Working Papers in AGRICULTURAL ECONOMICS

Dec. 89-11

An Analysis of Alternate Micro Level Models of Investment Behavior

by
Eddy L. LaDue
Lynn H. Miller
and
Joseph H. Kwiatkowski

Department of Agricultural Economics
New York State College of Agriculture and Life Sciences
A Statutory College of the State University
Cornell University, Ithaca, New York, 14853-7801

Table of Contents

Introduction..1

Theoretical Framework...2

The Data..3

Individual Investment Models....................................3

Heat Recovery System (binary logit).............................4

The Model...4

The Results...6

Projected Probabilities..10

Precooler (bivariate probit)...................................15

The Model..15

The Results..17

Projected Probabilities..24

Expansion Investment Models....................................24

Expansion, Replacement or No Investment (ordinal logit)........24

The Model..27

1986 Expansion Results...28

1986 Projected Probabilities...................................29

1985-86 Expansion Results......................................32

1985-86 Projected Probabilities................................32

Expansion vs. No Expansion, 1980-86 (binary logit).............37

The Model..37

The Results..37

Projected Probabilities..40

Summary and Conclusions..42

Appendix...50

References...51

AN ANALYSIS OF ALTERNATE MICRO LEVEL MODELS OF INVESTMENT BEHAVIOR

Eddy L. LaDue, Lynn H. Miller and *Joseph H. Kwiatkowski**

A knowledge of the investment behavior of farmers should allow policy makers to improve their estimates of farmer response to changing investment stimuli and improve the selection of policy instruments used to influence investment. Such knowledge may also allow farm suppliers to modify the demand for their products by addressing those factors that influence investment in their products or in items that use their products.

Much of prior research on investment behavior has focused on macro level models using aggregate data (Bischoff, Clark, Fisher, LeBlanc and Hrubovcak, Penson et al., and Lamm). Such models are useful in identifying gross investment relationships and have been instrumental in estimating the effect of changes in national policy variables, such as changes in the tax code (LeBlanc and Hrubovcak). However, because of the level of aggregation of data used, these models tend to abstract from the behavior of individuals and treat the entire economy or all of agriculture as a single entity. Such a procedure treats the conflicting behavior of individuals on a net basis, potentially resulting in the appearance of no action when individual actions are offsetting, and fails to identify the behavioral characteristics of those who are reacting to investment stimuli compared to those who do not invest or disinvest.

Considerable literature exists on the investment behavior of individuals. However, much of the research has been descriptive in nature (for a review, see Brase and LaDue). Attempts to quantify the behavioral relationships connected with investment at the firm level have been limited. Using a probit model, Hill and Kau found farm size, farm type, tenancy, operator age and specific corn crop variables significant in determining investment in grain dryers in Illinois. Similarly, farm size, tenancy and corn production and use variables were important determinants of grain bin investment in a tobit study by Dixon, Hill and Saffell. A more recent multivariate analysis of tractor and combine investments found soil type, value of machinery inventory, operator age and education to influence machinery investment decision making (Johnson, Brown and O'Grady). In a simulated investment environment, Gustafson, Barry and Sonka found that structural characteristics of the farm, including tenancy, leverage and age of the existing machinery complement to influence machinery investment. Most of these researchers found socio-economic factors to be important. Hill and Kau observed a threshold stimuli level required for investment. Gustafson et al., found that a desire for an even pace of investment was important but that interest rate was not. Johnson et al., illustrates that the management decision making process allows for inclusion of noneconomic factors.

The micro level investment behavior models reported to date generally discuss only a few of the variables identified as influencing investment by the more descriptive literature. This may be due to lack of data, a narrow view of investment behavior or tight control over the theoretical model design. More realistically, it likely reflects strict adherence to

---

* Professor and Research Support Specialists, respectively, Department of Agricultural Economics, Cornell University. The authors would like to thank Loren Tauer and Timothy Mount for helpful reviews of an earlier draft.

classical analysis procedures that one carefully designs a theoretical economic model and then fits the data to the model. If that is the case, one suspects that much "analysis" that would be useful to other researchers is left in the cutting room floor. Five researchers finding that something is not important may be just as useful to decision makers as a similar degree of agreement in the affirmative.

