economic behavior", Econometrica 40:221–251.

Nerlove, M. (1979), "The dynamics of supply: Retrospect and prospect", American Journal of Agricultural Economics 61:874–888.

Nerlove, M. (1983), "Expectations, plans and realizations in theory and practice", Econometrica 51:1251–1279.

Nerlove, M. (1988), "Analysis of business-test survey data by means of latent variable models", in: W. Franz, W. Gaab and J. Walters, eds., Theoretische und angewandte Wirtschaftsforschung, Heinz Konig zum 60 (Springer Verlag, Heidelberg).

Nerlove, M., and I. Fornari (1997), "Quasi-rational expectations, an alternative to fully rational expectations: An application to U.S. beef cattle supply", Journal of Econometrics 83:129–161.

Nerlove, M., D.M. Grether and J.L. Carvalho (1979), Analysis of Economic Time Series: A Synthesis (Academic Press, New York) Reprinted with corrections (1995).

Nerlove, M., and L. Raut (1997), "Growth models with endogenous population", in: M.R. Rosenzweig and O. Stark, eds., The Handbook of Population and Family Economics (Elsevier, New York).

Nerlove, M., and T. Schuermann (1995a), "Expectations: Are they rational, adaptive, or naive? An essay in simulation-based inference", in: G.S. Maddala, P.C.B. Phillips and T.N. Srinivasan, eds., Advances in Econometrics and Quantitative Economics (Basil Blackwell, Oxford).

Nerlove, M., and T. Schuermann (1995b), "Businessmen's expectations are neither rational nor adaptive", Unpublished working paper (Department of Agricultural and Resource Economics, University of Maryland).

Nerlove, M., and T.D. Soedjana (1996), "Slametans and sheep: Savings and small ruminants in semi-subsistence agriculture in Indonesia", Unpublished working paper (Department of Agricultural and Resource Economics, University of Maryland).

Nerlove, M., and M. Weeks (1992), "The construction of multivariate probability simulators with an application to the multinomial probit model", Mimeo (University of Pennsylvania).

Nerlove, M., D. Willson and B. Horvath (1992), "A reinterpretation of direct tests of forecast rationality using business survey data", in: K.H. Oppenländer and G. Poser, eds., Business Cycle Analysis by Means of Economic Surveys, Part 1 (Aldershot, Avebury).

Orazem, P., and J. Miranowski (1986), "An indirect test for the specification of expectation regimes", Review of Economics and Statistics 68:603–609.

Pesaran, M.H. (1987), The Limits to Rational Expectations (Basil Blackwell, Oxford).

Plott, C., and S. Sunder (1988), "Rational expectations and the aggregation of diverse information in laboratory security markets", Econometrica 56:1085–1118.

Pierce, D.A. (1975), "Forecasting in dynamic models with stochastic regressors", Journal of Econometrics 3:349–374.

Ravallion, M. (1985), "The informational efficiency of traders' price expectations in a Bangladesh rice market, Oxford Bulletin of Economics and Statistics 47:171–184.

Ravallion, M. (1987), Markets and Famines (Oxford University Press, Oxford).

Rosen, S. (1986), "Dynamic animal economics", American Journal of Agricultural Economics 68:547–557.

Runkle, D.E. (1991), "Are farrowing intentions rational forecasts?", American Journal of Agricultural Economics 73:594–600.

Sargent, T.J. (1973), "Rational expectations, the real rate of interest, and the natural rate of unemployment", in: Brookings Papers on Economic Activity 429–492.

Sargent, T.J. (1976a), "A classical macroeconometric model for the United States", Journal of Political Economy 84:207–237.

Sargent, T.J. (1976b), "The observational equivalence of natural and unnatural rate theories of macroeconomics", Journal of Political Economy 84:631–639.

Sargent, T.J. (1981), "Interpreting economic time series", Journal of Political Economy 89:213–248.

Sargent, T.J., and C.A. Sims (1977), "Business cycle modeling without pretending to have too much a priori information", in: C.A. Sims, ed., New Methods in Business Cycle Research (Federal Reserve Bank, Minneapolis).

Sargent, T.J., and N. Wallace (1973), "Rational expectations and the dynamics of hyperinflation", International Economic Review 14:328–350.

Savage, L. (1954), The Foundations of Statistics (Wiley, New York).

Sawyer, J. (1966), "Measurement and prediction, clinical and statistical", Psychological Bulletin 66:178–200.

Schmalensee, R. (1976), "An experimental study of expectations formation", Econometrica 44:17–41.

Schultz, T.W., and O.H. Brownlee (1941–42), "Two trials to determine expectation models applicable to agriculture", Quarterly Journal of Economics, 55:487–496.

Shackle, G.L.S. (1952), Expectation in Economics, 2nd ed. (University Press, Cambridge).

Shuford, E., A. Albert and H.E. Massengill (1966), "Admissible probability measurement procedures", Psychometrika 31:125–145.

Siegel, S., and D.A. Goldstein (1959), "Decision-making behavior in a two-choice uncertain outcome situation", Journal of Experimental Psychology 57:37–42.

Simon, H.A. (1979), "Rational decision making in business organizations", American Economic Review 69:493–513.

Simon, H. (1997), Models of Bounded Rationality, Vol. III: Empirically Grounded Economic Reason Part 4: Behavioral Economics and Bounded Rationality (The MIT Press, Cambridge, MA) 267–443.

Smith, V., G. Suchanek and A. Williams (1988), "Bubbles, crashes, and endogenous expectations in experimental spot asset markets", Econometrica 56:1119–1151.

Stael von Holstein, C.-A. (1970), Assessment and Evaluation of Subjective Probability Distributions (The Economic Research Institute at the Stockholm School of Economics, Stockholm).

Stock, J.H., and M.W. Watson (1988), "Testing for common trends", Journal of the American Statistical Association, 83:1097–1107.

Stokey, N.L., and R.E. Lucas (1989), Recursive Methods in Economic Dynamics (Harvard University Press, Cambridge).

Swenson, C. (1997), "Rational expectations and tax policy: Experimental market evidence", Journal of Economic Behavior and Organization, 32:433–455.

Theil, H. (1957), "A note on certainty equivalence in dynamic planning", Econometrica 25:346–349.

Treadway, A.B. (1967), "Rational entrepreneurial behavior and the dynamics of investment", Ph.D. Dissertation (The University of Chicago).

Tomek, W.G., and R.W. Gray (1970), "Temporal relationships among prices on commodity futures markets: Their allocative and stabilizing roles", American Journal of Agricultural Economics 52:372–380.

Tronstad, R., and T. McNeil (1989), "Asymmetric price risk: An econometric analysis of aggregate sow farrowings, 1973–86", American Journal of Agricultural Economics 71:630–637.

Turnovsky, S.J. (1970), "Empirical evidence on the formation of price expectations", Journal of the American Association 65:1441–1454.

Wallace, H.A. (1923), "What's on the corn judge's mind?", Journal of the American Society of Agronomy 15:300–304.

Wallis, K.F. (1980), "Econometric implications of the rational expectations hypothesis", Econometrica 48:49–73.

Williams, A. (1987), "The formation of price expectations in experimental markets", Journal of Money, Credit, and Banking 19:1–18.

Williams, J. (1986), The Economic Function of Futures Markets (Cambridge University Press).

Winkler, R. (1983), "Judgements under uncertainty", in: S. Kotz and N. Johnson, eds., Encyclopedia of Statistical Sciences, Vol. 4 (John Wiley and Sons, New York).

Working, H. (1942), "Quotations on commodity futures as price forecasts", Econometrica 10:39–52.

Zarnowitz, V. (1983), "The accuracy of individual and group forecasts from business outlook surveys", National Bureau of Economic Research Working Paper Number 1053.

Zellner, A., A. Hong and C. Min (1991), "Forecasting turning point in international output growth rates using Bayesian exponentially weighted autoregression, time-varying parameter and pooling techniques", Journal of Econometrics 49:275–304.

Chapter 4

THE AGRICULTURAL INNOVATION PROCESS: RESEARCH AND TECHNOLOGY ADOPTION IN A CHANGING AGRICULTURAL SECTOR

DAVID SUNDING and DAVID ZILBERMAN

Department of Agricultural and Resource Economics, University of California at Berkeley, Berkeley, CA

Contents

Handbook of Agricultural Economics, Volume 1, Edited by B. Gardner and G. Rausser

Abstract

The chapter reviews the generation and adoption of new technologies in the agricultural sector. The first section describes models of induced innovation and experimentation, considers the political economy of public investments in agricultural research, and addresses institutions and public policies for managing innovation activity. The second section reviews the economics of technology adoption in agriculture. Threshold models, diffusion models, and the influence of risk, uncertainty, and dynamic factors on adoption are considered. The section also describes the influence of institutions and government interventions on adoption. The third section outlines future research and policy challenges.

Keywords

innovation, diffusion, adoption, technology transfer, intellectual property

JEL classification: Q16

Technological change has been a major factor shaping agriculture in the last 100 years [Schultz (1964), Cochrane (1979)]. A comparison of agricultural production patterns in the United States at the beginning (1920) and end of the century (1995) shows that harvested cropland has declined (from 350 to 320 million acres), the share of the agricultural labor force has decreased substantially (from 26 to 2.6 percent), and the number of people now employed in agriculture has declined (9.5 million in 1920 vs. 3.3 million in 1995); yet agricultural production in 1995 was 3.3 times greater than in 1920 [United States Bureau of the Census (1975, 1980, 1998)]. Internationally, tremendous changes in production patterns have occurred. While world population more than doubled between 1950 and 1998 (from 2.6 to 5.9 billion), grain production per person has increased by about 12 percent, and harvested acreage per person has declined by half [Brown et al. (1999)]. These figures suggest that productivity has increased and agricultural production methods have changed significantly.

There is a large amount of literature investigating changes in productivity,[1] which will not be addressed here. Instead this chapter presents an overview of agricultural economic research on innovations – the basic elements of technological and institutional change. Innovations are defined here as new methods, customs, or devices used to perform new tasks.

The literature on innovation is diverse and has developed its own vocabulary. We will distinguish between two major research lines: research on innovation generation and research on the adoption and use of innovation. Several categories of innovations have been introduced to differentiate policies or modeling. For example, the distinction between innovations that are *embodied* in capital goods or products (such as tractors, fertilizers, and seeds) and those that are *disembodied* (e.g., integrated pest management schemes) is useful for directing public investment in innovation generation. Private parties are less likely to invest in generating disembodied innovations because of the difficulty in selling the final product, so that is an area for public action. Private investment in the generation of embodied innovations requires appropriate institutions for intellectual property rights protection, as we will see below.

The classification of innovations according to form is useful for considering policy questions and understanding the forces behind the generation and adoption of innovations. Categories in this classification include mechanical innovations (tractors and combines), biological innovations (new seed varieties), chemical innovations (fertilizers and pesticides), agronomic innovations (new management practices), biotechnological innovations, and informational innovations that rely mainly on computer technologies. Each of these categories may raise different policy questions. For example, mechanical innovations may negatively affect labor and lead to farm consolidation. Chemical and biotechnological innovations are associated with problems of public acceptance and environmental concerns. We will argue later that economic forces as well as the state of scientific knowledge affect the form of innovations that are generated and adopted in various locations.

---

[1] See Mundlak (1997), Ball et al. (1997), and Antle and McGuckin (1993).

Another categorization of innovation according to form distinguishes between process innovations (e.g., a way to modify a gene in a plant) and product innovations (e.g., a new seed variety). The ownership of rights to a process that is crucial in developing an important product may be a source of significant economic power. We will see how intellectual property rights and regulations affect the evolution of innovation and the distribution of benefits derived from them.

Innovations can also be distinguished by their impacts on economic agents and markets which affect their modeling; these categories include yield-increasing, cost-reducing, quality-enhancing, risk-reducing, environmental-protection increasing, and shelf-life enhancing. Most innovations fall into several of these categories. For example, a new pesticide may increase yield, reduce economic risk, and reduce environmental protection. The analysis of adoption or the impact of risk-reducing innovations may require the incorporation of a risk-aversion consideration in the modeling framework, while investigating the economics of a shelf-life enhancing innovation may require a modeling framework that emphasizes inter-seasonal dynamics.

Three sections on the generation of innovations follow in Section 1. The first introduces results of induced innovation models and the role of economic forces in triggering innovations; the second presents a political-economic framework for government financing of innovations; and the third addresses various institutions and policies for managing innovation activities. Section 2 discusses the adoption of innovations and includes four sections. The first section considers threshold models and models of diffusion as a process of imitation; the second presents adoption under uncertainty; the third addresses dynamic considerations on adoption; and the last two sections deal with the impact of institutional and policy constraints on adoption. Section 3 addresses future directions.

1. Generation of innovation

1.1. Induced innovations

There are several stages in the generation of innovations. These stages are depicted in Figure 1. The first stage is discovery, characterized by the emergence of a concept or results that establish the innovation. A second essential stage is development, where the discovery moves from the laboratory to the field, and is scaled up, commercialized, and integrated with other elements of the production process. In cases of patentable innovations, between the time of discovery and development there may also be a stage where there is registration for a patent. If the innovation is embodied, once it is developed it has to be produced and, finally, marketed. For embodied innovations, the marketing stage consists of education, demonstration, and sales. Only then does adoption occur.

Some may hold the notion that new discoveries are the result of inspiration occurring randomly without a strong link to physical reality. While that may sometimes be the case, Hayami and Ruttan (1985) formalized and empirically verified their theory of

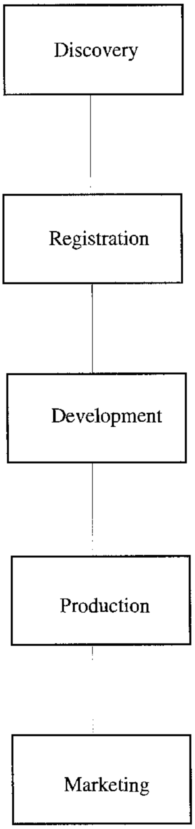

Figure 1.

induced innovations that closely linked the emergence of innovations with economic conditions. They argued that the search for new innovations is an economic activity that is significantly affected by economic conditions. New innovations are more likely to emerge in response to scarcity and economic opportunities. For example, labor shortages will induce labor-saving technologies. Environment-friendly techniques are likely to be linked to the imposition of strict environmental regulation. Drip irrigation and other water-saving technologies are often developed in locations where water constraints are binding, such as Israel and the California desert. Similarly, food shortages or high prices of agricultural commodities will likely lead to the introduction of a new high-yield variety, and perceived changes in consumer preferences may provide the background for new innovations that modify product quality.

The work of Boserup (1965) and Binswanger and McIntire (1987) on the evolution of agricultural systems supports the induced-innovation hypothesis. Early human groups, consisting of a relatively small number of members who could roam large areas of land, were hunters and gatherers. An increase in population led to the evolution of agricultural systems. In tropical regions where population density was still relatively small, farmers relied on slash-and-burn systems. The transition to more intensive farming systems that used crop rotation and fertilization occurred as population density increased even further. The need to overcome diseases and to improve yields led to the development of innovations in pest control and breeding, and the evolution of the agricultural systems we are familiar with. The work of Berck and Perloff (1985) suggests that the same phenomena may occur with seafood. An increased demand for fish and expanded harvesting may lead to the depletion of population and a rise in harvesting costs, and thus trigger economic incentives to develop alternative aquaculture and mariculture for the provision of seafood.

While scarcity and economic opportunities represent potential demand that is, in most cases, necessary for the emergence of new innovations, a potential demand is not sufficient for inducing innovations. In addition to demand, the emergence of new innovations requires technical feasibility and new scientific knowledge that will provide the technical base for the new technology. Thus, in many cases, breakthrough knowledge gives rise to new technologies. Finally, the potential demand and the appropriate knowledge base are integrated with the right institutional setup, and together they provide the background for innovation activities. These ideas can be demonstrated by an overview of some of the major waves of innovations that have affected U.S. agriculture in the last 150 years.

New innovations currently are linked with discoveries of scientists in universities or firms. However, in the past, practitioners were responsible for most breakthroughs. Over the years, the role of research labs in producing new innovations has drastically increased, but field experience is still very important in inspiring innovations. John Deere, who invented the steel plow, was a farmer. This innovation was one of a series of mechanical innovations that were of crucial importance to the westward expansion of U.S. agriculture in the nineteenth century. At the time, the United States had vast tracts of land and a scarcity of people; this situation induced a wide variety of labor-saving innovations such as the thresher, several types of mechanical harvesters, and later the tractor.

Olmstead and Rhode (1993) argue that demand considerations represented by the induced-innovation hypothesis do not provide the sole explanation for the introduction of new technologies. They conclude that during the nineteenth century, when farm machinery (e.g., the reaper) was introduced in the United States, land prices increased relative to labor prices, which seems to contradict the induced-innovation hypothesis. As settlement of the West continued and land became more scarce, land prices may have risen relative to labor, but the cost of labor in America relative to other regions was high, and that provided the demand for mechanical innovations. Olmstead and Rhode (1993) argue that other factors also affected the emergence of these innovations, including the expansion of scientific knowledge in metallurgy and mechanics (e.g., the Bessemer process for the production of steel, and the invention of various types of mechanical engines), the establishment of the input manufacturing industry, and the interactive relationship between farmers and machinery producers.

The infrastructure that was established for the refinement, development, and marketing of the John Deere plow was later used for a generation of other innovations, and the John Deere Company became the world's leading manufacturer of agricultural mechanical equipment. It was able to establish its own research and development (R&D) infrastructure for new mechanical innovations, had enough financial leverage to buy the rights to develop other discoveries, and subsequently took over smaller companies that produced mechanical equipment that complemented its own. This pattern of evolution, where an organization is established to generate fundamental innovations of a certain kind, and then later expands to become a leading industrial manufacturer, is repeated in other situations in and out of agriculture.

It seems that during the settlement period of the nineteenth century, most of the emphasis was on mechanical innovation. Cochrane (1979) noted that yield per acre did not change much during the nineteenth century, but the production of U.S. agriculture expanded drastically as the land base expanded. However, Olmstead and Rhode (1993) suggest that even during that period there was heavy emphasis on biological innovation. Throughout the settlement period, farmers and scientists, who were part of research organizations such as the Agricultural Research Service (ARS) of the United States Department of Agriculture (USDA), and the experiment stations at the land-grant universities in the United States, experimented with new breeds, both domestic and imported, and developed new varieties that were compatible with the agro-climatic conditions of the newly settled regions. These efforts maintained per-acre yields.

Once most of the arable agricultural land of the continental United States was settled, expansion of agricultural production was feasible mostly through increases in yields per acre. The recognition of this reality and the basic breakthroughs in genetics research in the nineteenth century increased support for research institutions in their efforts to generate yield-increasing innovations. Most of the developed countries established agricultural research institutions. After World War II, a network of international research centers was established to provide agricultural innovations for developing countries. The establishment of these institutions reflected the recognition that innovations are products of R&D activities, and that the magnitude of these activities is affected by economic incentives.

Economic models have been constructed to explain patterns of investment in R&D activities and the properties of the emerging innovations. Evenson and Kislev (1976) developed a production function of research outcomes particularly appropriate for crop and animal breeding. In breeding activities, researchers experiment with a large number of varieties to find the one with the highest yield. The outcome of research efforts depends on a number of plots. In their model, the yield per acre of a crop is a random variable that can assume numerous values. Each experiment is a sampling of a value of this random variable and, if experiments are conducted, the experiment with the highest value will be chosen. Let Y_n be yield per acre of the nth experiment and n assumes value from 1 to N. The outcome of n experiments is $Y_N^* = \max\{Y_1, \dots, Y_N\}$. Y_N^* is the maximum value of the n experiment. Each Y_n can assume the value in the range of $(0, Y_X)$

with probability density $g(Y_n)$ so that $Y_{\max} g(Y_n)\,\mathrm{d}Y_n = 1$. The outcome of research on N plots Y_N^* is a random variable with the expected value $\mu(N) = \mathrm{E}\{\max_{n=1,N} Y_n\}$.

Evenson and Kislev (1976) showed that the expected value of Y_N^* increases with the number of the experiment, i.e., $\mu_N = \partial \mathrm{E}Y^*/\partial N > 0$, $\mu_{NN} = \partial^2 \mathrm{E}Y^*/\partial N^2 < 0$. As in Evenson and Kislev, consider the determination of optimal research levels when a policymaker's objective is to maximize net expected gain from research. Assume that the research improves the productivity of growers in a price-taking industry with output price P and acreage L. The new innovation is adopted fully and does not require extra research cost. The optimal research program is determined by solving

$$\max_N PL\big(U(N)\big) - C(N).$$

The first-order condition is

$$PL\mu_N - C_N = 0, \tag{1}$$

where C_N is the cost of the Nth research plot, and $C_N > 0$, $C_{NN} > 0$. Condition (1) implies that the optimal number of experiments is such that the expected value of the marginal experiment, $(PL\mu_N)$, equals the marginal cost of experiments, (C_N). Furthermore, the analysis can show that the research effort increases with the size of the region, $(\partial N/\partial L > 0)$, and the scarcity of the product, $(\partial N^*/\partial P > 0)$. Similarly, lower research costs will lead to more research effort.

The outcome of research leading to innovations is subject to much uncertainty and, in cases where a decision-maker is risk averse, risk considerations will affect whether and to what extent experiments will be undertaken. For simplicity, consider a case where decision-makers maximize a linear combination of mean and variance of profits, and thus the optimization problem is

$$\max_N PL\big[\mu(N) - C(N)\big] - \frac{1}{2}\phi P^2 L^2 \sigma^2(N),$$

where $\sigma^2(N)$ is the variance of Y_N^*, the maximum value of yield of N experiments, and ϕ is a risk-aversion coefficient. The variance of maximum outcome of N experiments declines with N in most cases so that $\sigma_N^2 = \partial\sigma^2(N)/\partial N < 0$. The first-order condition determining N is

$$PL\mu_N - \phi\sigma_N^2 P^2 L^2 - C_N = 0. \tag{2}$$

Under risk aversion, N is determined so that the marginal effect of an increase of N or expected revenues plus the marginal reduction in the cost of risk bearing is equal to the marginal cost of experiments. A comparison of conditions (1) and (2) suggests that the risk-reducing effect of extra experiments will increase the marginal benefit of experiments under risk aversion. Thus, a risk-averse decision-maker who manages a line of research, is likely to carry out more experiments than a risk-neutral decision-maker. Note, however, that expected profits under risk aversion are smaller than under risk neutrality since risk-neutral decision-makers do not have a risk-carrying cost. If experimentation has a significant fixed cost ($C(N) = C_0 + C_1(N)$), there may be situations when risk aversion may prevent carrying out certain lines of research that would be done under risk neutrality. Furthermore, one can expand the model to show that risk considerations may lead risk-averse decision-makers to carry out several substitutable research lines simultaneously in order to diversify and reduce the cost of risk bearing. Thus, uncertainty about the research outcome may deter investment in discovery research, but it may increase and diversify the research efforts once they take place.

There has not been much research on investment in certain lines of research over time. However, the Evenson–Kislev model suggests that there is a decreasing expected marginal gain from experiments. If a certain yield was established after an initial period of experimentation, the model can be expanded to show that the greater the initial yield, the smaller the optimal experiment in the second period. That suggests that the number of experiments carried out in a certain line of research will decline over time, especially once significant success is obtained, or when it is apparent that there are decreasing marginal returns to research. On the other hand, technological change that reduces the cost of innovative efforts may increase experimentation. Indeed, we have witnessed, over time, the tendency to move from one research line to another and, thus, both dynamic and risk considerations tend to diversify innovative efforts.

The Evenson–Kislev model explains optimal investment in one line of research. However, research programs consist of several research lines. The model considers a price-taking firm that produces Y units of output priced at P and also generates its own technology through innovative activities (research and development). There are J parallel lines of innovation, and j is the research line indicator, $j = 1, \dots, J$. Let V_j be the price of one unit of the jth innovation line and m_j be the number of units used in this line. Innovations affect output through a multiplicative effect to the production function, $g(m_i, \dots, m_j)$, and by improving input use effectiveness. The producers use I inputs, and i is the input indicator, $i = 1, \dots, I$. Let the vector of inputs be $m = \{m_i, \dots, m_j\}$. We distinguish between the actual unit of input i used by the producer, X_i, and the effective input e_i where $e_i = h_i(m)X_i$. Thus, it is assumed that a major effect of the innovation is to increase input use efficiency, and the function $h_i(m)$ denotes the effect of all the lines of input effectiveness. An innovative line j may increase effectiveness of input i, and in this case $\partial h_i/\partial m_j > 0$. Thus, the production function of the producer is

$$Y = g(m) f\bigl(X_1 h_1(m), X_2 h_2(m), \dots, X_I h_I(m)\bigr).$$

For simplicity, assume that, without any investment in innovation, $h_i(m) = 1$, for all i; thus, $Y = f(X_1, \dots, X_2)$. The producer has to determine optimal allocation of resources among inputs and research lines. In particular, the choice problem is

$$\max_{X_i,m_i} pg(m)f\left[X_1h_1(m), X_2h_2(m), X_3h_3(m), X_Ih_I(m)\right] - \sum_{i=1}^{I} w_iX_i - \sum_{j=1}^{J} v_im_i,$$

where w_i is the price of the ith input and v_j is the price of one unit of the jth line of innovation. The first-order condition to determine use of the ith input is

$$pg(m)\frac{\partial F}{\partial e_i}h_i(w) - w_i = 0 \quad \forall i. \tag{3}$$

Input i will be chosen at a level where the value of marginal product of input i's effective units, $pg(m)\frac{\partial F}{\partial e_i}$, is equal to the price of input i's effective units, which is $w_i/h_i(m)$. If the innovations have a positive multiplicative effect, $g(m) > 1$, and increase input use efficiency, $h_i(m) > 1$, then the analysis in [Khanna and Zilberman (1997)] suggests that innovations are likely to increase output but may lead to either an increase or decrease in input use. Input use is likely to increase with the introduction of innovations in cases where they lead to substantial increases in output. Modest output effects of innovations are likely to be associated with reduced input use levels.[2]

The optimal effort devoted to innovation line j is determined according to

$$\frac{\partial g}{\partial m_i}pf(m) + g(m)p\sum_{i=1}^{I}\frac{\partial h_i}{\partial m_j}X_i - v_j = 0 \quad \forall j. \tag{4}$$

Let the elasticity of the multiplicative effect of innovation with respect to the level of innovation j be denoted by $\varepsilon^{g}_{m_j} = \frac{\partial g}{\partial m_j}\frac{m_j}{g(m)}$, and let the elasticity of input i's effectiveness coefficient, with respect to the level of innovation j, be $\varepsilon^{h_i}_{m_j} = \frac{\partial h_i}{\partial m_j}\frac{m_j}{h_i}$. Using (3), the first-order condition (4) becomes

$$PY\left[\varepsilon^{g}_{m_j} + \sum_{i=1}^{I} S_i\varepsilon^{h_i}_{m_j}\right] - m_jv_j = 0, \tag{5}$$

where $S_i = w_iX_i/PY$ is the revenue share of input i. Condition (5) states that, under optimal resource allocation, the expenditure share (in total revenue of innovation line j) will be equal to the sum of elasticities of the input effectiveness, with respect to research line j, and the elasticity of the multiplicative output coefficient with respect to this research line. This condition suggests that more resources are likely to be allocated to

---

[2] Khanna and Zilberman (1997) related the impact of technological change on input use to the curvature of the production function. If marginal productivity of e_i declines substantially with an increase in e_i, the output effects are restricted and innovation leads to reduced input use.

research lines with higher productivity effects that mostly impact inputs with higher expenditure shares that have a relatively lower cost.[3]

Risk considerations provide part of the explanation for such diversification, but whether innovations are complements or substitutes may also be a factor. When the tomato harvester was introduced in California, it was accompanied by the introduction of a new, complementary tomato variety [de Janvry et al. (1981)]. McGuirk and Mundlak's (1991) analysis of the introduction of high-yield "green revolution" varieties in the Punjab shows that it was accompanied by the intensification of irrigation and fertilization practices.

The induced innovation hypothesis can be expanded to state that investment in innovative activities is affected by shadow prices implied by government policies and regulation. The tomato harvester was introduced following the end of the Bracero Program, whose termination resulted in reduced availability of cheap immigrant workers for California and Florida growers. Environmental concerns and regulation have led to more intensive research and alternatives for the widespread use of chemical pesticides. For example, they have contributed to the emergence of integrated pest management strategies and have prompted investment in biological control and biotechnology alternatives to chemical pesticides.

Models of induced innovation should be expanded to address the spatial variability of agricultural production. The heterogeneity of agriculture and its vulnerability to random events such as changes in weather and pest infestation led to the development of a network of research stations. A large body of agricultural research has been aimed at adaptive innovations that develop practices and varieties that are appropriate for specific environmental and climatic conditions. The random emergence of new diseases and pests led to the establishment of research on productivity maintenance aimed at generating new innovations in response to adverse outcomes whenever they occurred.

The treatment of the mealybug in the cassava in Africa is a good example of responsive research. Cassava was brought to Africa from South America 300 years ago and became a major subsistence crop. The mealybug, one of the pests of cassava in South America, was introduced to Africa and reduced yields by more than 50 percent in 1983–84; without treatment, the damage could have had a devastating effect on West Africa [Norgaard (1988)]. The International Institute of Tropical Agriculture launched a research program which resulted in the introduction of a biological control in the form of a small wasp, *E lopezi*, that is a natural enemy of the pest in South America. Norgaard estimated the benefit/cost ratio of this research program to be 149 to 1, but his calculation did not take into account the cost of the research that established the methodology of biological control, and the fixed cost associated with maintaining the infrastructure to respond to the problem.

Induced innovation models such as Binswanger's (1974) are useful in linking the evolution of innovations to prices, costs, and technology. However, they ignore some

---

[3] Binswanger (1974) proves these assertions under a very narrow set of conditions.

of the important details that characterize the system leading to agricultural innovations.[4]

Typically, new agricultural technologies are not used by the entities that develop them (e.g., universities and equipment manufacturers). Different types of entities have their distinct decision-making procedures that need to be recognized in a more refined analysis of agricultural innovations. The next subsection will analyze resource allocation for the development of new innovations in the public sector, and that will be followed by a discussion of specific institutions and incentives for innovation activities (patents and intellectual property rights) in the private sector.

Induced innovations by agribusiness apply to innovations beyond the farm gate. In much of the post World War II period, there has been an excess supply of agricultural commodities in world markets. This has led to a period of low profitability in agriculture, requiring government support. While increasing food quantity has become less of a priority, increasing the value added to food products has become a major concern of agriculture and agribusiness in developed nations. Indeed, that has been the essence of many of the innovations related to agriculture in the last 30 years. Agribusiness took advantage of improvements in transportation and weather-controlled technologies that led to innovations in packing, storage, and shipping. These changes expanded the availability as well as the quality of meats, fruits, and vegetables; increased the share of processing and handling in the total food budget; and caused significant changes in the structure of both food marketing industries and agriculture.

It is important to understand the institutional setup that enables these innovations to materialize. While there has not been research in this area, it seems that the availability of numerous sources of funding to finance new ventures (e.g., venture capital, stock markets, mortgage markets, credit lines from buyers) enables the entities that own the rights to new innovations to change the way major food items are produced, marketed, and consumed.

1.2. Political economy of publicly funded innovations

Applied R&D efforts are supported by both the public and private sectors because of the innovations they are likely to spawn. Public R&D efforts are justified by the public-good nature of these activities and the inability of private companies to capture all the benefits resulting from farm innovations.

Studies have found consistently high rates of returns (above 20 percent) to public investment in agricultural research and extension, indicating underinvestment in these activities, see [Alston et al. (1995), Huffman (1998)]. Analysis of patterns of public spending for R&D in agriculture shows that federal monies tend to emphasize research

---

[4] The Binswanger model (1974) is very closely linked to the literature on quantifying sources of productivity in agriculture. For an overview of this important body of literature, which benefited from seminal contributions by Griliches (1957, 1958) and Mundlak, see [Antle and McGuckin (1993)].

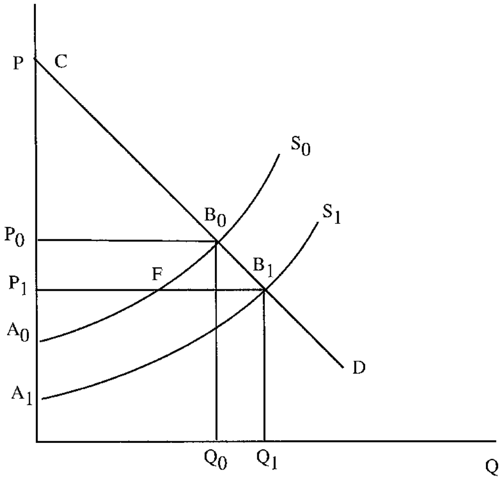

Figure 2.

on commodities that are grown in several states (e.g., wheat, corn, rice), while individual states provide much of the public support for innovation-inducing activities for crops that are specialties of the state (e.g., tomatoes and citrus in Florida, and fruits and vegetables in California). The process of devolution has also applied to public research and, over the years, the federal share in public research has declined relative to the state's share. Increased concern for environmental and resource management issues over time led to an increase in relative shares of public research resources allocated to these issues in agriculture [Huffman and Just (1994)].

Many of the studies evaluating returns to public research in agriculture (including Griliches' 1957 study on hybrid corn that spawned the literature) rely on partial equilibrium analysis, depicted in Figure 2.