The objectives of this research are twofold: (1) to compare alternate model specifications to test the apparent contribution or importance of a variety of behavioral characteristics including those identified as important to investment decisions by descriptive studies, and (2) to assess the behavioral characteristics of farmers who undertake net business investment through expansion of the business.

In the discussion that follows, four basic models are used to assess investment behavior. Each of these are discussed in turn after a review of the data used is presented. The first two models deal with investment in a particular piece of dairy farm technology. The first is a binary logit model of investment in heat recovery systems. The second is a bivariate probit model of investment for precoolers. The third and fourth models deal with expansion investment by farm businesses in New York State. The third is an ordinal logit model designed to elicit behavioral characteristics of farmers making no investment, purchasing replacement items only, or expanding the business. The fourth model is a binary logit model of expansion versus no expansion of the business. For each basic model, the economic and mathematical specification of the model are presented followed by the results. Following discussion of the four models, some conclusions are presented.

Theoretical Framework

One approach to identification of the basic forces influencing investment is to start with the firm level neoclassical model of optimal capital accumulation (Jorgenson) where net worth (N) of the firm is given by:

$$(1) \quad N = \int_{0}^{\infty} e\text{-}rt[P(t)Q(t)-w(t)L(t)-q(t)I(t)]$$

Where: P = Price of production (output)
Q = Quantity of output produced
w = Price of variable inputs
L = Quantity of inputs used
q = Price of capital
I = Investment in durable goods

From this model, it is clear that investment is a function of the prices of output, inputs and capital, the production function which establishes the level of output as a function of the amount of inputs and capital used and the time value of money or discount rate. Since optimal investment at any time is a function of future values of these variables, investment is determined by their expected value. Recognizing the lack of correspondence between the sale price of a used asset and the remaining flow of services from the asset implies a finite horizon and makes investment a function of the planning horizon. This lack of correspondence is particularly evident for buildings which often suffer a high level of capital loss upon construction, and new machinery which suffers a large decline in market value upon delivery at the farm.

Since investment is based on expected future income streams which are not known with certainty, expected income is probabilistic in nature. To reflect the fact that operators may value nonuniform probabilistic income streams differently, the model must be placed in a utility framework. Thus, the utility of expected income becomes a basic factor which may influence investment.

If the variables contained in the model represent the basic forces influencing investment, most of the variables identified by the literature as influencing investment are proxies for one or more of the basic forces. Some are more direct proxies than others but few could be called direct proxies for the basic forces influencing investment. In general, studies of investment behavior have not had access to direct measures of the basic forces. The proxy nature of the variables being assessed implies that some variables may be important to investment behavior (are good proxies) while others are not. Thus, comparison of alternate model specifications will be important in assessing behavior.

In developing the models discussed below, effort was made to avoid including more than one proxy for the same basic force in a model unless there was good reason to believe that the proxies would be complementary in reflecting the basic force rather duplicative. In selecting our models, we considered: (1) appropriate proxy sets, (2) prior research results, and (3) the specific characteristics of the investment.

The Data

The data used in this study were collected as part of a survey of a random sample of upstate New York farm businesses (Kelleher and Bills). Counties on Long Island and adjacent to New York City were excluded. A personal interview was used to obtain information on production and management practices, energy use and technology adoption as well as investment behavior. Reasonably complete information on investment related variables was obtained on 756 farms. Data were collected during the Spring of 1987 and are basically cross sectional observations.

Farms were defined as places selling $10,000 or more of agricultural products during the 1986 calendar year. Seventy-six percent of the farms surveyed were dairy farms. Representation of other farm types included: other livestock, six percent; cash crop, six percent; fruit, four percent; vegetable, two percent; horticulture, two percent, and miscellaneous, four percent. The 756 farms used in this analysis had average total farm assets of $423,000 and average annual gross receipts of $127,000.

Individual Investment Models

Investment behavior was investigated for two individual investment items: a heat recovery system and a precooler. Since one of the objectives of this research was to investigate the importance of various variables to investment behavior, several alternate models are expected to be compared. Selection of a "best model" by comparing several, invokes the optimism principle (Picard and Cook) resulting in potentially biased coefficients and statistical measure of goodness of fit. Practically, the model appears to classify farmers better than it does in fact. To address the problem, the sample was split into an estimating sample which was used to test alternate model specifications and a holdout sample which was used to determine (validate) the statistical properties of the final model. Observations were assigned to the two samples using a computerized random assignment process.

Heat Recovery System

A heat recovery system uses the heat removed from the milk at the bulk tank to preheat water going to the water heater. Heated refrigerant from the bulk tank is used to heat the water which cools the refrigerant before it is cycled back to the bulk tank. Since dairy farms must cool all milk from body temperature to 32-40 degrees Fahrenheit and use large amounts of hot water in the milking and cleaning process, substantially lower energy expenditures can be experienced when using a heat recovery system.

The Model

A binary logit model is used where the probability of investing in a heat recovery system (Y) is estimated as:

$$(2) \quad \hat{Y}_i = 1- \left[\frac{1}{1 + \exp\ (\alpha + \beta' X_i)}\right]$$

Where: $\hat{Y}_i$ is the predicted probability that a farmer i will invest in a heat recovery system given the values for Y_i.
α constant.
exp the base of the natural logarithm.
β' denotes the vector of regression parameters.
X_i denotes the values of the factors related to investment in a heat recovery system for farmer i.

To estimate this model it is reformulated as:

$$(3) \quad Y_i^* = \alpha + (\beta' X_i)$$

Where: Y_i^* is the log of the odds ratio of investing in a heat recovery system, i.e., $Y_i^* = \ln(\frac{Y_i}{1-Y_i})$

The model was estimated using the supplemental LOGIST procedure from the Statistical Analysis Systems Institute (Harrell). The dependent variable was one for farms with a heat recovery system and zero for those without such a system. Only data from dairy farms were used. A holdout sample was selected randomly from the total sample and set aside for development of model statistics for the final model selected. The estimating sample included 261 farms; the holdout sample included 267 farms. Sample sizes differ because of the random process used for assigning farms.

Since 38.6 percent of the farms in the sample had invested in a heat recovery system, this sample prior probability of investment was used as the cut-off point for classification of farmers. Farms for which the probability of investment in a heat recovery system exceeded the cut-off point were classified as investors. Those with lower probabilities were classified as not investing in a heat recovery system. This is consistent with Beaver (1966) who states that the optimal cut-off point is one that will minimize the percentage of incorrect classifications and Maddala (1983) who points out that, following Bayes theorem, if the costs of misclassifying are equal for both types of error, then the optimal cut-off point will be the prior population probability of being in a class. For this study, the costs of committing Type I or Type II error are equal. Although the population probability is not known, it is established from the sample probabilities.

The variables included in the models were those identified in the literature as influencing agricultural investment. The initial model contained eight economic variables. Number of cows was used as an indicator of size. Economies of scale imply that a minimum number of cows are needed to justify investment in a heat recovery system, resulting in a higher probability of investment with larger size. However, as herd size increases beyond some point, it appears likely that the added economic incentive would increase the probability of investment in a heat recovery only at a decreasing rate. Thus, cow numbers squared is added to allow for the expected curvelinear relationship.

The four geographic regions of New York State represent differences in soil and climate resources, input costs and milk prices. Region one has a variable resource base, somewhat higher input costs, higher milk prices and considerable urban pressure. Region two is largely hill and valley soils. Region three has the best soil and climate resources. Region four has modest soil resources, colder temperatures, lower input costs, lower milk prices and few alternatives to dairy farming. Although those factors could have an influence on heat recovery adoption, the expected sign of the included variables is indeterminate.

The existence of a parlor or pipeline milking system indicates a willingness to adopt milking technology. Also, parlors and pipelines are often more recent investments than a bucket or bucket/transfer system implying a higher probability of milking system investment since heat recovery systems have become available. Both of these factors would encourage heat recovery investment.

Education has been found to be positively correlated with investment in new technology (Funk). Education facilitates the evaluation process as well as management of the new asset. Education is measured as years of education.

The management index was constructed from respondents answers to questions about use of farm records, input buying strategies, marketing procedures, personnel management practices and short term goal setting behavior (LaDue and Kwiatkowski). The index has four levels with the best combination of management practices coded four. Given the profitability of heat recovery system for most farm situations, it is expected that better managers will be more likely to adopt.