The model considers an agricultural industry facing a negatively sloped demand curve D. The initial supply is denoted byS_0, and the initial price and quantity are P_0 and Q_0, respectively. Research, development, and extension activities led to adoption of an innovation that shifts supply to S_1, resulting in price reduction to P_1, and consumption gain Q_1.[5] The social gain from the innovation is equal to the area $A_0B_0B_1A_1$ in Figure 2 denoted by G. If the investment leading to the use of the innovation is denoted by I, the net social gain is $NG = G - I$, and the social rate of return to appropriate research development and extension activities is NG/I.

The social gain from the innovation is divided between consumers and producers. In Figure 2, consumer gain is equal to the area $P_0B_0B_1P_1$. Producer gain is $A_0FA_1B_1$

---

[5] Of course, actual computation requires discounting and aggregation, and benefits over time, and may recognize the gradual shift in supply associated with the diffusion process.

because of lower cost and higher sales, but they lose $P_0B_0FP_1$ because of lower price. If demand is sufficiently inelastic, producers may actually lose from public research activities and the innovations that they spawn. Obviously, producers may not support research expenditures on innovations that may worsen their well-being, and distributional considerations affect public decisions that lead to technological evolution.[6]

This point was emphasized in Schmitz and Seckler's (1970) study of the impact of the introduction of the tomato harvester in California. They showed that society as a whole gained from the tomato harvester, while farm workers lost from the introduction of this innovation. The controversy surrounding the tomato harvester [de Janvry et al. (1981)] led the University of California to de-emphasize research on mechanical innovations.

De Gorter and Zilberman (1990) introduced a simple model for analyzing political economic considerations associated with determining public expenditures on developing new agricultural technologies. Their analysis considers a supply-enhancing innovation. They consider an industry producing Y units of output. The cost function of the industry is $C(Y, I)$ and depends on output and investment in R&D where the level is I. This cost function is well behaved and an increase in I tends to reduce cost at a decreasing rate $\partial c/\partial I < 0$, and $\partial^2 c/\partial I^2 > 0$ and marginal cost of output $\partial^2 c/\partial I \partial Y < 0$. Let the cost of investment be denoted by r and the price of output by P. The industry is facing a negatively sloped demand curve, $Y = D(P)$. The gross surplus from consumption is denoted by the benefit function $B(Y) = \int_0^Y P(z)\,\mathrm{d}z$, where $P(Y)$ is inverse demand.

Social optimum is determined at the levels of Y and I that maximize the net surplus. Thus, the social optimization problem is

$$\max_{Y,I} B(Y) - C(Y, I) - rI,$$

and the first-order optimality conditions are

$$\frac{\partial B}{\partial Y} - \frac{\partial C}{\partial Y} = 0 \Rightarrow P(Y) = \frac{\partial C}{\partial Y}, \tag{6}$$

and

$$-\frac{\partial C}{\partial I} - \gamma = 0. \tag{7}$$

Condition (6) is the market-clearing rule in the output market, where price is equal to marginal cost. Condition (7) states the optimal investment in R&D at a level where

---

[6] Further research is needed to understand to what extent farmers take into consideration the long-term distributional effects of research policy. They may be myopic and support a candidate who favors any research, especially when facing a pest or disease.

the marginal reduction in production cost because of investment in R&D is equal to the cost of investment. The function $-\partial C/\partial I$ reflects a derived demand for supply-shifting investment and, by our assumptions, reducing the price of investment (γ) will increase its equilibrium level. Condition (7) does not likely hold in reality. However, it provides a benchmark with which to assess outcomes under alternative political arrangements.

De Gorter and Zilberman (1990) argued that the political economic system will determine both the level of investment in R&D and the share of the burden of financing it between consumers (taxpayers) and producers. Let Z be the share of public investment in R&D financed by producers. Thus, $Z = 0$ corresponds to the case where R&D is fully financed by taxpayers, and $Z = 1$ where R&D is fully financed by producers. The latter case occurs when producers use marketing orders to raise funds to collectively finance research activities. There are many cases in agriculture where producers compete in the output market but cooperate in technology development or in the political arena [Guttman (1978)].

De Gorter and Zilberman (1990) compare outcomes under alternative arrangements, including the case where producers both determine and finance investment in R&D. In this case, I is the result of a constrained optimization problem, where producer surplus, $PS = P(Y)Y - C(Y, I)$, minus investment cost, rI, is maximized subject to the market-clearing constraint in the output market $P(Y) = \partial C/\partial Y$. When there is internal solution, the first-order optimality condition for I is

$$-\frac{\partial C}{\partial I} - \eta = r, \tag{8}$$

where

$$\eta = -Y\frac{\partial^2 C}{\partial Y \partial C}\left[1 - \left(\frac{\partial^2 C}{\partial Y^2}\right)\Big/\left(\frac{\partial P}{\partial I}\right)\right].$$

The optimal solution occurs at a level where the marginal cost saving due to investment minus the term η, which reflects the loss of revenues because of price reduction, is equal to the marginal investment cost, r. The loss of revenues because of a price reduction due to the introduction of a supply-enhancing innovation increases as demand becomes less elastic. A comparison of (8) to (7) suggests that under-investment in agricultural R&D is likely to occur when producers control its level and finance it, and the magnitude of the under-investment increases as demand for the final product becomes less elastic. Below a certain level of demand elasticity, it will be optimal for producers not to invest in R&D at all. If taxpayers (consumers) pay for research but producers determine its level, the optimal investment will occur where the marginal reduction in cost due to the investment is equal to η, the marginal loss in revenue due to price reduction. When the impact of innovation on price is low (demand for final product is highly elastic), producer control may lead to over-investment if producers do not pay for it. However, when $\eta > r$, and expansion of supply leads to significant price reduction, even when taxpayers pay for public agricultural research, producer determination of its level will lead to under-investment.

The public sector has played a major role in funding R&D activities that have led to new agricultural innovations, especially innovations that are disembodied or are embodied but non-shielded. Rausser and Zusman (1991) have argued that choices in political-economic systems are effectively modeled as the outcome of cooperative games among parties. Assume that two groups, consumers/taxpayers and producers, are affected by choices associated with investment in the supply-increasing innovation mentioned above. The political-economic system determines two parameters. The first is the investment in the innovation (I) and the second is the share of the innovation cost financed by consumers. Let this share be denoted as z; thus, the consumer will pay $zc(I)$ for the innovation cost. It is assumed that the investment in the innovation is non-negative ($I \geqslant 0$), but z is unrestricted ($z > 1$ implies that the producers are actually subsidized).

The net effects of the investment and finance of innovations on consumers/taxpayers' welfare and producers' welfare are $\Delta CS(I) - zc(I)$ and $\Delta PS(I) - (1-z)c(I)$, respectively. The choice of the innovation investment and the sharing coefficients are approximated by the solution to the optimization problem

$$\max_{I,Z}\big(\Delta CS(I) - zc(I)\big)^{\alpha}\big(\Delta PS(I) - (1-z)c(I)\big)^{1-\alpha}, \tag{9}$$

where α is the consumer weight coefficient, $0 \leqslant \alpha \leqslant 1$. The optimization problem (9) (i) incorporates the objective of the two parties; (ii) leads to outcomes that will not make any of the parties worse off; (iii) reflects the relative power of the parties (when α is close to one, consumers dominate decision-making but the producers have much of the power when $\alpha \to 0$); and (iv) reflects decreasing marginal valuation of welfare gained by most parties.[7]

After some manipulations, the solutions to this optimization problem are presented by

$$\frac{\Delta CS(I)}{\partial I} + \frac{\partial \Delta PS(1)}{\partial I} = \frac{\partial C}{\partial I}; \tag{10}$$

$$\frac{\alpha_1}{1-\alpha_1} = \frac{\Delta CS(I) - zc(I)}{\Delta PS(I) - (1-z)c(I)}. \tag{11}$$

Equation (10) states that innovation investment will be determined when the sum of the marginal increase in consumer and producer surplus is equal to the marginal cost of investment innovation. This rule is equivalent to equating the marginal cost of innovation investment with its marginal impact on market surplus (since $\Delta MS = \Delta PS + \Delta CS$).

---

[7] $\partial PG/\partial I > 0$, $\partial^2 PG/\partial I^2 < 0$, $\partial CS/\partial I > 0$, $\partial^2 CS/\partial I^2 < 0$.

Equation (11) states that the shares of two groups in the total welfare gain are equal to their political weight coefficients. Thus, if α_1 is equal to, say, 0.3 and consumers have 30 percent of the weight in determining the level and distribution of finance of innovation research, then they will receive 30 percent of the benefit. Producers will receive the other 70 percent. Equation (9) suggests that the political weight distribution does not affect the total level of investment in innovation research that is socially optimal, but only affects the distribution of benefits. If farmers have more political gain in determining the outcome because of their intense interest in agricultural policy issues, they will gain much of the benefit from innovation research.

The cooperative game framework is designed to lead to outcomes where both parties benefit from the action they agree upon. Since both demand and supply elasticities for many agricultural commodities are relatively low, producer surplus is likely to decline with expanded innovation research. When these elasticities are sufficiently low, farmers as a group will directly lose from expanded innovation research unless compensated. Thus, in certain situations and for some range of products, positive innovation research is not feasible unless farmers are compensated. This analysis suggests a strong link between public support for innovation research and programs that support farm income. In such situations innovation research leads to a significant direct increase in consumer surplus through increased supplies and a reduction in commodity prices. It will also result in an increase in farmer subsidies by taxpayers. Thus, for a range of commodities with low elasticities of output supply and demand, consumers/taxpayers will finance public research and compensate farmers for their welfare losses. For commodities where demand is quite elastic, say about 2 or 3, and both consumers and producers gain significantly from the fruits of innovation research, both groups will share in financing the research. When demand is very elastic and most of the gain goes to producers, the separate economic frameworks suggest that they are likely to pay for this research significantly, but if their political weight in the decision is quite important (α close to 1), they may benefit immensely from the fruits of the innovation research, but consumers may pay for a greater share of the research.

While this political analysis framework is insightful in that it describes the link between public support for agricultural research and agricultural commodity programs, it may be off the mark in explaining the public investment in innovation research in agriculture, since there is a large array of studies that argues that the rate of return for agricultural research is very high, and thus there is under-investment. One obvious limitation of the model introduced above is that it assumes that the outcomes of research innovation are certain. However, there is significant evidence that returns for research projects are highly skewed. A small number of products may generate most of the benefits, and most projects may have no obvious outcome at all. This risk consideration has to be incorporated explicitly into the analysis determining the level of investment in innovation research. Thus, when consumers consider investment I in innovation research, they are aware that each investment level generates a distribution of outcome, and they will consider the expected consumer surplus gain associated with I. Similarly, producers are aware of the uncertainty involved with innovation research, and they will consider the expected producer surplus associated with each level in assessing the various levels of innovation research.

1.3. Policies and institutions for managing innovation activities

The theory of induced innovations emphasizes the role of general economic conditions in shaping the direction of innovation activities. However, the inducement of innovations also requires specific policies and institutions that provide resources to would-be innovators and enable them to reap the benefits from their innovations.

Patent protection is probably the most obvious incentive to innovation activities. Discoverers of a new patentable technology have the property right for its utilization for a well-defined period of time (17 years in the U.S.). An alternative tool may be a prize for the discoverer of a new technology, and Wright (1983) presents examples where prizes have been used by the government to induce creative solutions to difficult technological problems. A contract, which pays potential innovators for their efforts, is a third avenue in motivating innovative activities. Wright (1983) develops a model to evaluate and compare these three operations. Suppose that the benefits of an innovation are known and equal to B. The search for the innovation is done by n homogeneous units, and the probability of discovery is $P(n)$, with

$$\frac{\partial P}{\partial n} > 0, \quad \frac{\partial^2 P}{\partial n^2} > 0.$$

The cost of each unit is C. The social optimization problem to determine optimal research effort is

$$\max_n P(n)B - nC,$$

and socially optimal u is determined when

$$\frac{\partial P}{\partial N} B = C. \tag{12}$$

The expected marginal benefit of a research unit is equal to its cost. This rule may be used by government agents in determining the number of units to be financed by contracts. On the other hand, under prizes or patents, units will join in the search for the innovation as long as their expected net benefits from the innovation, $P(N)B/N$, are greater than the unit cost C. Thus, optimal N under patents is determined when

$$\frac{P(N)}{N} B = C. \tag{13}$$

Assuming decreasing marginal probability of discovery, average probability of discovery for a research unit is greater than the marginal probability, $P(N)/N > \partial P/\partial N$.

Thus, a comparison of (12) with (13) suggests that there will be over-investment in experimentation under patents and prizes. In essence, under patents and prizes, research units are *ex ante*, sharing a common reward and, as in the classical "Tragedy of the Commons" problem, will lead to overcrowding. Thus, when the award for a discovery is known, contracts may lead to optimal resource allocation.

Another factor that counters the oversupply of research efforts under patent relative to contracts is that the benefits of the innovation under patent may be smaller than under contract. Let B_p be the level of benefits considered for deriving

$$\frac{\mathrm{d}L_1^r}{\mathrm{d}L} = \eta \frac{L_1^r}{L} + (r - \eta)R,$$

the research effort under the patent system. B_p is equal to the profits of the monopolist patent owner. Let B_c be the level of benefits considered in determining η_c, the research effort under contract. If η_c is determined by a social welfare maximizing agent, B_c is the sum of consumers' and producers' surplus from the use of the innovation. In this case $B_c > B_N$. Thus, in the case of full information about the benefits and costs, more research will be conducted under contracts if

$$\frac{B_c}{B_p} > \frac{\frac{P(\eta_p)}{\eta_p}}{\frac{\partial P}{\partial n}(\eta_c)}.$$

In many cases, the uncertainty regarding the benefits of an innovation at the discovery and patent stages is very substantial. Commercialization of a patent may require significant investment, and a large percentage of patents are not utilized commercially [Klette and Griliches (1997)]. Commercialization of an innovation requires upscaling and development, registration (in the case of chemical pesticides), marketing, and development of production capacity for products resulting from the patents. Large agribusiness firms have the resources and capacity to engage in commercialization, and they may purchase the right to utilize patents from universities or smaller research and development firms. Commercialization may require significant levels of research that may result in extra patents and trade secrets that strengthen the monopoly power of the commercializing firm. Much of the research in the private sector is dedicated to the commercialization and the refinement of innovations, while universities emphasize discovery and basic research. Thus, Alston, Norton, and Pardey (1995) argue that private-sector and public-sector research spending are not perfect substitutes. Actually, there may be some complementarity between the two. An increase in public sector research leads to patentable discoveries, and when private companies obtain the rights to the patents, they will invest in commercialization research. Private sector companies have recognized the unique capacity of universities to generate innovations, and this has resulted in support for university research in exchange for improved access to obtain rights to the innovations [Rausser (1999)].

1.4. Factors beyond the farm gate

Over the years, product differentiation in agriculture has increased along with an increase in the importance of factors beyond the farm gate and within specialized agribusiness. This evolution is affecting the nature and analysis of agricultural research. Economists have recently addressed how the vertical market structure of agriculture conditions the benefits of agricultural research, and also how farm-level innovation may contribute to changes in the downstream processing sector.

One salient fact about the food-processing sector is that it tends to be concentrated. The problem of oligopsonistic competition in the food processing sector has been addressed by Just and Chern (1980), Wann and Sexton (1992), and Hamilton and Sunding (1997). Two recent papers by Hamilton and Sunding (1998) and Alston, Sexton, and Zhang (1997) point out that the existence of noncompetitive behavior downstream has important implications for the impacts of farm-level technological change.

Consider a situation where the farm sector is competitive and sells its product to a monopsonistic processing sector. Let X denote the level of farm output, R be research expenditures, W be the price paid for the farm output, P be the price of the final good, and f be the processing production function. The monopsonist's problem is then

$$\max_{X} Pf(X) - W(X, R)X. \tag{14}$$

Since the farm sector is competitive, W is simply the marginal cost of producing the raw farm good. It is natural to assume that $\partial W/\partial X > 0$ since supply is positively related to price and $\partial W/\partial R < 0$ since innovation reduces farm costs. Second derivatives of the marginal cost function are more ambiguous. Innovations that increase crop yields may tend to make the farm supply relation more elastic, and in this case, $\partial^2 W/\partial X \partial R < 0$. However, industrialization may result in innovations that limit capacity or increase the share of fixed costs in the farm budget. In this case, $\partial^2 W/\partial X \partial R > 0$ and the farm supply relation becomes less elastic as a result of innovation.