Farmers with higher income expectations would be expected to have more funds available for investment and greater optimism for the future resulting in greater investment. In addition to data on their 1986 cash income, farmers were asked to indicate how that income had changed since 1980 and 1985, and how they expected it to change by 1990. These data were used to estimate 1980, 1985 and 1990 expectations as; (1) lower, (2) the same as, or (3) higher, than 1986.

Interest rate has long been considered an important factor influencing investment. Presumably lower interest rates make more investments financially feasible and, thus, results in greater investment. Age has frequently been found to be related to investment. Younger farmers invest as they are trying to increase their level of income. Older farmers who have reached a reasonable income and size level reduce investment and then disinvest as they near or reach retirement age.

The Results

The initial results are labeled Model 1 in Table 1. The variables that were not significant at the 0.10 level are age, interest rate and management ability. One of the three regional dummies was also insignificant, though the other two were significant. The sign on the dummy variable for those expecting future (1990) income to be greater than 1986 was unexpected.

The results achieved with age could be explained in two ways. First, age may represent a number of correlated variables such as education and size and when they are included in the model, age becomes unimportant. Alternately, the low dollar investment required for a heat recovery system and the high profitability of the investment may make it a high priority even for the young who are limited by resources to modest total investment and the old who are reducing total business size. Model 2 (not presented) excluded age as a variable and all coefficients and statistical tests were similar to those presented for Model 1.

The apparent unimportance of interest rate may also stem from the particular characteristics of this investment. The modest investment required likely implies that many farmers could make the investment from equity rather than borrowed capital. Also, the high relative level of income from the investment for businesses that have attained the minimum size likely makes the investment profitable for a wide range of interest rates. An alternate explanation would be measurement error in the interest rate variable. The rate used was the rate paid for the expansion investment for expanders and average rate paid for expanding farmers with the same primary credit source for nonexpanders. The actual rate that a farmer might pay on a precooler investment may differ from this rate.

Table 1. *Comparison of Heat Recovery Models*
Estimating Sample

Variable	Model 1	3	4	5	6	7
	---Model Coefficients and P Values[a]---					
Intercept	-7.221	-8.461	-8.239	-7.757	-6.549	-1.938
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)
Cows	.01	.012	.012	.011	.008	.016
	(.02)	(.01)	(.01)	(.01)	(.04)	(.00)
Cows ^{2} (00)	-.001	-.001	-.001	-.001	-.001	-.001
	(.08)	(.06)	(.04)	(.07)	(.14)	(.00)
Pipeline	2.365	2.374	2.358	2.329	2.442	2.552
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)
Parlor	3.210	3.187	3.184	3.010	2.968	3.115
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)
Education	.318	.319	.319	.260	.226	
	(.00)	(.00)	(.00)	(.00)	(.00)	
Region 1	.491	.514	.495	.266		
	(.32)	(.29)	(.31)	(.57)		
Region 2	1.351	1.317	1.395	1.275		
	(.00)	(.00)	(.00)	(.00)		
Region 4	1.100	1.102	1.102	.811		
	(.03)	(.03)	(.03)	(.09)		
Greater 1990 Inc.	-.879	-.833	-.817			
	(.02)	(.03)	(.03)			
Lesser 1990 Inc.	-1.414	-1.369	-1.353			
	(.02)	(.02)	(.02)			
Management	.142	.109				
	(.43)	(.54)				
Int. Rate	-.117					
	(.20)					
Age	.001					
	(.93)					
	---Model Statistics---					
Chi Square	90.6	94.6	92.7	83.7	70.5	64.1
P Value	0.00	0.00	0.00	0.00	0.00	0.00
R	.45	.47	.47	.45	.42	.41
C Statistic	.84	.84	.84	.83	.80	.79
	---Correct Classification Percentages---					
Total	77.5	76.9	77.6	73.7	72.3	70.9
With H.R.	79.1	76.3	77.9	73.7	72.6	66.3
Without H.R.	76.6	77.2	77.4	73.8	72.0	73.5
Class Efficiency	77.6	76.9	77.6	73.8	72.2	70.7

[a] P values are in parentheses under the coefficient. P value indicates the probability that the coefficient is zero.

Model 3 excludes interest rates from the variable list, resulting in only modest change in model coefficients and statistics (Table 1).

The reason for the insignificance of the management variable is less clear. It may be that the investment is simple enough that profitability is easily identified from herd size even for those with limited management skills. Alternately, education may sufficiently represent managerial capacity for this investment. Eliminating management as a variable results in Model 4 which was little different from Model 3. Some model statistics declined slightly while the classification efficiency improved modestly.

The inconsistent sign achieved with the income expectations variables led to an investigation of other income variables. Four alternates were tested as variations of Model 4. Because investment in the heat recovery system actually took place prior to the date of the survey, income expectations may have been more influenced by prior income levels. The four alternate income measures were; (1) a composite income variable calculated as a sum of net farm income, 25 percent of 1980, 50 percent of 1985 and 25 percent of 1986, (2) 1986 income, (3) 1990 expected income measured in dollars, and (4) 1980 income. When substituted into Model 4 one at a time, all of these variables had low coefficient values and were insignificant at the 0.2 level. From this, we conclude that income level and income expectations are unimportant to investment in a heat recovery system. It appears that the profitability of this particular investment can be identified with sufficient clarity that farmers are little influenced by general expectations about future profitability or income levels.

Excluding income measures results in Model 5. All coefficients have an acceptable sign and are either significant at the 0.1 level or tied to variables that are significant. For the total model, the adjusted pseudo R is 0.45 which is good for cross sectional farm level data. The C statistic[1] of 0.83 is also acceptable for this type of study.

The model correctly classifies over 73 percent of all farmers. The correct classification rates for those with heat recovery is similar to the rate for those without heat recovery. This compares quite favorably to the conditional probability rate of 52.6 percent. The conditional probability rate is the rate of correct classifications expected assuming one only knew the proportion of the population that had heat recovery systems[2].

---

1 With a binary model the C statistic is equivalent to the area under the receiver operating characteristic curve (Hanley and McNeil). Thus, the statistic has a range of 0.5 to 1.0 with 0.5 indicating no apparent discriminatory power and 1.0 indicating perfect discriminatory power. In this case the C statistic represents the probability that a randomly chosen farm with a heat recovery system will be correctly rated by being given a higher probability than a randomly drawn farm without a heat recovery system.

2 The naive conditional probability of a farm is calculated using the prior probability of the farm being in the two different groups. For this model the prior probability of having a heat recovery system is 38.6 percent. Thus, the conditional prior probability of correctly

To examine the possibility that some other variables frequently identified as influencing investment behavior may be important in heat recovery investment, variations of Model 5 were run incorporating these variables one at a time.

Survey respondents were asked to compare their likelihood of investing in a more risky investment to other farmers. Those indicating a greater likelihood are identified as risk tolerant. Those less likely to invest are identified as more risk averse and those who indicated they were like other farmers were called average risk. These divisions do not correspond to risk lover, risk neutral and risk averse because farmers on average have been found to be somewhat risk averse. It does provide a measure of relative risk aversion. It likely contains some self selection bias since it is based on self perception.

Given the low level of investment required in a heat recovery system the risk with investment is likely not perceived as significant. When this measure of risk aversion was incorporated as dummy variables, the coefficients either had the wrong sign, were statistically insignificant or both.

Partnerships and corporations invest more than individual proprietorship on both an absolute and percent of current asset basis (LaDue and Kwiatkowski). Such a propensity could sweep heat recovery investment along with it and then be important in determining investment behavior. However, when dummy variables for partnership and corporations (omitting sole proprietorship) were included in the model, the coefficients were not significantly different from zero at the .2 level. It appears that the small amount of investment required results in adoption even by those who invest only small amounts.

Survey respondents were asked if they expected to continue farming for the next 10 years and, if so, to indicate their primary goal from a list of eight goals. Those indicating their goals were allocated to two groups generally described as income increasing goals or improved family living goals. Dummy variables for the two general goal categories were added to Model 5, with an exit from farming in the next 10 years as the omitted variable. Coefficients were insignificant at the .05 level. In this case, it is likely that these farms that expected to leave farming in the next 10 years provided a somewhat random base for the dummy variables. Given the small magnitude of the investment and a payback period of less than 10 years, the expectation of being out of business likely had little effect on investment.