Totally differentiating the solution to (14), it follows that the change in farm output following an exogenous increase in research expenditures is

$$\frac{\mathrm{d}X}{\mathrm{d}R} = \frac{-\left(P\frac{\partial f}{\partial X} - \frac{\partial^2 W}{\partial X \partial R}X - \frac{\partial W}{\partial R}\right)}{\mathrm{SOC}}.$$

The numerator is of indeterminate sign, while the denominator is the monopsonist's second-order condition, and thus negative. The first and third terms of the numerator are positive and negative, respectively, by the assumptions of positive marginal productivity in the processing sector, and the marginal cost-reducing nature of the innovation. This last effect is commonly termed the "shift" effect of innovation on the farm supply relation. There is also a "pivot" effect to consider, however, which is represented by the second term in the numerator. As pointed out earlier, this term can be either positive or negative depending on the form of the innovation. In fact, if public research makes the farm supply curve sufficiently inelastic, then a cost-reducing innovation can actually reduce the equilibrium level of farm output. Hamilton and Sunding (1998) make this point in the context of a more general model of oligopsony in the processing sector. They point out that an inelastic pivot increases the monopsonist's degree of market power and increases its ability to depress farm output. If the farm supply relation becomes sufficiently inelastic following innovation, this effect can override the output-enhancing effect of cost-reduction. Note further that the "pivot" effect only matters when there is imperfect competition downstream; the second term in the numerator disappears if the processing sector is competitive. Thus, in the case of perfect downstream competition, reduction of the marginal cost of farming is a sufficient condition for the level of farm output to increase.

The total welfare change from farm research is also affected by downstream market power. In the simple model above, social welfare is given by the following expression:

$$SW = \int_0^{Y(X(R))} P(Z)\,\mathrm{d}Z - \int_0^{X(R)} W(Z,R)\,\mathrm{d}Z, \tag{15}$$

where $P(Z)$ is the inverse demand function for the final good. The impact of public research is then

$$\frac{\mathrm{d}SW}{\mathrm{d}R} = \left(P\frac{\partial f}{\partial X} - W\right)\frac{\mathrm{d}X}{\mathrm{d}R} - \int_0^X \frac{\partial W}{\partial R}\,\mathrm{d}Z.$$

This expression underscores the importance of downstream market structure. Under perfect competition, the wedge between the price of the final good and its marginal cost is zero, and so the first term disappears. In this case, the impact of farm research on social welfare is determined completely by its impact on the marginal cost of producing the farm good.[8] When the processing sector is imperfectly competitive, however, some interesting results emerge. Most importantly, if farm output declines following the cost-reducing innovation (which can only occur if the farm supply relation becomes more inelastic), then social welfare can actually decrease. This argument was developed in Hamilton and Sunding (1998), who describe the final outcome of farm-level innovation as resulting from two forces: the social welfare improving effect of farm cost reduction and the welfare effect of changes in market power in the processing industry.

Hamilton and Sunding (1998) show that the common assumption of perfect competition may seriously bias estimates of the productivity of farm-sector research. Social returns are most likely to be overestimated when innovation reduces the elasticity of the farm supply curve, and when competition is assumed in place of actual imperfect competition. Further, Hamilton and Sunding demonstrate that all of the inverse supply functions commonly used in the literature preclude the possibility that $\partial^2 W/\partial X \partial R > 0$,

---

[8] This point has also been noted recently in Sunding (1996) in the context of environmental regulation.

and thus rule out, *a priori*, the type of effects that result from convergent shifts. More flexible forms and more consideration of imperfect competition are needed to capture the full range of possible outcomes.

The continued development of agribusiness is leading to both physical and intellectual innovation. Feed suppliers, in an effort to expand their market, contributed to the evolution of large-scale industrialized farming. This is especially true in the poultry sector. Until the 1950s, separate production of broilers and chickens for eggs was scarce. The price of chicken meat fluctuated heavily, and that limited producers' entry into the emerging broiler industry. Feed manufacturers provided broiler production contracts with fixed prices for chicken meat, which led to vertical integration and modern industrial methods of poultry production. These firms not only offer output contracts, but they also provide production contracts and contribute to the generation of production technology. Recently, this same phenomenon has occurred in the swine sector, where industrialization has reduced the cost of production.

But agribusiness has spurred the development of another set of quality-enhancing innovations. Again, some of the most important developments have been in the poultry industry. Tyson Foods and other companies have produced a line of poultry products where meats are separated according to different categories, cleaned, and made ready to be cooked. The development of these products was based on the recognition of consumers' willingness to pay to save time in food preparation. In essence, the preparation of poultry products has shifted labor from the household to the factory where it can be performed more efficiently.

In addition to enhancing the value of the final product, the poultry agribusiness giants introduced institutional technological innovations in poultry production [Goodhue (1997)]. Packing of poultry has shifted to rather large production units that have contractual agreements with processors/marketers. The individual production units receive genetic materials and production guidance from the processor/marketer, and their pay is according to the relative quality. This set of innovations in production and marketing has helped reduce the relative price of poultry and increase poultry consumption in the United States and other countries over the last 20 years. Similar institutional and production innovations have occurred in the production of swine, high-value vegetables, and, to some extent, beef. These innovations are major contributors to the process of industrialization of agriculture. While benefiting immensely from technology generated by university research, these changes are the result of private sector efforts and demonstrate the important contributions of practitioners in developing technologies and strategies.

2. Technology adoption

2.1. *Adoption and diffusion*

There is often a significant interval between the time an innovation is developed and available in the market, and the time it is widely used by producers. Adoption and diffusion are the processes governing the utilization of innovations. Studies of adoption behavior emphasize factors that affect if and when a particular individual will begin using an innovation. Measures of adoption may indicate both the timing and extent of new technology utilization by individuals. Adoption behavior may be depicted by more than one variable. It may be depicted by a discrete choice, whether or not to utilize an innovation, or by a continuous variable that indicates to what extent a divisible innovation is used. For example, one measure of the adoption of a high-yield seed variety by a farmer is a discrete variable denoting if this variety is being used by a farmer at a certain time; another measure is what percent of the farmer's land is planted with this variety.

Diffusion can be interpreted as aggregate adoption. Diffusion studies depict an innovation that penetrates its potential market. As with adoption, there may be several indicators of diffusion of a specific technology. For example, one measure of diffusion may be the percentage of the farming population that adopts new innovations. Another is the land share in total land on which innovations can be utilized. These two indicators of diffusion may well convey a different picture. In developing countries, 25 percent of farmers may own or use a tractor on their land. Yet, on large farms, tractors will be used on about 90 percent of the land. While it is helpful to use the term "adoption" in depicting individual behavior towards a new innovation and "diffusion" in depicting aggregate behavior, in cases of divisible technology, some economists tend to distinguish between intra-firm and inter-firm diffusion. For example, this distinction is especially useful in multi-plant or multi-field operations. Intra-firm studies may investigate the percentage of a farmer's land where drip irrigation is used, while inter-firm studies of diffusion will look at the percentage of land devoted to cotton that is irrigated with drip systems.

2.1.1. The S-shaped diffusion curve

Studies of adoption and diffusion behaviors were undertaken initially by rural sociologists. Rogers (1962) conducted studies on the diffusion of hybrid corn in Iowa and compared diffusion rates of different counties. He and other rural sociologists found that in most counties diffusion was an S-shaped function of time. Many of the studies of rural sociologists emphasized the importance of distance in adoption and diffusion behavior. They found that regions that were farther away from a focal point (e.g., major cities in the state) had a lower diffusion rate in most time periods. Thus, there was emphasis on diffusion as a geographic phenomenon.

Statistical studies of diffusion have estimated equations of the form

$$Y_t = K\left[1 + \mathrm{e}^{-(a+bt)}\right]^{-1}, \tag{16}$$

where Y_t is diffusion at time t (percentage of land for farmers adopting an innovation), K is the long-run upper limit of diffusion, a reflects diffusion at the start of the estimation period, and b is a measure of the pace of diffusion.

With an S-shaped diffusion curve, it is useful to recognize that there is an initial period with a relatively low adoption rate but with a high rate of change in adoption.

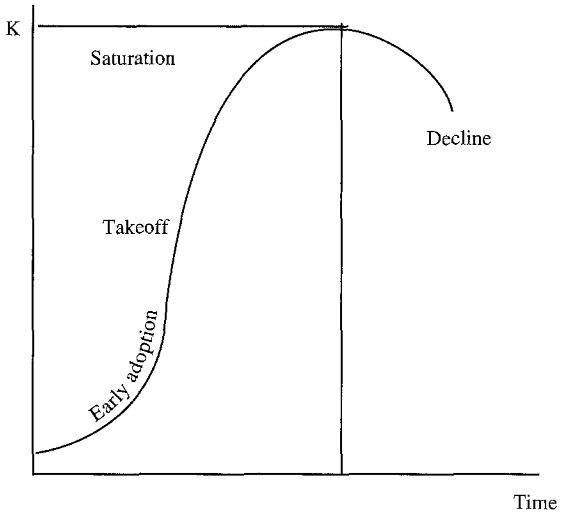

Figure 3.

Figure 3 shows this as a period of introduction of a technology. Following is a takeoff period when the innovation penetrates the potential market to a large extent during a short period of time. During the initial and takeoff periods, the marginal rate of diffusion actually increases, and the diffusion curve is a convex function of time. The takeoff period is followed by a period of saturation where diffusion rates are slow, marginal diffusion declines, and the diffusion reaches a peak. For most innovations, there will also be a period of decline where the innovation is replaced by a new one (Figure 3).

Griliches' (1957) seminal study on adoption of hybrid corn in Iowa's different counties augmented the parameters in (16) with information on rates of profitability, size of farms in different counties, and other factors. The study found that all three parameters of diffusion function (K, a, and b) are largely affected by profitability and other economic variables. In particular, when $\Delta\pi$ denotes the percent differential in probability between the modern and traditional technology, Griliches (1957) found that $\partial a/\partial\Delta\pi$, $\partial K/\partial\Delta\pi$, and $\partial b/\partial\Delta\pi$ are all positive. Griliches' work (1957, 1958) spawned a large body of empirical studies [Feder et al. (1985)]. They confirmed his basic finding that profitability gains positively affect the diffusion process. The use of S-shaped diffusion curves, especially after Griliches (1957) introduced his economic version, has become widespread in several areas. S-shaped diffusion curves have been used widely in marketing to depict diffusion patterns of many products, for example, consumer durables. Diffusion studies have been an important component of the literature on economic development and have been used to quantitatively analyze the processes through which modern practices penetrate markets and replace traditional ones.

2.1.2. Diffusion as a process of imitation

The empirical literature spawned by Griliches (1957, 1958) established stylized facts, and a parallel body of theoretical studies emerged with the goal of explaining its major findings. Formal models used to depict the dynamics of epidemics have been applied by Mansfield (1963) and others to derive the logistic diffusion formula. Mansfield viewed diffusion as a process of imitation wherein contacts with others led to the spread of technology. He considered the case of an industry with identical producers, and for this industry the equation of motion of diffusion is

$$\frac{\partial Y_t}{\partial t} = bY_t\left(1 - \frac{Y_t}{K}\right). \tag{17}$$

Equation (17) states that the marginal diffusion at time t ($\partial Y/\partial t$, the actual adoption occurring at t) is proportional to the product of diffusion level Y_t and the unutilized diffusion potential $(1 - Y_t/K)$ at time t. The proportional coefficient b depends on profitability, firm size, etc. Marginal diffusion is very small at the early stages when $Y_t \to 0$ and as diffusion reaches its limit, $Y_t \to K$. It has an inflection point when it switches from an early time period of increasing marginal diffusion ($\partial^2 Y_t/\partial t^2 > 0$) to a late time period of decreasing marginal diffusion ($\partial^2 Y/\partial t^2 < 0$). For an innovation that will be fully adopted in the long run ($K = 1$),

$$\frac{\partial Y_t}{\partial t} = bY_t(1 - Y_t),$$

the inflection point occurs when the innovation is adopted by 50 percent of producers. Empirical studies found that the inflection point occurs earlier than the simple dynamic model in (17) suggests. Lehvall and Wahlbin (1973) and others expanded the modeling of the technology diffusion processes by incorporating various factors of learning and by separating firms that are internal learners (innovators) from those that are external learners (imitators). This body of literature provides a very sound foundation for estimation of empirical time-series data on aggregate adoption levels. However, it does not rely on an explicit understanding of decision-making by individual firms. This criticism led to the emergence of an alternative model of adoption and diffusion, the threshold model.

2.1.3. The threshold model

Threshold models of technology diffusion assume that producers are heterogeneous and pursue maximizing or satisfying behavior. Suppose that the source of heterogeneity is farm size. Let L denote farm size and $g(L)$ be the density of farm size. Thus, $g(L)\Delta L$ is the number of farms between $L - \Delta L/2$ and $L + \Delta L/2$. The total number of farms is then $N = \int_0^\infty g(L)\,\mathrm{d}L$, and the total acreage is $\overline{L} = \int_0^\infty Lg(L)\,\mathrm{d}L$.

Suppose that the industry pursued a traditional technology that generated π_0 units of profit per acre. The profit per acre of the modern technology at time t is denoted by $\pi_1(t)$ and the profit differential per acre is $\Delta\pi_t$. It is assumed that an industry operates under full certainty, and adoption of modern technology requires a fixed cost that varies over time and at time t is equal to F_t. Under these assumptions, at time t there will be a cutoff farm size, $L_t^C = F_t/\Delta\pi_t$, upon which adoption occurs. One measure of diffusion at time t is thus

$$Y_t^1 = \frac{\int_{L_t^C}^{\infty} g(L)\,\mathrm{d}L}{N}, \tag{18}$$

which is the share of farms adopting at time t. Another measure of diffusion of time t is

$$Y_t^2 = \frac{\int_{L_t^C}^{\infty} Lg(L)\,\mathrm{d}L}{\overline{L}}, \tag{19}$$

which is the share of total acres adopting the modern technology at time t.

The diffusion process occurs as the fixed cost of the modern technology declines over time ($\partial F_t/\partial t < 0$) or the variable cost differential between the two technologies increases over time ($\partial\Delta\pi_t/\partial t > 0$). The price of the fixed cost per farm may decrease over time because the new technology is embodied in new indivisible equipment or because it requires an up-front investment in learning. "Learning by doing" may reduce fixed costs through knowledge accumulation. The profit differential often will increase over time because of "learning by using". Namely, farmers will get more yield and save cost with more experience in the use of the new technology.

The shape of the diffusion curve depends on the dynamics of farm size and the shape of farm size distribution. Differentiation of (18) obtains marginal diffusion under the first definition

$$\frac{\partial Y_t^1}{\partial t} = -\frac{g(L_t^C)}{N}\frac{\partial L_t^C}{\partial t}. \tag{20}$$

Marginal diffusion at time t is equal to the percentage of farms adopting technology at this time. It is expressed as $\partial L_t^C/\partial t$ times the density of the farm size distribution at L_t^C, $g(L_t^C)$.

The dynamics of diffusion associated with the threshold model are illustrated in Figure 4. Farm size distribution is assumed to be unimodal. When the new innovation is introduced, only farms with a size greater than L_0^C will adopt. The critical size declines over time and this change triggers more adoption. The marginal adoption between the first and second year is equal to the area $abL_2^C L_1^C$. Figure 4 assumes that the marginal decline in L_t^C is constant because of the density function's unimodality. Marginal diffusion increases during the initial period and then it declines, thus leading to an S-shaped diffusion curve. It is plausible that farm size distribution (and the distribution of other

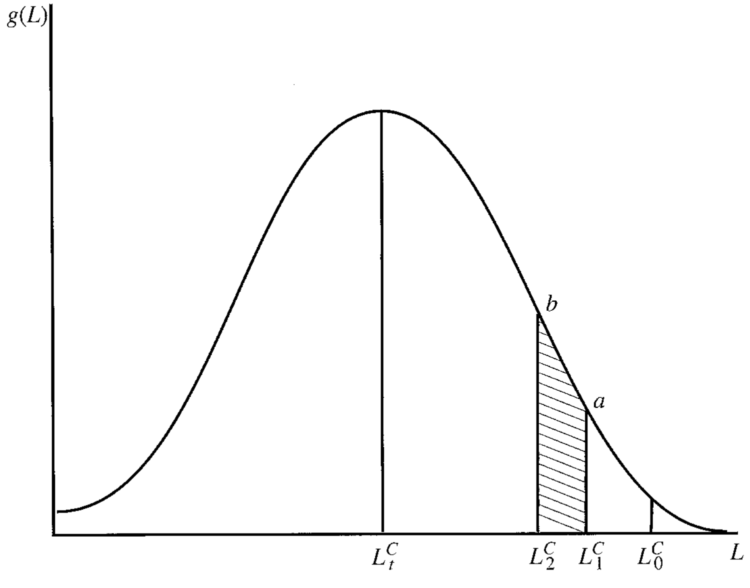

Figure 4.

sources of heterogeneity) will be unimodal and that combined with a continuous decline of L_t^C will lead to S-shaped behavior.[9]

The threshold model was introduced by Paul David (1969) to explain adoption of grain harvesting machinery in the United States in the nineteenth century. He argued that the main source of heterogeneity among farmers was farm size and he derived the minimum farm size required for adoption of various pieces of equipment. Olmstead and Rhode (1993) review historical documents that show that, in many cases, much smaller farms adopted some of the new machinery because farmers cooperated and jointly purchased harvesting equipment. This example demonstrates some of the limitations of the threshold model, especially when heterogeneity results from differences in size.