It has been theorized that proximity to urban areas will reduce the incentive to invest in real estate type assets. Since the heat recovery system is built into facilities that will not likely be moved if the farm is sold, it is plausible that distance to an urban area could influence investment. Countering that is the relatively quick payback on the heat

---

classifying a farm given this knowledge is (0.386)(0.386) + (0.614)(0.614) = 52.6 percent. To calculate the model's efficiency, the model's correct classification rates for each group are substituted. Thus, the efficiency is (0.386)(0.663) + (0.614)(0.735) = 0.707.

recovery investment itself. When distance to a city of 20,000 or more population was incorporated in Model 5 the coefficient sign was negative and not significant at the .2 level.

Given the high P values for two of the three regional dummy variables, the importance of region in the equation could be questioned. In Model 6 region is dropped. This resulted in a lower coefficient for herd size and a higher intercept. Overall model statistics deteriorated somewhat, although the reduction in classification ability was modest.

The only variable in Model 6 that is farmer rather than farm related is education. Thus, a parsimonious predictive model developed by dropping education from Model 6 results in Model 7. This model could be used to predict heat recovery investment without knowledge of farm operator characteristics and with only two pieces of information on the farm itself. The loss in model statistics and correct classification percentages from dropping the education variable is modest.

Based on the assessment that Model 6 is the "best" economic model, in that it has acceptable model statistics with relatively few problems and that Model 7 may be useful as a predictive model for situations without information on the farmer, these two models were refit using the holdout sample (Table 2). The holdout sample statistics represent the true statistical characteristics of the model. As expected, the model statistics and correct classification percentages were somewhat lower for the holdout sample than were observed with the estimating sample. However, the statistics are still quite acceptable for cross sectional farm data.

The only surprise was the lower value obtained for the education coefficient and that the education variable becomes insignificant. This appears to imply that either the two samples are significantly different in education characteristics or that the education variable is somewhat unstable in its effect.

Clearly the most important determinants of investment in a heat recovery system are the type of milking technology employed. The more capital intensive the technology, the more likely investment in a heat recovery system.

Projected Probabilities

Using the probability form of the equation, the estimated equation can be used to calculate the probability of investment in a heat recovery system for farms with different characteristics. To obtain the model coefficient that should have the highest likelihood of being accurate the final form of the equations were fit using the entire sample (estimating plus holdout samples)[3]. These coefficients were used to calculate the probabilities shown in Figures 1-3. The probability of a bucket or

---

[3] Only the coefficients generated by this process are used. The statistics are presented for information only.

transfer milking system farm investing in a heat recovery system is less than 10 percent (Figure 1). The likelihood of investment by pipeline owners ranges from about 35 percent for those with small herds to 70 percent for those with large herds. About half of the parlor owners with small herds invest in heat recovery compared to 85 percent for those with large herds. Figure 1 shows the effect of milking system and herd size on the probability of investment with education held at the average 12.55 years.

Table 2. *Validation and Heat Recovery Model Values*

Variables	Holdout Sample: Model 6	Holdout Sample: Model 7	All Observations: Model 6	All Observations: Model 7
	---Model Coefficients and P Values---			
Intercept	-3.972	-3.338	-5.099	-3.459
	(.00)	(.00)	(.00)	(.00)
Cows	.011	.012	.009	.009
	(.01)	(.01)	(.00)	(.00)
Cows ^{2} (00)	.001	.001	.001	.001
	(.03)	(.03)	(.01)	(.02)
Pipeline	2.196	2.224	2.333	2.360
	(.00)	(.00)	(.00)	(.00)
Parlor	2.496	3.556	2.760	2.838
	(.00)	(.00)	(.00)	(.00)
Education	.057		.134	
	(.49)		(.01)	
	---Model Statistics---			
Chi Square	63.2	66.3	128.9	123.0
P Value	0.00	0.00	0.00	0.00
R	.39	.40	.41	.41
C Statistic	.77	.77	.77	.78
	---Correct Classification Percentages---			
Total	66.9	68.5	67.8	69.3
With H.R.	84.8	84.8	72.9	71.0
Without H.R.	53.6	56.8	64.4	68.2
Class Efficiency	65.6	67.6	67.7	69.3

The level of education significantly increased the probability of investment (Figure 2). Only 30 percent of the pipeline owners with a 10th grade education could be expected to have a heat recovery system compared to nearly 50 percent for those with a college education. Similarly education increased the probability of investment by parlor owners from about 50 percent to over 70 percent. These probabilities were calculated for average herd sizes for each milking system: 62.7 cows for pipeline systems and 127.3 cows for parlor systems.