---

[9] To have an S-shaped behavior, $f^2Y_t^1/ft^2 > 0$ for an initial period with $t < \hat{t}$ and $f^2Y_t^1/ft^2 < 0$ for $t > \hat{t}$. Differentiation of (20) yields

$$\frac{\partial^2 Y_1^t}{\partial t^2} = -\frac{1}{N}\left[\frac{\partial g(L_t^C)}{\partial L_t^C}\left(\frac{\partial L_t^C}{\partial t}\right)^2 + g\left(L_t^C\right)\frac{\partial^2 L_t^C}{\partial t^2}\right].$$

Assuming unimodal distribution, let L_t^C be associated with the model of $g(L)$. As long as $L_t^C > L_{t5}^C \partial g(L_t^C)/\partial L_t^C < 0$, then $L_t^C < L_{t5}^C \partial g(L_t^C)/\partial L_t^C > 0$. At the early periods, $\partial^2 L_t^C/\partial t^2$ may be small or even negative, but as t increases the marginal decline in L_t^C gets smaller and $\partial^2 L_t^C/\partial t^2$ may be positive. Thus, the change of the sign of both elements of $\partial^2 Y_1^t/\partial t^2$ will contribute to S-shaped behavior.

The threshold model also applies in other cases where heterogeneity results from differences in land quality or human capital. For example, Caswell and Zilberman (1986) argue that modern irrigation technologies augment land quality, and predicted that drip and sprinkler irrigation will be adopted on lands where water-holding capacity is below a certain threshold. They also showed that adoption of these technologies by growers who rely on groundwater will be dependent on well depth. Akerlof's (1976) work on the "rat race" suggests that differences in human capital establish thresholds and result in differences in the adoption of different technologies and practices.

The threshold models shifted empirical emphasis from studies of diffusion to studies of the adoption behavior of individual farmers and a search for sources of heterogeneity. Two empirical approaches have been emphasized in the analysis of monthly cross-sectional data on technological choices and other choices of parameters and characteristics of individual firms. In the more popular approach, the dependent variables denote whether or not certain technologies are used by a farm product or unit at a certain period, and econometric techniques like logit or probit are used to explain discrete technology choices. The dependent variable for the second approach denotes the duration of technologies used by farms. (They answer the question, How many years ago did you adopt a specific technology?) Also, limited variable techniques are used to explain the technology data. Qualitatively, McWilliams and Zilberman (1996) found that the two approaches will provide similar answers, but analysis of duration data will enable a fuller depiction of the dynamics of diffusion.

2.1.4. Geographic considerations

Much of the social science literature on innovation emphasizes the role of distance and geography in technology adoption [Rogers (1962)]. Producers in locations farther away from a regional center are likely to adopt technologies later. This pattern is consistent with the findings of threshold models because initial learning and the establishment of a new technology may entail significant travel and transport costs, and these costs increase with distance.

Diamond's (1999) book on the evolution of human societies emphasizes the role of geography in the adoption of agricultural technologies. China and the Fertile Crescent have been source regions for some of the major crops and animals that have been domesticated by humans. Diamond argues that the use of domestic animals spread quickly throughout Asia and laid the foundation for the growth of the Euro-Asian civilizations that became dominant because most of these societies were at approximately the same geographic latitude, and there were many alternative routes that enabled movement of people across regions. The diffusion of crop and animal systems in Africa and the Americas was more problematic because population movement occurred along longitudinal routes (south to north) and thus, technologies required substantial adjustments to different climatic conditions in different latitudes. Diamond argues that there were other geographic barriers to the diffusion of agricultural technologies. For example, the slow evolution of agricultural societies in Australia and Papua New Guinea is explained by their distance from other societies, which prevented diffusion of practices from elsewhere.

Geography sets two barriers to adoption: climatic variability and distance. Investment in infrastructure to reduce transportation costs (e.g., roads and telephone lines) is likely to accelerate adoption. One reason for the faster rate of technological adoption in the United States is the emergence of a national media and the drastic reduction in the cost of access that resulted from the establishment of railroads, the interstate highway system, and rural electrification.

Distance is a major obstacle for adoption of technologies in developing countries. The impediment posed by distance is likely to decline with the spread of wireless communication technologies. It is a greater challenge to adopt technologies across different latitudes and varying ecological conditions. The establishment of international research centers that develop production and crop systems for specific conditions is one way to overcome this problem.

2.2. *Risk considerations*

The adoption of a new technology may expand the amount of risk associated with farming. Operators are uncertain about the properties and performance of a new technology, and these uncertainties interact with the random factors affecting agriculture. The number of risks associated with new technologies gives rise to several modeling approaches, each emphasizing aspects of the problem that are important for different types of innovations. In particular, some models will be appropriate for divisible technologies and others for lumpy ones, and some will explicitly emphasize dynamic aspects while others will be static in nature.

Much of the agricultural adoption literature was developed to explain adoption patterns of high-yield seed varieties (HYV), many of which were introduced as part of the "green revolution". Empirical studies established that these technologies were not fully adopted by farmers in the sense that farmers allocated only part of their land to HYV while continuing to allocate land to traditional technologies. Roumasset (1976) and others argued that risk considerations were crucial in explaining these diversifications, while having higher expected yield also tended to increase risk.

A useful approach to model choices associated with adoption of HYV is to use a static expected utility portfolio model to solve a discrete problem (whether or not to adopt the new technology at all); adoption can also be modeled as a continuous optimization problem in which optimal land shares devoted to new technologies and variable inputs are chosen, see [Just and Zilberman (1988), Feder and O'Mara (1981)].

To present these choices formally, consider a farmer with $\bar{L}$ acres of land, which can be allocated among two technologies. Let i be a technology variable, where $i = 0$ indicates the traditional technology, and $i = 1$ the modern one. Let the indicator variable be $\delta = 0$ when the modern technology is adopted (even if not adopted on all the land), and $\delta_1 = 0$ when the modern technology is not adopted. When $\delta = 0$, L_0 denotes land allocated to traditional technology and L_1 denotes land allocated to the new variety. The fixed cost associated with adoption of the new technology is $\bar{k}$ dollars. Profits per acre under the traditional and modern technologies are π_0 and π_1, respectively, and both are random variables. For convenience, assume that all the land is utilized when the traditional variety is used. Assume that the farmer is risk averse with a convex utility function $U(W)$ where W is wealth after operation and $W = W_0 + \Pi$ when W_0 is the initial wealth level and Π is the farmer's profit.

The optimal resource allocation problem of the farmer is

$$\max_{\substack{\delta=0,1\\ L_1,L_0}} \mathrm{EU}\big[W_0 + \underset{\substack{\uparrow\\ \left\{\begin{array}{c}\text{profits when modern}\\ \text{technology is adopted}\end{array}\right\}}}{\delta(\pi_0 L_0 + \pi_1 L_1 - k)} + (1-\delta)\pi_0 L\big] \quad \text{subject to } \underset{\substack{\uparrow\\ \left\{\begin{array}{c}\text{profits when adoption}\\ \text{does not occur}\end{array}\right\}}}{L_0 + L_1} \leqslant \overline{L}. \tag{21}$$

Just and Zilberman (1988) considered the case where the profits under both technologies are normally distributed, the expected value of profit per acre under technology i is m_i, the variance of profit per acre of technology i is σ_i^2, and the correlation of the per acre profits of the technologies is ρ. They demonstrated that when the modern technology is adopted ($\delta = 1$) on part of the land, but all of the land is utilized, the optimal land allocation to the modern technology (L_1^*) is approximated by the function $L_1^r(\overline{L})$. Formally,

$$L_1^* = L_1^r(\overline{L}) = \frac{\mathrm{E}(\Delta\pi)}{\phi r(\Delta\pi)} + R\overline{L}, \tag{22}$$

where $\mathrm{E}(\Delta\pi) = m_1 - m_0$ is the difference in expected profits per acre between the modern and traditional technology. $v(\Delta\pi) = v(\pi_1 - \pi_0) = \sigma_0^2 + \sigma_1^2 - 2\rho\sigma_1\sigma_0$ is the variance of the difference of profit per acre of the two technologies. Further,

$$R = \frac{\sigma_0(\sigma_0 - \rho\sigma_1)}{v(\Delta\pi)} = \frac{1}{2}\cdot\frac{\partial v(\Delta\pi)}{\partial\sigma_0}\frac{\sigma_0}{v(\Delta\pi)}$$

is a measure of the responsiveness of $v(\Delta\pi)$ to changes in σ_0, and ϕ is the Arrow-Pratt measure of absolute risk aversion, dependent upon expected wealth.

Numerous adoption studies have addressed the case where the modern technology increased mean yield per acre, $\mathrm{E}(\Delta\pi) > 0$, and had high variance as compared to the traditional technology, $\sigma_1^2 > \sigma_0^2$. These assumptions will be used here. First, consider the case where profits under the traditional technology are not risky, ($\sigma_0^2 = 0$). From condition A, $L_1^* = \mathrm{E}(\Delta\pi)/\phi\sigma_1^2$, adoption does not depend directly on farm size (only indirectly, through the impact of I on risk aversion), and adoption is likely to increase as the expected gain from adoption $\mathrm{E}(\Delta\pi)$ increases and the risk of the modern technologies (σ_1^2) decreases.

When $\sigma_0^2 > 0$ and ϕ is constant, Equation (22) suggests that L_1^r is a linear function of farm size $\overline{L}$. The slope of L_1^r is equal to R, and assuming $\sigma_1^2 > \sigma_0^2$, $R = \sigma_0(\sigma_0 - \rho\sigma_1) - v(\Delta\pi)$ is smaller than one. When the profits of two technologies are highly correlated, $\rho > \sigma_0/\sigma_1$, $R < 0$, $\mathrm{d}L_1^R/\mathrm{d}\overline{L} < 0$, and acreage of the modern technology declines with farm size. This occurs because the marginal increase with acreage (variance of profits) is larger than the marginal increase of expected profits that slow the growth or even reduce (when $\rho > \sigma_0/\sigma$) the acreage of the modern and more risky technology of larger farms.

Assume now that absolute risk aversion is a function of farm size (a proxy of expected wealth) denoted by $\phi(L)$. In this case, Just and Zilberman showed that the marginal effect of increase on the area of the modern technology is

$$\frac{\mathrm{d}L_1^r}{\mathrm{d}\overline{L}} = \eta\frac{L_1^r}{\overline{L}} + (r - \eta)R,$$

where $\eta = -\phi'\overline{L}/\phi$ is the elasticity of absolute risk aversion and is assumed to be between 0 ($\eta = 0$ implies constant absolute risk aversion) and 1 ($\eta = 1$ implies constant relative risk aversion coefficient, $\phi(L) \cdot L =$ constant). In this more general case, L_1^r may be a nonlinear function of $\overline{L}$ and may have a negative slope in cases of high correlation and small η.

Optimal land allocation to the modern technology, L_1^*, is constrained to be between 0 and $\overline{L}$. Thus, it may be different than L_1^r defined in (22). In cases with small η (ϕ does not change much with $\overline{L}$), the increase in risk (variance of profits) with size is much greater than the increase in expected profit with size. When $\overline{L}$ is close to zero, $L_1^R > \overline{L}$ and thus where farm size is below a critical level, $\overline{L}_b$,[10] the modern technology should be fully adopted if it is optimal. From (22) the adoption of the modern technology is optimal if it pays for the extra investment it entails. Thus, farms below another critical size, $\overline{L}_a$, cannot pay for the modern technology and do not adopt it.

Figure 5 depicts some plausible relations between L_1^* and $\overline{L}$. The segment *Oabcd* depicts the behavior of L_1^* when $R > 0$ and $\overline{L}_b > \overline{L}_a$. If $\overline{L}_b > \overline{L}_a$ and $R < 0$, L_1^* is depicted by *Oabce*. If $\overline{L}_a > \overline{L}_b$, and $R > 0$, L_1^* is depicted by *Ogh* and if $\overline{L}_a > \overline{L}_b$ and $R < 0$, L_1^* is depicted by *Ogle*. In the last two cases, there is no full adoption of the modern technology.

Feder, Just, and Zilberman (1985) report the results of several studies that show that when adoption occurs, the full share of modern technologies declines with farm size among adopters. These findings are consistent with all the scenarios in Figure 5.

2.3. Mechanisms to address product performance and "fit risk"

Adopters of new technologies, especially if embodied in high capital costs that entail significant irreversible investment, face uncertainty with respect to the performance of

---

[10] At $\overline{L} = \overline{L}_b$, $L_1^r(\overline{L})_b = \overline{L}_b$.

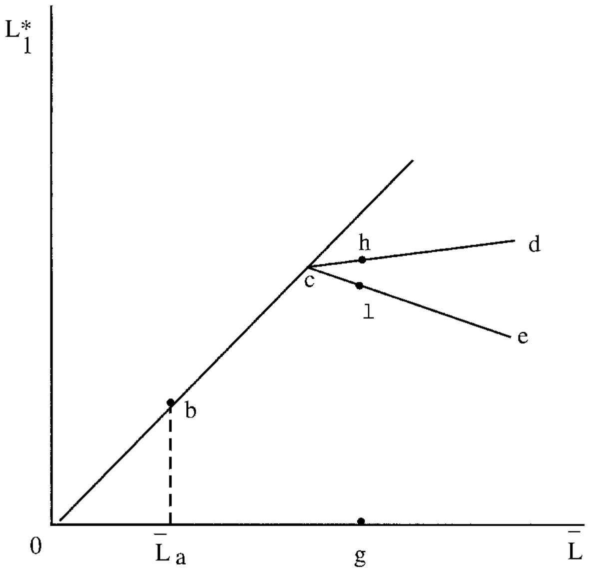

Figure 5.

the product, its reliability, and appropriateness of their operation. When a farmer buys a piece of machinery – be it a combine, harvester, seeder, or cultivator – and it has a breakdown or major malfunction, it may cost a farmer much of his revenues. Conceptually, one may think about several solutions to address some risk, including insurance. The prevailing approach to address such risk is to form a product-backup system. To address the financial risks that are associated with the repair cost of a broken or malfunctioning product, especially in the early life of the product, manufacturers introduced mechanisms such as warranties and established dealerships equipped to repair breakdowns. Thus, the combination of a warranty agreement and a well-functioning technical support system significantly reduces the amount of reliability risk associated with new products.

Significant elements of agribusinesses, such as mechanic shops, are devoted to the repair and maintenance of new capital equipment. The availability and quality of performance of this support will determine the risk farmers face in adoption decisions and, thus, their ability to carry risk. One of the main advantages of large farming operations is their in-house capacity to handle repairs, breakdowns, and maintenance of equipment. That makes them less dependent on local dealers and repair shops, and reduces their risk of having to purchase (in many cases) new products.

The value of the capacity to address problems of product equipment failure swiftly and efficiently is intensified by timing considerations. In many regions, harvesting seasons are short. Leaving a wheat crop unharvested for an extra day or two may expose it to damage due to rain, hail, or pests, thereby decreasing its yield. Market prices of perishable fruits and vegetables are significantly dependent on the timing of harvest;

a one-week delay in harvesting early season fruits or vegetables for shipping can reduce prices by factors of 30 to 40 percent [Parker and Zilberman (1993)]. This timing consideration increases the value of a well-functioning product system. It may provide an explanation for the maintenance of excess capacity to harvest or conduct other vital activities. Of course, the extent to which farmers maintain excess capacity depends on how well the product support system functions. The agricultural community may establish customs and other social and institutional arrangements for mutual help in a crisis situation associated with a breakdown of equipment.

Adoption of new technology entails risk with respect to its appropriateness to the farm and its performance. Results of prior testing by manufacturers represent performance and conditions that may not be exactly similar to those of farmers. New technologies may also require special skills and training. Institutional arrangements to reduce the risk associated with the adoption of new technologies have been introduced. They include product information and demonstration such as educational materials in various media formats as well as hands-on demonstrations. The farmer may go to a dealership to see farm machinery in operation or the equipment may be loaned to the farmer for a supervised and/or unsupervised trial period. For new seed varieties, manufacturers will send farmers samples of seeds for examination. Many farmers will plant small trial plots.

When university researchers are the providers of new seeds, extension plays a major role in demonstration. In the case of new seed varieties and equipment developed by the private sector, extension plays an important role in demonstrating efficacy in local conditions as well as making objective judgments on manufacturers' claims regarding new products.