The combined effect of herd size and education level explains a large part of the variation in heat recovery system investment (Figure 3).

Figure 1. Probability of Owning a Heat Recovery System by Type of Milking System and Herd Size

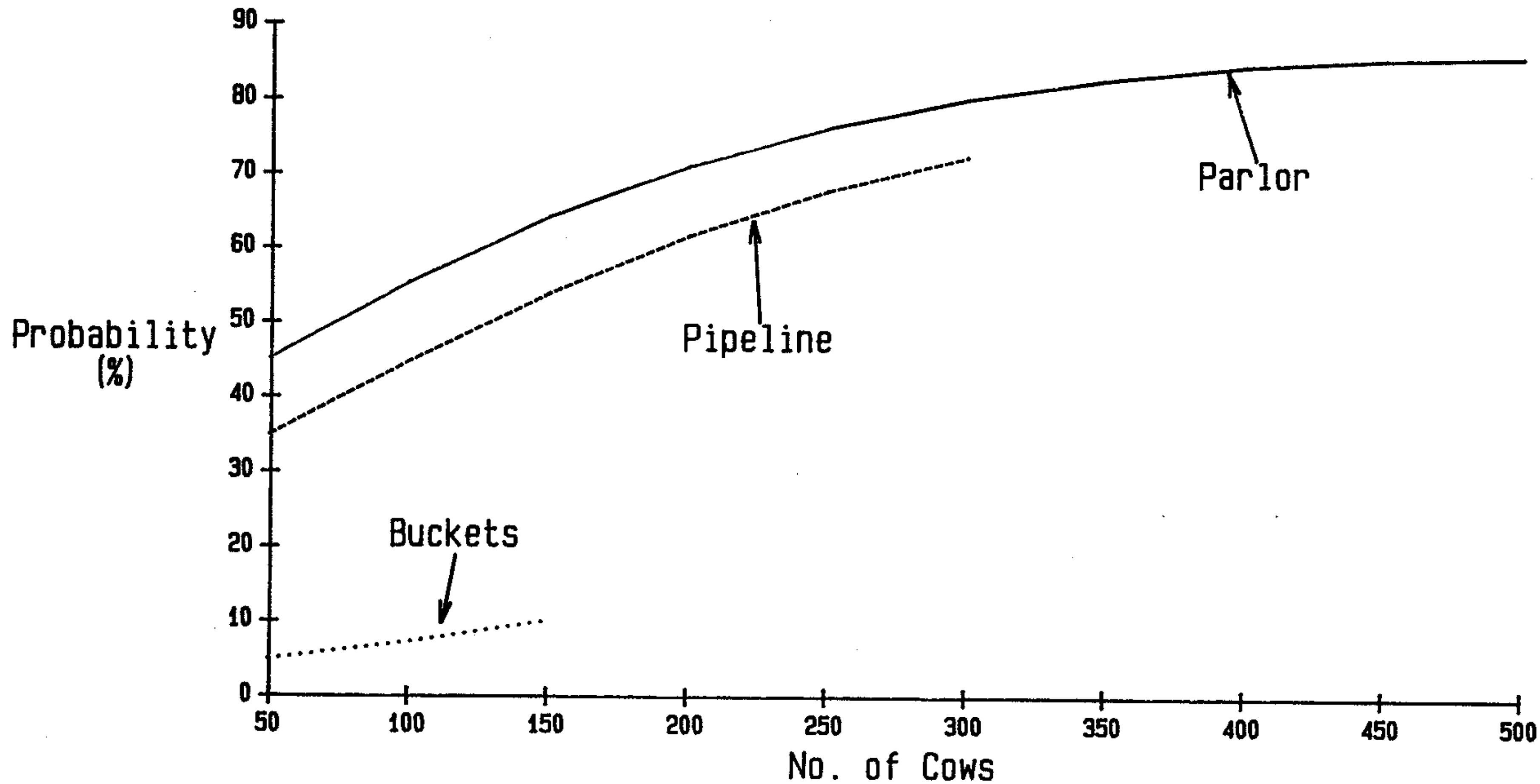


Figure 2. Probability of Owning a Heat Recovery System by Education Level of Operator Pipeline and Parlor Milking Systems