In addition to various types of extension, the reduction of risk associated with performance and the appropriateness of new technologies is addressed by arrangements such as money-back guarantees. With money-back guarantees, the farmer is given the option to return the product. In this case, obviously the price of the product includes some payment for this option [Heiman et al. (1998)]. However, the money-back guarantee agreement allows farmers longer periods of experimentation with new products. Generally money-back guarantees are not complete and a fraction of the original cost is not returned.

Sometimes renting is used as a mechanism to reduce the risk associated with investment in new products. For example, when sprinkler irrigation was introduced in California, the main distributor of sprinklers in the state was a company called Rain for Rent. This company rented sprinkler equipment to farmers. Over time, the practice of renting sprinkler equipment became much less common and more new sprinkler equipment was purchased. In some cases, farmers use custom services for an initial trial with new technologies, and invest in the equipment only when they feel more secure and certain about its properties.

Many of the marketing strategies, including warranties, money-back guarantees, and demonstrations that are part of businesses throughout the economy, were introduced by agricultural firms including John Deere and International Harvesting. Currently, hundreds of millions of dollars are spent on promotion and education in the use of new products. Unfortunately, not much research has been conducted to understand this aspect of agricultural and technological change in agriculture. It seems, however, that a large body of empirical evidence regarding geographic concentration of new technologies and geographic patterns of technology adoption may be linked to considerations of marketing and product support efforts. New technologies are more likely to be adopted earlier near market centers where dealers and product supports are easily available. Agricultural industries and certain types of technologies may be clustered in certain regions, especially in the earlier life of a new technology, and these regions will generally be located in areas that have technical support and expertise associated with the maintenance and development of the technologies. It seems that considerations of marketing and geographic locations are two areas where more research should be done.

2.4. *Dynamic considerations*

The outcome of technology adoption is affected by dynamic processes that result in changes in prices of capital goods and input, learning by producers and users of capital goods, etc. Some of these processes have random components and significant uncertainty over time. Some of these dynamic considerations have been introduced to recent microlevel models of adoption behavior.

2.4.1. *Optimal timing of technology adoption*

The earlier discussion on threshold models recognized that timing of adoption may vary across production units reflecting differences in size, human capital, land quality, etc. The above analysis suggests that, at each moment, decision-makers select technologies with the best-expected net benefits (or expected net present values adjusted by risk). Thus, when a new technology is available decision-makers continuously evaluate whether or not to adopt; when the discounted expected benefits of adoption are greater than the cost, the technology will be adopted. This approach may lead to suboptimal outcomes because decision-makers do not consider the possibility of delaying the technology choice to take advantage of favorable dynamic processes or to enable further learning. These deficiencies have been corrected in recent models.

2.4.2. *Learning by using, learning by doing, and adoption of new technologies*

Consider a farmer who operates with a traditional technology and is considering adopting a new one that requires a fixed investment. The increase in temporal profit from adoption at time t increases as more experience is gained from the use of this technology. This gain in experience represents learning by doing. Let t_0 be the time of adoption and assume that self-experience is the only source of learning by doing. The increase in operational profits in $t > 0$ is $\Delta\pi(t - t_0)$, $\partial\pi/\partial t > 0$. Let the fixed cost of investment in firm t_0 be denoted by $K(t_0)$. The process of learning by using reduces the manufacturing cost of fixed assets and results in reduction in $K(t_0)$ over time. It is reasonable to assume that the effects of both learning processes decline over time. Thus,

$$\frac{\partial^2 \pi(t-t_0)}{\partial t} < 0, \frac{\partial K(t_0)}{\partial (t_0)} < 0, \frac{\partial^2 K(t_0)}{\partial t^2} > 0.$$

When the farmer disregards the learning processes in determining the time of adoption, adoption will occur when the temporal gain of adoption equals the extra periodical fixed cost. Let r denote discount rate and assume the economic life of the new technology is infinite. At t_0,

$$\Delta\pi(0) = rK(t_0).$$

When the learning processes are taken into account, the marginal reduction in investment cost, because of learning by using, tends to delay adoption, and the marginal benefits from learning by using may accelerate the time of adoption. The optimal conditions that determine t_0 in this more general case are

$$\underset{\substack{\text{Extra profit}\\ \text{from adoption}}}{\overset{(+)}{\Delta\Pi(0)}} - \underset{\substack{\text{Investment}\\ \text{cost}}}{\overset{(-)}{rK(t_0)}} + \underset{\substack{\text{Learning by}\\ \text{doing effect}}}{\overset{(-)}{\frac{\partial K(t_0)}{\partial t_0}}} + \underset{\substack{\text{Learning by}\\ \text{using effect}}}{\overset{(+)}{\int_0^\infty e^{-rt}\frac{\partial \pi(t)}{\partial t}\,\mathrm{d}t}} = 0.$$

In cases where the new technology increases the productivity of an agricultural crop with constant returns to scale,

$$\Delta\Pi(t-t_0) = \Delta\pi(t-t_0)\cdot L,$$

where L is acreage. In this case, both the extra profit from adoption and the learning-by-using effects will increase with farm size and lead larger farms to be early adopters. Higher interest rates will tend to retard adoption because they will increase the investment cost per period and reduce the learning-by-using effect.

2.4.3. Adoption under irreversibility and uncertainty

Adoption sometimes entails irreversible investments with uncertain payoffs. Delay of an adoption decision may enable the producer to obtain more information, reducing overall uncertainty, and increasing expected discounted benefits by avoiding irreversible investment when it is not worthwhile. This observation can be illustrated by the following example that analyzes adoption decisions in a simple, two-period model.

The adoption decision requires an initial investment of \$100. The returns from adoption consist of \$50 at the initial period, \$30 with probability of .5 (low returns case), and \$150 with probability of .5 (high returns case) in the second period. Let r be the discount rate. According to the neoclassical investment theory, adoption should occur at the initial period of the expected net benefit of this decision, and ($ENPV_0$) is positive when

$$ENPV_0 = 50 + \frac{1}{1+r}[0.5 \cdot 30 + 0.5 \cdot 150] - 100 = \frac{90}{1+r} - 50.$$

The standard expected net present value criteria will suggest adoption in the initial period when the discount rate is smaller than 0.8 (since $ENPV_0 > 0$ when $90/(1+r) > 50$ for $r < 0.8$). However, the farmer's set of choices includes an option to wait until the second period and adopt only in the case of high returns. The investment associated with adoption is irreversible, and waiting to observe the returns in the second period enables avoiding investment in the case of low returns. The expected net present value with this approach is

$$ENPV_1 = 0.5 + \frac{(150 - 100)}{1+r} = \frac{25}{1+r} > 0.$$

When $r = 0.5$, $ENPV_0 = 90/1.5 - 50 = 10$, $ENPV_1 = 25/1.5 = 16\frac{2}{3}$, then the "wait and see" approach is optimal. This approach removes the downside risk of the low-return case in the second period. The value added by waiting and retaining flexibility in light of new information is called "option value" and in this example is defined below as follows:

$$0V = \max[NPV_1 - NPV_0, 0] = \max\left[50 - \frac{65}{1+r}, 0\right].$$

In the case of $r = 0.5$, the option value is 6–2/3 and waiting to see the outcome of the second period is optimal. In the case with $r < 0.3$, the option value is 0 and adoption in the initial period is optimal.

This example is a simple illustration of a more complex, multi-period model of adoption. Suppose a farmer employs two technologies, traditional and modern. The temporal profit from each of the technologies depends on a random variable, S_t. This may be the price of output or input, or it may be the value of a physical variable (climatic condition) that affects profitability. The modern technology usually generates more profits but requires a fixed investment. Let the difference in temporal profit between the two technologies in period t be $\Delta\Pi(S_t) = \Pi_1(S_t) - \Pi_0(S_t)$. Assume that the temporal gain from adoption increases with $S_t (\partial\Delta\Pi/\partial S_t > 0)$. Let the cost of the investment in the new technology be denoted by K, and the discount rate be denoted by r. The farmer has to determine when to adopt the modern technology. Let T be the period of adoption. The farmer's optimization problem is

$$\max_{\substack{T\\ [T=0,1,\ldots,\infty]}} \sum_{i=T}^{\infty} \mathrm{E}_{S_t}\left[\frac{\Delta\pi(S_t)}{(1+r)^t}\right] - \frac{K}{(1+r)^T},$$

where $\mathrm{E}_{S_t}(\cdot)$ denotes expectation with respect to S_t. The nature of the solution depends on the assumption regarding the evolution of the sequence of random variables S_t. For example, suppose $S_t = S_{t-1} + \varepsilon_t$ where all the ε_t's are independently and identically distributed random variables whose means are zero. (If they are normally distributed, S_t is generated by a "random walk" process.) This approach has been very successful in the analysis of options in finance, and Dixit and Pindyck (1994) and McDonald and Siegel (1986) applied it to the analysis of capital investments. They viewed investments with unrestricted timing as "real options" since the decision about when to undertake an investment is equivalent to the decision about when to exercise an option. McDonald and Siegel (1986) considered a continuous time model to determine the time of investment. They assumed that the S evolves according to a Wiener process (which is a differential continuous version of the process described above) and used the Ito calculus to obtain formulas to determine the threshold for adoption, $\overline{S}$. Their analysis suggests that the threshold level of $\overline{S}$ increases as the variance of the temporal random variable ε_t increases.

Their framework was applied by Hasset and Metcalf (1992) to assess adoption of energy conservation in the residential sectors. Thurow, Boggess, and Moss (1997) applied the real option approach to assess how uncertainty and irreversibility considerations will affect adoption of free-stall dairy housing, a technology that increases productivity and reduces pollution. The source of uncertainty in their case is future environmental regulation. Using simulation techniques, they showed that when investment is optimal under the real option approach, expected annual returns are more than twice the expected annual returns associated with adoption under the traditional net present value approach. Thus, the real value approach may lead to a significant delay in adoption of the free-stall housing and occurs when pollution regulations are very stiff.

Olmstead (1998) applied the real value approach to assess adoption of modern irrigation technology when water prices and availability are uncertain. Her simulation suggests that the water price leading to adoption under the real option approach is 133 percent higher than the price that triggers adoption under the standard expected net present value approach. In her simulation, the average delay in adoption associated with the real option approach is longer than 12 years.

There have been significant studies of adoption of irrigation technologies and, while adoption levels seemed to respond significantly to economic incentives, adoption did not occur in many of the circumstances when it was deemed to be optimal using the expected present value criteria. Much of the adoption occurs during drought periods when water prices escalate drastically [Zilberman et al. (1994)]. The option value approach provides a good explanation of the prevalence of adoption during crisis situations.

The analysis of adoption behavior using "real options" models holds much promise and is likely to be expanded. In many cases, not all the adoption investment is "sunk cost". Some of it can be recovered. For example, capital goods may be resold, and added human capital may increase earning opportunities. The delay caused by adoption costs and uncertainties will likely be shorter if these costs are more recoverable, and institutions that reduce irreversibilities (rental of capital equipment, money-back guarantee agreements) are apt to increase and accelerate adoption.

The real option approach provides new insight and is very elegant, but it does not capture important aspects of the dynamics of adoption. It assumes that decision-makers know the distribution of random events that determine profitability when it is more likely that a learning process is going on throughout the adoption process, and adopters adjust their probability estimates as they go along. Furthermore, while adoption requires a fixed initial investment, it also may entail incremental investments, especially when the intensity of use of a new technology changes over time. Thus, a more complete dynamic framework for analyzing adoption should address issues of timing, learning, and sequential investment. Some scholars [Chavas (1993)] have introduced models that incorporate these features, but this research direction requires more conceptual and empirical work.

2.4.4. The Cochrane treadmill

A key issue in the economics of innovation and adoption is to understand the impact of technology change on prices and, in particular, the well-being of the farm population over time. When a supply-increasing innovation is adopted to a significant degree, it will lead to reduction in output prices, especially in agricultural commodities with low elasticity of demand. When it comes to adoption of a new technology, Cochrane (1979) divided the farming population into three subgroups – early adopters, followers, and laggards. The early adopters may be a small fraction of the population, in which case the impact of their adoption decision on aggregate supply and, thus, output prices is relatively small. Therefore, these individuals stand to profit from the innovation.

The followers are the large share of the farm sector who tend to adopt during the take-off stage of the innovation. Their adoption choice will eventually tend to reduce prices, which reduces profits as well. This group of adopters may gain or lose as a result of innovation.

Finally, the laggards (the third group) are the farmers who either adopt at the lag stage of the adoption process or do not adopt at all. These individuals may lose from technological change. If they do not adopt, they produce the same quantity as before, at low prices; and if they adopt, the significant price effect may sweep the gain associated with higher yields. Thus, Cochrane argues that farmers, on the whole, are not likely to gain from the introduction of innovation in agriculture, except for a small group of early adopters. Introduction of new technology may lead to structural change and worsen the lot of some of the small farms. The real gainers from technological change and innovation in agriculture are likely to be consumers, who pay less for their food bill.

Kislev and Schori-Bachrach (1973) developed conceptual and empirical models based on Cochrane's analysis using data from Israel. They show that small subgroups of farmers are the early innovators who adopt the new technologies. When there is a wave of new technologies, these individuals, who have a higher education and other indicators of human capital, will consistently be able to take advantage of technology change and profit. The rest of the farming population does not do as well from technological change.

The Cochrane results are modified in situations where agricultural commodities face perfectly elastic demand, for example, when adopting industry export goods from a small country. In this case, the impact of increased profitability associated with the introduction of a new technology will lead to an increase in land rents which may occur some time after the innovation was introduced. Thus, the early adopters, even if they are farm operators, may be able to make an above-normal profit as a result of their adoption decision, but most of the followers will not gain much from the adoption decision because their higher revenues will be reduced by an increase in rent. Laggards and nonadopters may lose because the higher rent may reduce their profits. Again, landowners will be gainers from the innovations, and not farmers who own no land. Thus, this extension of Cochrane's model reaches the same conclusion-that at least some farmers do not benefit from technological change as much as other agents in the population.

Cochrane's modeling framework was used to argue that, in spite of the high technological change that occurs in agriculture and its dynamic nature, farmers may not be better off and actually some of them may be worse off from innovations. That may justify the "farm problem" that occurred in much of the twentieth century where the well-being of farmers became worse relative to other sectors of the population. Cochrane's basic framework was not introduced formally. Zilberman (1985) introduced the dynamics of the threshold model of adoption that identified conditions under which the quasi-rents of farmers decline over time. His model did not take into account the changes in structure that may be associated with innovation agriculture. When innovations are embodied in technology packages that are both yield-increasing (high-yield varieties) and labor-saving (tractors and other machinery), and agricultural demand is inelastic, then technological change will reduce quasi rent per acre and make operations in the farm sector less appealing to a large segment of the population. Thus the early adopters are likely to accumulate more of the land, increasing their farm size. Over time, structural change will result in a relatively small farm sector, and earnings per farm may actually increase as farms become much bigger. Gardner's (1988) findings show that, in relative terms, the farm population is now as well off or even better off than the nonfarm population, especially in the United States. His findings are consistent with the process of technological change that led to the accumulation of resources by small subgroups of the farm population while the rest migrated to the urban sector where earnings were better. But in addition to the gains from technological change, the adopters may also have benefited from a commodity program that slowed the decline in prices as well as the processes of globalization that makes demand more elastic over time.

A more formal and complete understanding of the distribution and price implications of technological change over time is a challenge for further research on the economics of technology adoption. Stoneman and Ireland (1983) argue that firms producing the components of new technology recognize the dynamics of adoption; they design their production and establish technology component prices accordingly, taking advantage of the monopolistic power. Thus there is a clear linkage between the economics of innovation and adoption that should be investigated further. An understanding of these links is essential for the design of better patent policy and public research strategies.

2.5. *Institutional constraints to innovation*

While agricultural industries tend to be competitive, the perfectly competitive model does not necessarily apply since farmers may face a significant number of institutional constraints and policies which affect their behavior significantly and result in outcomes that are different from those predicted by the perfectly competitive model. This institutional constraint may be especially important in the area of technological change and adoption. Some of the most important constraints relate to credit as well as tenure relationships, as addressed below. Note that institutional constraints may affect the patterns of adoption of new technologies, but on the other hand, the introduction of new technologies may affect the institutional structure and operation of agricultural industries. We will concentrate on the first problem but will address both.

2.5.1. *Credit*

Asymmetric information between lenders and borrowers and the uncertain conditions in agriculture and financial markets have led to imperfections in the credit market, most notably credit constraints that affect adoption behavior [Hoff et al. (1993)]. In many cases, farmers use some of their own equity to finance at least part of their investments. In other cases, assets such as land or the crop itself are used as collateral for financing a new technology. The exact formulation of the credit constraint faced by farmers is quite tricky, but it is not unreasonable to approximate as a linear function of acreage. The reason is that, in many cases, land is the major asset of a farming operation.

Just and Zilberman (1983) introduced a credit constraint in their static model of adoption under uncertainty. They assume that investment in the new technology is equal to $k + \alpha L_1$ when α is investment per acre in the modern technology. The constraint on credit per acre is m dollars. Thus, the farm credit constraint is $m\overline{L} \geqslant k + \alpha L_1$. If $m < \alpha_1$, there will be full adoption. However, if $m > \alpha$, the credit constraint will not bind for larger farms. Figure 6 depicts some plausible outcomes for the second case. Consider a case where $R > 0$ and $\overline{L}_a < \overline{L}_b$. Without the credit constraints, optimal allocation of land to the modern technology, as a function of farm size, is depicted by *0abcd* in Figure 6. There may be several scenarios under the credit constraints.

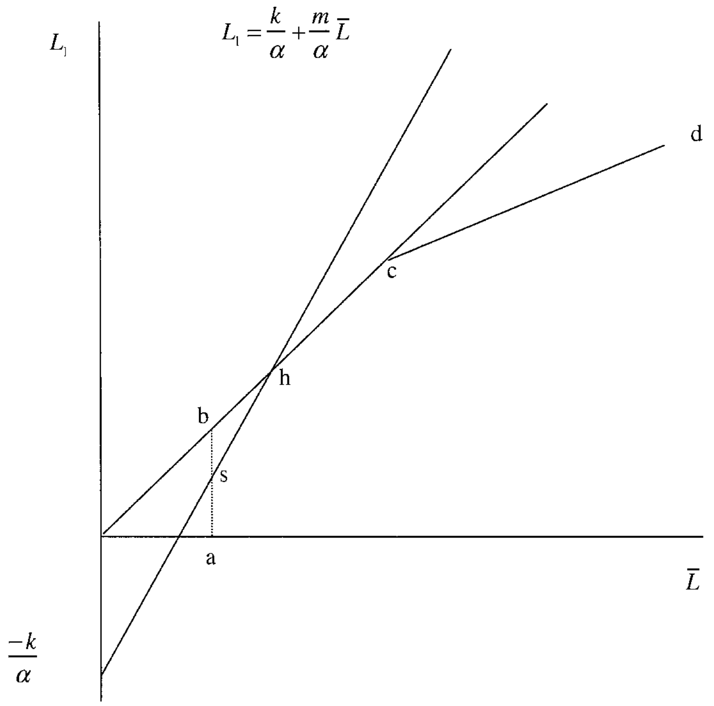

Figure 6.

In terms of Figure 6, when credit is a binding constraint, $L_1 < -\frac{k}{\alpha} + \frac{m}{\alpha}\bar{L}$. Small farms (with sizes in the range *0a*) will be non-adopters. Somewhat larger farms, in the range *bh*, will be credit-constrained partial adopters. Even larger farms (in the range *hc*) will specialize in the new technology, and farms of the largest size (corresponding to *cd*) will be risk diversifiers when *m* is smaller. Policies to remove credit constraints will be beneficial, especially to smaller farmers, and will enable some to adopt and others to extend their intensity of adoption.

The credit constraints per acre may be affected by the lender's perception of the profitability of agriculture (and farmland prices that reflect the profitability). Initial subsidization of credit early in the diffusion process that will enhance adoption will provide evidence that may change (improve in the case of a valuable technology) the lender's perception of the profitability of the industry and the modern technology, and lead to a relaxation of credit constraints. It will thus facilitate further adoption.

The interest rate and other financial charges may be differentiated according to size. Banks may perceive smaller farms to be more risky, so they may need to compensate for the fixed cost of loan processing, etc.[11] If the price of credit is higher for smaller

---

[11] There is significant evidence in the development literature that smaller operators face a higher interest cost.

farms, that extra hurdle will reduce the minimal farm size that is required for new technology adoption and will slow adoption by smaller-sized farms. Thus, advantageous credit conditions may be another reason larger farms adopt new technologies earlier. The reduction of institutions such as the Grameen Bank in Bangladesh and organizations such as the Bank of America in the United States, which in the beginning of the century facilitated loans to smaller operations, may be a crucial element in accelerating the process of technological change in improving adoption.

The financial crisis of the 1980s has led to a realization of the significance of risk associated with emphasizing collateral considerations in loan generation. The value of assets such as land is highly correlated to the profitability of agriculture and, in periods of crises and bankruptcies, land will be less valuable as collateral. That will lead to an increased emphasis on "ability to pay" as a criterion for loan generation. Thus, farmers need to provide sufficient guarantees about the profitability of their investment and their future ability to repay a loan. This may put investment in new technologies at a disadvantage because many of them do not have a sufficient track record that will assure banks of their economic viability. Banks may lack the personnel that are able to correctly assess new technologies and their economic value [Agricultural Issues Center (1994)].

One approach to overcoming this obstacle is by credit subsidies for a new technology, which may be appropriate in situations when investments generate positive externalities. However, an alternative and more prevalent solution is the provision of finance or a loan guarantee by the input manufacturer that leads to a reduction of the financial constraints on farmers. Furthermore, it reduces the fixed cost of adoption since it reduces the cost of searching for a loan. (One of the major implications of restricted availability of credit is the higher cost of finance, even for people who eventually obtain the credit.) Indeed, some of the major automobile and heavy equipment companies have their own subsidiaries or contractual arrangements that provide financing for new purchases of equipment, and seed companies often play an important role in the provision of credit. In many cases farmers may obtain loans for credit provisions through cooperatives or government policies (see chapter on credit).

2.5.2. Tenure

There is a distinct separation between ownership and the operation of agricultural land throughout the world. About 50 percent of the farmland in the United States is operated by individuals who do not own the land, and the financial arrangements between owners and operators vary. In the development literature, there is a significant emphasis on the importance of tenure systems on technology adoption. Most of the literature takes tenure as given and assesses its impact on adoption of technologies. However, this impact depends on the arrangements as well as the nature of the technology. Furthermore, as we will argue later, the introduction of new technologies may result in new tenure relationships.

The simplest relationships are land rent contracts where operators pay a fixed rent to landowners. Several factors will determine how these contracts affect adoption behavior. In the case of short-term contracts, when operators are not secure in maintaining the same land for a long time, the likelihood that they will adopt a technology that requires investment in the physical infrastructure and improvement of the land is very low. In these cases, rental relationships may be a significant deterrent for the adoption of innovations. On the other hand, the fixed-rate rent will not be a major deterrent of adoption if the innovation does not require a significant modification of the physical infrastructure, or if it augments or is dependent upon the human or physical capital of the operator. For example, an operator may purchase a tractor to reduce the cost of his operation. The necessary condition for adoption in this case is that the operator rent a sufficient amount of land every year in order to recapture and repay the investment. Actually, in some cases, the existence of a well-functioning land rental market may accelerate adoption of technologies that require a significant scale of operation. In fact, some farmers may augment the land utilized by them by renting land from others, thus enabling them to adopt large equipment. This was the situation, for example, in California when the cotton harvester was introduced. Therefore, it is useful to distinguish between large operators who use rental agreements to increase the acreage under their control (the rental agreements may facilitate adoption) and small operators without land of their own. For these operators, due to the credit constraints, lack of land may be a deterrent for adoption, even for technologies that do not improve the land and related assets.

2.5.3. *Complementary inputs and infrastructure*

The introduction of new technologies may increase demand for complementary inputs and when the supply of these inputs is restricted, adoption will be constrained. High-yield "green revolution" varieties require increased water and fertilizer use. McGuirk and Mundlak's (1991) analysis of the adoption of high-yield varieties in the Punjab showed that adoption was constrained by the availability of water and fertilizer. Private investment in the drilling of wells, and private and public investment in the establishment of fertilizer production and supply facilities removed these constraints and contributed to the diffusion of modern wheat and rice varieties in the Punjab. The adoption of high-yield maize varieties in the Punjab was much lower than wheat and rice, mostly because of disease problems. Adoption rates in maize might have been higher if complementary disease-control technologies were available.

Some of the complementary input constraints are eased or eliminated with the appropriate infrastructure. Effective research and extension programs may devise solutions to pest problems thus enabling the adoption of vulnerable varieties. Some of the modeling and analysis of diffusion [Mahajan and Peterson (1985)] suggests that the diffusion rates in regions that are farther from commercial centers are lower. To some extent this reflects barriers for professional support and more limited and costly access to complementary inputs. Improvement in transportation infrastructure may thus be useful for enhancing adoption.

2.6. *Adoption and farm policy*

Agriculture in developing and developed countries has been subject to government interventions that, in turn, affect technological change. Generally speaking, agricultural policies in developed countries aim to raise and stabilize agricultural incomes and, in some cases, to curtail supplies, while agricultural outputs have been taxed in developing nations. In both cases agricultural inputs have tended to be subsidized.

In recent years, agriculture has been subject to new environmental policies that control and affect the use of certain inputs that may cause pollution. The following is a discussion of the impacts of different policies on technological change.

2.6.1. *Price supports*

Just, Rausser, and Zilberman (1986) and Just et al. (1988) developed a framework, relying on the model presented in Equation (21), to analyze the impact of agricultural policies on technology adoption for farmers operating under uncertainty. They analyze various policies by tracing their impacts on price distributions of inputs and outputs as well as constraints (i.e., credit) on adoption. Price supports increase the mean of prices received by farmers and reduce their variability by setting lower price bounds. When the new technology has a yield-increasing effect (for example, high-yield variety), and if it is also perceived to have higher risk, price-support policies tend to increase its relative profitability, which leads to an increase in both the extent and intensity of adoption. McGuirk and Mundlak (1991) argue that the introduction of guaranteed markets for Punjabi food grain production by the government procurement policy (which was in essence a price support policy) enhanced the adoption of high-yield wheat and rice varieties in this region.

The mechanism through which price supports impact the adoption behavior of farms of different sizes varies. Smaller farms may increase their adoption because of price supports (their impact on credit) and the reduction in the minimum size required to justify adoption. Larger farms that may be risk diversifiers will increase the share of modern technologies on their land because of the mean effect and the reduction in risk. Price supports may also enhance adoption of mechanical innovations when they increase the relative profitability of operations with a new technology and thus reduce the size threshold required for adoption. Price supports may enhance adoption also through their impact on credit. When the ability to obtain credit depends on expected incomes, price supports will increase adoption when credit is constrained.

2.6.2. *Combined output price supports and land diversion policies*

In the United States as well as in some European countries, the subsidization of prices has been accompanied by a conditional reduction in acreage. The higher and most secure prices on at least part of the land provide incentives for farmers to adopt yield-increasing varieties. On these lands, they raise the value of property and expected income, which increases their capacity to obtain credit that may enhance the adoption of all types of technologies.

Specific elements of the support program vary over time. In recent years, the base for support has not been the actual yield, but the average base yield that is dependent on the average past performance of either the farmer or the region. The acreage that provides the base for entitlement to the benefits of a diversion program also depends on past performance. According to the specifics of a program, farmers might expand their yield or acreage in order to expand their entitlements in the future. Thus, adoption of high-yielding technologies, or technologies that may be especially beneficial with marginal land, is more likely to occur with price supports/diversion policies.[12] The historical record that provides a base for future program entitlements may, on the other hand, provide disincentives to adopt new crops or to introduce nonprogram crops to certain areas and thus reduce the flexibility of farming. The 1996 Farm Act in the United States makes entitlements that are independent of most farming activities, including choice of crop. However, even under this bill, land that is entitled to income support is somewhat restricted in its choice of crops, and that may retard the adoption and introduction of new crops to some of the major field crop regions of the United States.

Cochrane (1979) argued that the commodity programs in the United States played a major role in the adoption of mechanical and chemical innovations by reducing risk and increasing profitability per acre. The commodity programs as well as the increases in demand and prices during and after World War II led to modernization and structural change in U.S. agriculture. De Gorter and Fisher (1993) used a dynamic model to show that the combination of price supports and land diversion led to intensification of farming in the United States. Lichtenberg's (1989) work demonstrates the importance of economic incentives for the adoption of center-pivot irrigation in Nebraska and other Midwestern states, and suggests that expansion of the irrigated land base in these states benefited from the support programs of the 1970s and 1980s.

2.6.3. Output taxation

Taxation of agricultural outputs, prevalent especially in developing countries, has a disastrous effect on technological change. It reduces the incentive to adopt yield-increasing technologies, increases the scale of operation that justifies financing purchases of new equipment, and depresses the price of agricultural land, thus reducing the ability to borrow. Furthermore, with lower prices, there are incentives to apply intensively modern inputs, which are associated in many cases with the adoption of modern, high-yield varieties in developing countries. The low growth of Argentinian agriculture between 1940 and 1973 is a result of output taxation and other policies that reduced relative prices of agricultural products and slowed investments and technological change in this sector [Cavallo and Mundlak (1982)].

---

[12] The work of Zilberman (1984) provides a rigorous argument on the impact of programs such as deficiency payments and diversion policies on the expansion of acreage and supply.

2.6.4. *Trade liberalization and macroeconomic policies*

The adoption of innovations is likely to be significantly influenced by policies that affect the general economy. This may include trade and exchange rate policies as well as macroeconomic and credit policies. Macroeconomic policies that lead to high interest rates may reduce adoption because investment in new technologies is more costly. Adoption of mechanical innovations may suffer more significantly with high interest rates, while farmers may switch to technologies that are labor-intensive.

Changes in international trade regimes will affect various regions differently according to their relative advantage. The opening of markets in the United States led to the introduction of high-value varieties in different communities in Central America [Carletto et al. (1996)]. This change in cropping was combined with the establishment of a new infrastructure and the construction of packinghouses and transportation facilities. Thus, when a change in trade rules seems permanent, it may lead to a complete overhaul of the infrastructure, and that may enable adoption of new crops and modernization.

Favorable pricing because of trade barriers enables growers in Europe, Japan, and some parts of the United States to adopt yield-increasing varieties, to invest and develop greenhouse technologies, and to expand the capacity of different technologies, including irrigated agriculture, in situations that would not have warranted it under free trade. The growth and investment in the agricultural sector in both Argentina and Chile suffered during periods when international trade was constrained, and benefited from trade liberalization [Coeymans and Mundlak (1993), Cavallo and Mundlak (1982)].

2.6.5. *Environmental policies*

A wide array of environmental regulations affects technologies available for agriculture. Pesticide bans provide a strong incentive for the development of alternatives at the manufacturer level and for the adoption of alternative strategies including nonchemical treatment, biological control, etc. On the other hand, the lack of availability of chemicals may retard adoption of high-yield varieties or new crops that are susceptible to a particular pest, especially in cases where nonchemical alternatives are not very effective. The elimination of DBCP (with its unique capacity to treat soil-borne diseases) in the mid-1980s in California led, on the one hand, to the abandonment of some grape acreage and a switch to other crops. At the same time, it enhanced the adoption of drip irrigation that enabled applications of alternatives in other areas.

2.6.6. *Input subsidies*

There is a wide body of literature [Caswell (1991)] that shows that subsidized water pricing tends to retard the adoption of modern irrigation technologies. However, subsidized input led to the adoption of high-yield varieties and "green revolution" technologies in countries like India. They also increased profitability and thus have an indirect positive impact on adoption through credit effects. Similarly, subsidization of pesticides and fertilizers led to the adoption of high-yield varieties and chemical-intensive technologies in developing and developed countries alike, which is also likely to result in problems of environmental pollution since the environmental side effects of agriculture are often the result of excessive residues. Alternatively, elimination of subsidies and especially taxation of chemical inputs may lead to adoption of more precise application technologies that will reduce residues and actually may increase yield [Khanna and Zilberman (1997)].

2.6.7. Conditional entitlements of environmental programs

Governments have recognized that they can use entitlements to support programs conditional on certain patterns of behavior. Therefore, in recent years there have been attempts to link entitlements to income supports, policies, and other subsidization to certain patterns of environmental behavior. A program like the Environmental Quality Incentives Program (EQIP) in the United States attempts to induce farmers to adopt practices such as low-tillage and soil testing, and to reduce the application of chemicals in exchange for entitlements for some support. In some cases, the benefits of such a program are short-lived and farmers may quit using modern practices once the program benefits disappear. On the other hand, especially when it comes to new, untested technologies, elements of learning-by-doing and experience may improve the profitability of those technologies that have some environmental benefits so that farmers recognize their economic advantages. Thus, the adoption of such technologies may persist in the long run.

3. Future directious

Research on agricultural technology evolves from the technology and the institutions associated with it. At present, agriculture is undergoing a technological revolution as evidenced by the introduction of biotechnology and precision technology. We are also witnessing related processes of industrialization, product differentiation, and increased vertical integration in agriculture [Zilberman et al. (1997)]. These changes raise new issues and introduce new challenges. Several significant changes have been observed thus far from the emergence of biotechnology [Zilberman et al. (1998)].

With many past technologies, university research identified some of the basic concepts while most of the innovations were done in industries. However, with biotechnology, universities are the source of numerous new discoveries, and technology transfer from universities to industries has triggered the creation of leading products and companies. The unwillingness of private firms to develop university innovations without exclusive rights motivated the establishment of offices of technology transfer that identified buyers who would share the rights to develop university innovations. Each arrangement provides new sources of funding to universities since royalties are divided among universities, researchers, and departments. Thus far, income from technology transfer revenues has paid less than 5 percent of university research budgets. However, in some areas (biology and medicine) it made a difference. Most of the royalties were associated with fewer than 10 innovations [Parker et al. (1998)], reinforcing our existing knowledge that benefits to research tend to concentrate on a small number of critical innovations. Established companies were not willing to buy the rights to develop some of the most radical, yet important, university innovations and biotechnology. Thus, offices of technology transfer, working with venture capitalists, helped to establish new upstart companies, some of which became leading biotechnology firms (e.g., Genentech, Chiron and Amgen). As these companies grew and became successful, some of the major multinationals bought a majority of shares in these companies. Thus, most of the activities in biotechnology have been in medical biotechnology. However, 1996 was the breakthrough year for agricultural biotechnology as millions of acres were planted with pest-resistant varieties of cotton and soybeans. In agricultural biotechnology, we see again the importance of small startups from the collaborations between university researchers and venture capitalists. Most of the startups in agricultural biotechnology have been acquired by giants like Monsanto and DuPont.

The evolution of biotechnology suggests that the university is becoming a major player in industrial development, and it affects the structure and competitiveness of industries. University researchers working with venture capitalists generate new avenues of product development. Sometimes they may force some of the giant companies to change their product development strategy, and may even give up some of their monopolistic power. Other forms of contractual relationships between university researchers and industries are being established. For example, industries support certain lines of research for an exclusive option to purchase the rights for technology. Furthermore, some researchers suddenly find themselves wearing another hat, that of a partner in a technology company, and that may affect the way universities run their patterns of payments and support for researchers. Given these new realities, there is a need for both empirical and conceptual research on innovations and the relationships between public and private research. We need to better understand the existing arrangements of royalties, sharing of royalties within the university, the relationship between publications and patents, and the effect of university research and industrial structure, etc.

With computers, biotechnology, and other new technologies, most of the value is now embodied in specific knowledge. The Cohen–Boyer patent once generated the largest revenues to universities. In this case, companies paid for the right to use a process for genetic manipulation. The key to biotechnology is the process of innovation (which specifies how to conduct specific manipulation) and product innovation (which specifies what type of outcomes can be controlled by which genes). New genetic engineering products will be produced by combining certain procedures and items of knowledge that are protected by certain patent rights. In principle, the developers of new products should pay the royalties to whoever owns the patents. Thus the markets for rights to different types of knowledge will emerge.

A new research agenda is suggested to address the economics of intellectual property rights. In particular, it should address pricing rules for different types of intellectual property rights and the design of biotechnology products given the price structures for different processes and product innovations. An important area of understanding is the pricing of international property rights within complex international systems where protection of intellectual property rights is not always feasible and where there are significant disparities in income.

The research in intellectual property rights will also have implications on the issues of biodiversity and compensation for developing countries for genetic materials that are embodied in their natural resources. Other related issues include the incentives for and integration of research to develop basic foods; the alleviation of starvation in the poorest countries; how new emerging industrial orders in agriculture and biotechnology can provide appropriate technologies to these countries; defining the role of international research institutes and other public entities (e.g., the United Nations and global organizations) in conducting research aimed at the poorest countries; and what type of payment arrangement should exist between research units focused on developing countries and commercial firms in the more developed nations.

Materials and chemicals that were previously produced by chemical procedures may be produced through modified biological organisms. First, biotechnology in agriculture will produce alternative forms of pest control and pest-resistant varieties but, over time, it will produce higher quality food products and new products such as pharmaceuticals and fine chemicals, see [Zilberman et al. (1997)]. With biotechnology the value added of seeds will increase to include some of the rent that was accrued to chemicals. Pesticide manufacturers already have become major players in biotechnology and are taking over seed companies in order to obtain a channel to market their products. Often the owners of the rights to patents try to capture some of the rent through contracting; thus, biotechnology will provide both the incentives to enhance contractual arrangements and vertical integration in agriculture. Some of the recent mergers and acquisitions in agricultural biotechnology can be explained by attempts to obtain rights to intellectual property and access to markets [Rausser et al. (1999)]. Finally, biotechnology causes firms with agricultural characteristics (for example, dairies, livestock operations, and even field crop operations) to produce products in areas that are not traditionally agricultural (pharmaceutical, oils, coloring). As the borderline between agriculture and industry becomes fuzzier, new models replace the competitive models as the major paradigm to assess agriculture.

The new product lines and the new types of industrial organization that may occur with biotechnology will raise environmental concerns and management issues. Biotechnology, thus far, has had a good track record, but it could have a negative potential. The design of the regulatory framework will significantly affect the structure of biotechnology industries and their impact. A more restrictive registration process, for example, may lead to a more concentrated biotechnology. Thus, it will become a research and policy challenge to modify the registration process and to balance the risks and benefits associated with biotechnology through monitoring over time. The optimal design of intellectual property rights agreements in biotechnology will become another issue of major concern. Patent rights that are too broad will lead to concentration in industries. It may stymie competition but may encourage a small number of firms to invest heavily in new products. Biotechnology patent protections that are too narrow may prevent significant investment in a costly research line.

Over the last 30 or 40 years, precision technologies have evolved that adjust input use to variation over space and time and reduce residues. The use of precision technologies is still in its infancy. The development of computer and satellite technology suggests a new, vast potential, but it has had limited use thus far. However, new products are continuously being introduced, and some types of precision technologies will play a major role in the future of agriculture. One challenge in improving precision technology will be to develop the software and management tools that will take advantage of new information. That will present a significant challenge to researchers in farm management. Other issues involve the development of institutions that take advantage of network externalities associated with knowledge and that accumulate and distribute information that is pertinent to farm management. The pricing of knowledge will also become a major issue of research within the context of precision farming.

Another important issue associated with precision farming is the potential for improving environmental quality. The adoption of precision farming may be induced by environmental regulation. The link between environmental regulation, research, development, and the adoption of new products needs to become clearer and provide insight to improve institutions and incentives. Most of the research on technology and innovation thus far has been done within regional bounds, but one of the main challenges of the future is to analyze issues of research and development within an international context. We need to better understand issues of technology transfer and intellectual property rights within nations. In some cases we need to better understand the mechanisms of collaboration between nations to address either global problems or to take advantage of increases in returns to scale. International food research centers and some existing binational research and development arrangements have become very prominent.[13] These types of arrangements may become more important in the future and should be further investigated. Furthermore, the relationship between the private and the public sectors in research and development should be viewed in a global context. A multinational corporation may change the research activities and infrastructure between nations in response to changes in economic conditions, and the activities of such private organizations depend both on national and international public sector policies.

International aspects of research and development are especially important in light of trade agreements such as GATT and NAFTA, and there is very little knowledge on how international trade agreements affect research and development. However, this type of knowledge is crucial because R&D is becoming a key element in the evolution of agricultural industries. An important issue to address, of course, is the development of research infrastructure on global problems, for example, private global climate change. Thus far, this research has been conducted by individual nations without much coordination of finance, finding, and direction. As we recognize our interdependence and

---

[13] For evaluation of the Binational Agricultural Research Development (BARD) fund between Israel and the United States, see [Just et al. (1988)].

the importance of issues such as global management of natural resources, fisheries, and biodiversity, we need to determine what type of mechanism we should use to enhance efficiency and research in knowledge development on a global basis.

In addition to an abundance of new research topics on innovations that should be addressed in the future, there are new research techniques and paradigms that seem very promising for the future. The new evolution in finance examines investments within the context of dynamics, and uncertainty should be further incorporated to assess the economics and management of research. Research activities should be evaluated as part of the management portfolio and financial activities of firms and concerns. The use of financial tools will provide new avenues for pricing research products and international property rights. However, tools, while very useful, have limits of their own. We need to better understand what kinds of processes, in terms of technology and economic and physical forces, give rise to the stochastic processes that are used in financial management. We need to better understand the dynamics of uncertain events and how they affect markets. Research agendas that link general equilibrium modeling with financial tools are an important challenge to economics in general but will be very important in the area of agricultural research and development.

Much of the research has emphasized technical innovations but it may be just as important to understand institutional innovations. What are the reasons for the emergence of institutions such as futures markets, farmer cooperatives, product quality warranties, etc.? To what extent are these institutions induced by economic conditions? How do human capital and political structures affect the emergence of institutions? Zilberman and Heiman (1997) suggest that economic research contributed to the emergence of institutional innovations (e.g., Keynesian macroeconomic policies, emission, etc.). But this topic needs to be studied in-depth which will enable better assessment of investments in social science research. Research on the emergence of institutions will benefit if we have a better understanding of how institutions actually work and the main features that characterize them.

Innovative activities are critically dependent on human capacity to make decisions and learn. The assumption of full rationality that characterizes many economic models is unrealistic. It will be useful to borrow the modeling approach from psychology and other behavioral sciences, and develop models of learning, adoption, and other choices that recognize bounded rationality. Thus far, there is much successful research in other areas, in particular, on uncertainty, and such direction will be very important in the study of innovation and technology.

Technological innovation and institutional change have a profound effect on the evolution of the agricultural sector. The agricultural economic literature on innovation clearly documents that innovations do not occur randomly, but rather that incentives and government policies affect the nature and the rate of innovation and adoption. Both the generation of new technologies and their adoption are affected by intentional public policies (e.g., funding of research and extension activities), unintended policies (e.g., manipulation of commodity prices), and activities of the private sector. One of the challenges of designing technology policies in agriculture is to obtain an optimal mix of public and private efforts. Design of these policies will require improved understanding of the economics of complex processes of innovation, learning, and adoption in a myriad of institutional and technological settings. Economists have made many notable advances through their research on innovation and adoption, but there remains much to be discovered.

References

Agricultural Issues Center (1994), "Financing agriculture in California's new risk environment", in: S.C. Blank, ed., S. Weber, technical ed., Proceedings of a Conference on December 1, 1993 in Sacramento, California (Agricultural Issues Center, University of California, Davis, CA).

Akerlof, G. (1976), "The economics of caste and of the rate race and other woeful tales", Quarterly Journal of Economics 90:591–617.

Alston, J.M., G.W. Norton and P.G. Pardey (1995), Science under Scarcity: Principles and Practice for Agricultural Research Evaluation and Priority Setting (Cornell University Press, Ithaca, NY).

Alston, J., R. Sexton and M. Zhang (1997), "The effects of imperfect competition on the size and distribution of research benefits", American Journal of Agricultural Economics 79(4):1252–1265.

Antle, J.M., and T. McGuckin (1993), "Technological innovation, agricultural productivity, and environmental quality", in: G.A. Carlson, D. Zilberman and J.A. Miranowski, eds., Agricultural and Environmental Resource Economics (Oxford University Press, New York).

Ball, V.E., J.-C. Bureau, R. Nehring and A. Somwaru (1997), "Agricultural productivity revisited", American Journal of Agricultural Economics 79(4):1045–1063.

Berck, P., and J.M. Perloff (1985), "The commons as a natural barrier to entry: Why there are so few fish farms", American Journal of Agricultural Economics 67(2):360–363.

Binswanger, H.P. (1974), "A microeconomic approach to induced innovation", Economic Journal 84(336):940–958.

Binswanger, H.P., and J. McIntire (1987), "Behavioral and material determinants of production relations in land abundant tropical agriculture", Journal of Economic Development and Cultural Change 36(1):73–100.

Boserup, E. (1965), Conditions of Agricultural Growth (Aldine Publishing Company, Chicago).

Brown, L.R., G. Gardner and B. Halweil (1999), "Beyond Malthus. Nineteen dimensions of the population challenge", in: L. Starke, ed., The Worldwatch Environmental Alert Series (W.W. Norton & Company, New York).

Carletto, C., A. de Janvry and E. Sadoulet (1996), "Knowledge, toxicity, and external shocks: The determinants of adoption and abandonment of non-traditional export crops by smallholders in Guatemala", Working paper 791 (Department of Agricultural and Resource Economics, University of California, Berkeley).

Caswell, M. (1991), "Irrigation technology adoption decisions: Empirical evidence", in: A. Dinar and D. Zilberman, eds., The Economics and Management Water and Drainage in Agriculture (Kluwer Academic Publishers, Boston).

Caswell, M.F., and D. Zilberman (1986), "The effects of well depth and land quality on the choice of irrigation technology", American Journal of Agricultural Economics 68(4):798–811.

Cavallo, D., and Y. Mundlak (1982), "Agriculture and economic growth in an open economy: The case of Argentina", Research Report 36 (International Food Policy Research Institute, Washington, DC).

Chavas, J.-P. (1993), "On sustainability and the economics of survival", American Journal of Agricultural Economics 75(1):72–83.

Cochrane, W.W. (1979), The Development of American Agriculture: A Historical Analysis (University of Minnesota Press, Minneapolis).

Coeymans, J.E., and Y. Mundlak (1993), "Sectoral growth in Chile: 1962–82", Research Report 95 (International Food Policy Research Institute, Washington, DC).

David, P.A. (1969), "A contribution to the theory of diffusion", Memorandum no. 71 (Stanford Center for Research in Economic Growth).

de Gorter, H., and E.O. Fisher (1993), "The dynamic effects of agricultural subsidies in the United States", Journal of Agricultural and Resource Economics 18(2):147–159.

de Gorter, H., and D. Zilberman (1990), "On the political economy of public good inputs in agriculture", American Journal of Agricultural Economics 72(1):131–137.

de Janvry, A., P. Leveen and D. Runsten (1981), "The political economy of technological change: Mechanization of tomato harvesting in California", Working paper 177 (Department of Agricultural and Resource Economics, University of California, Berkeley).

Diamond, J. (1999), Guns, Germs and Steel (W.W. Norton and Company, New York).

Dixit, A., and R. Pindyck (1994), Investment Under Uncertainty (Princeton University Press, Princeton).

Evenson, R., and Y. Kislev (1976), "A stochastic model of applied research", Journal of Political Economics 84(2):265–281.

Feder, G., R.E. Just and D. Zilberman (1985), "Adoption of agricultural innovations in developing countries: A survey", Economic Development and Cultural Change 33(2):255–298.

Feder, G., and G.T. O'Mara (1981), "Farm size and the adoption of green revolution technology", Economic Development and Cultural Change 30:59–76.

Gardner, B. (1988), The Economics of Agricultural Policies (Macmillan Publishing Company, New York).

Goodhue, R. (1997), Agricultural complementarities and coordination: Modeling value differentiation and production contracting, Ph.D. Dissertation (Department of Agricultural and Resource Economics, University of California, Berkeley).

Griliches, Z. (1957), "Hybrid corn: An exploration in the economics of technological change", Econometrica 25(4):501–522.

Griliches, Z. (1958), "Research costs and social returns: Hybrid corn and related innovations", Journal of Political Economy 66(5):419–431.

Guttman, J.M. (1978), "Interest groups and the demand for agricultural research", Journal of Political Economy 85(3):467–484.

Hamilton, S.F., and D.L. Sunding (1997), "The effect of farm supply shifts on concentration and market power in the food processing sector", American Journal of Agricultural Economics 79:524–531.

Hamilton, S.F., and D.L. Suuding (1998), "Returns to public investments in agriculture with imperfect downstream competition", American Journal of Agricultural Economics 80(4):830–838.

Hasset, K.A., and G.E. Metcalf (1992), Energy Tax Credits and Residential Conservation Investment (NBER, Cambridge).

Hayami, Y., and V.W. Ruttan (1985), Agricultural Development: An International Perspective (Johns Hopkins University Press, Baltimore).

Heiman, A., J. Zhao and D. Zilberman (1998), "Modeling money-back guarantees as financial options", Mimeograph (Department of Agricultural and Resource Economics, University of California, Berkeley).

Hoff, K., A. Braverman and J. Stiglitz (1993), The Economics of Rural Organization: Theory, Practice and Policy (Oxford University Press, New York, NY).

Huffman, W.E. (1998), "Finance, organization and impacts of U.S. agricultural research: Future prospects", Paper prepared for conference "Knowledge Generation and Transfer: Implications for Agriculture in the 21st Century", University of California, Berkeley, June 18–19, 1998.

Huffman, W.E., and R.E. Just (1994), "Funding, structure, and management of public agricultural research in the United States", American Journal of Agricultural Economics 76(4):744–759.

Just, R., and W. Chern (1980), "Tomatoes, technology, and oligopsony", Bell Journal of Economics 11:584–602.

Just, R.E., G.C. Rausser and D. Zilberman (1986), "Modelling the effects of policy on farmers in developing agriculture", International Journal of Development Planning Literature 1(3):287–300.

Just, R.E., and D. Zilberman (1983), "Stochastic structure, farm size, and technology adoption in developing agriculture", Oxford Economic Papers 35(2):307–328.