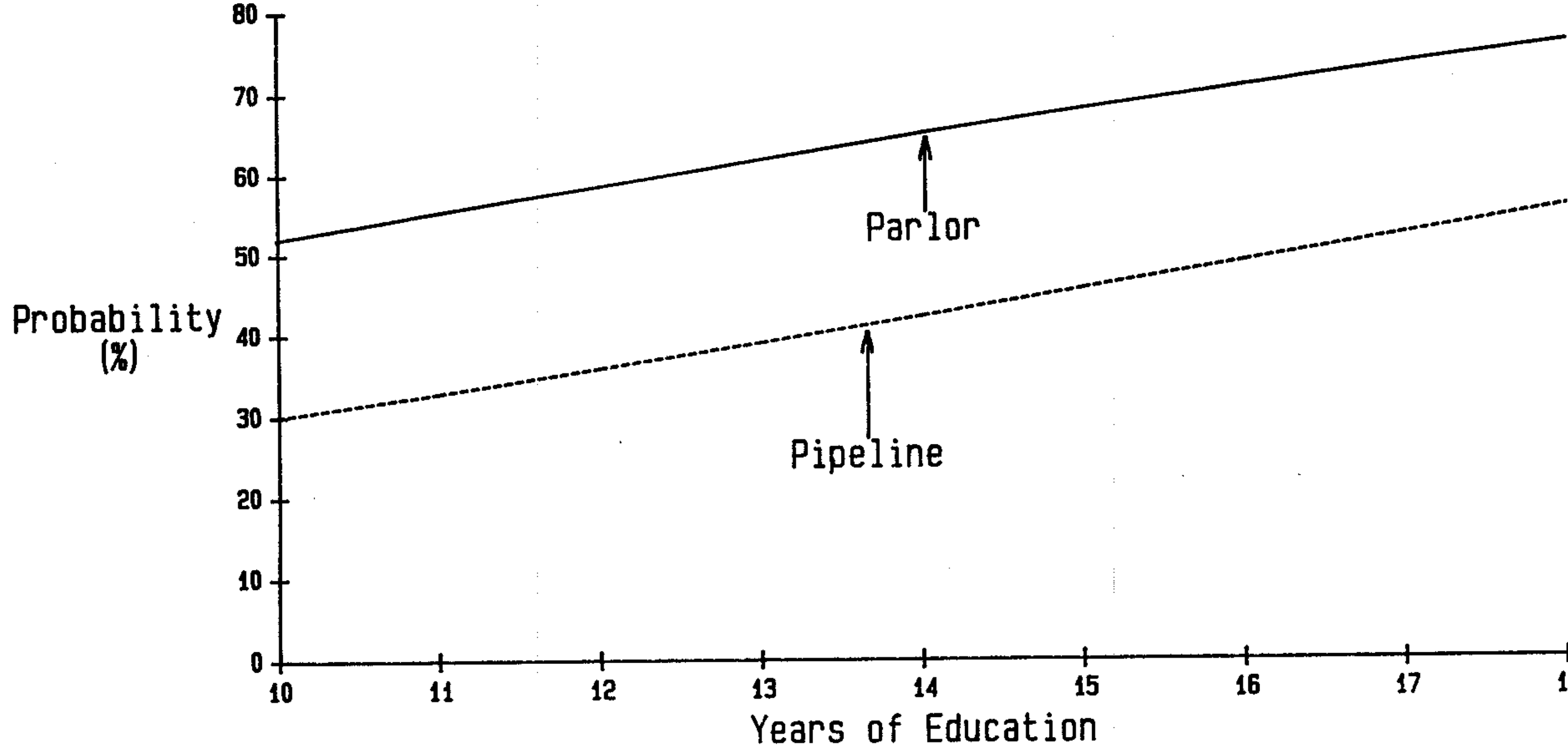
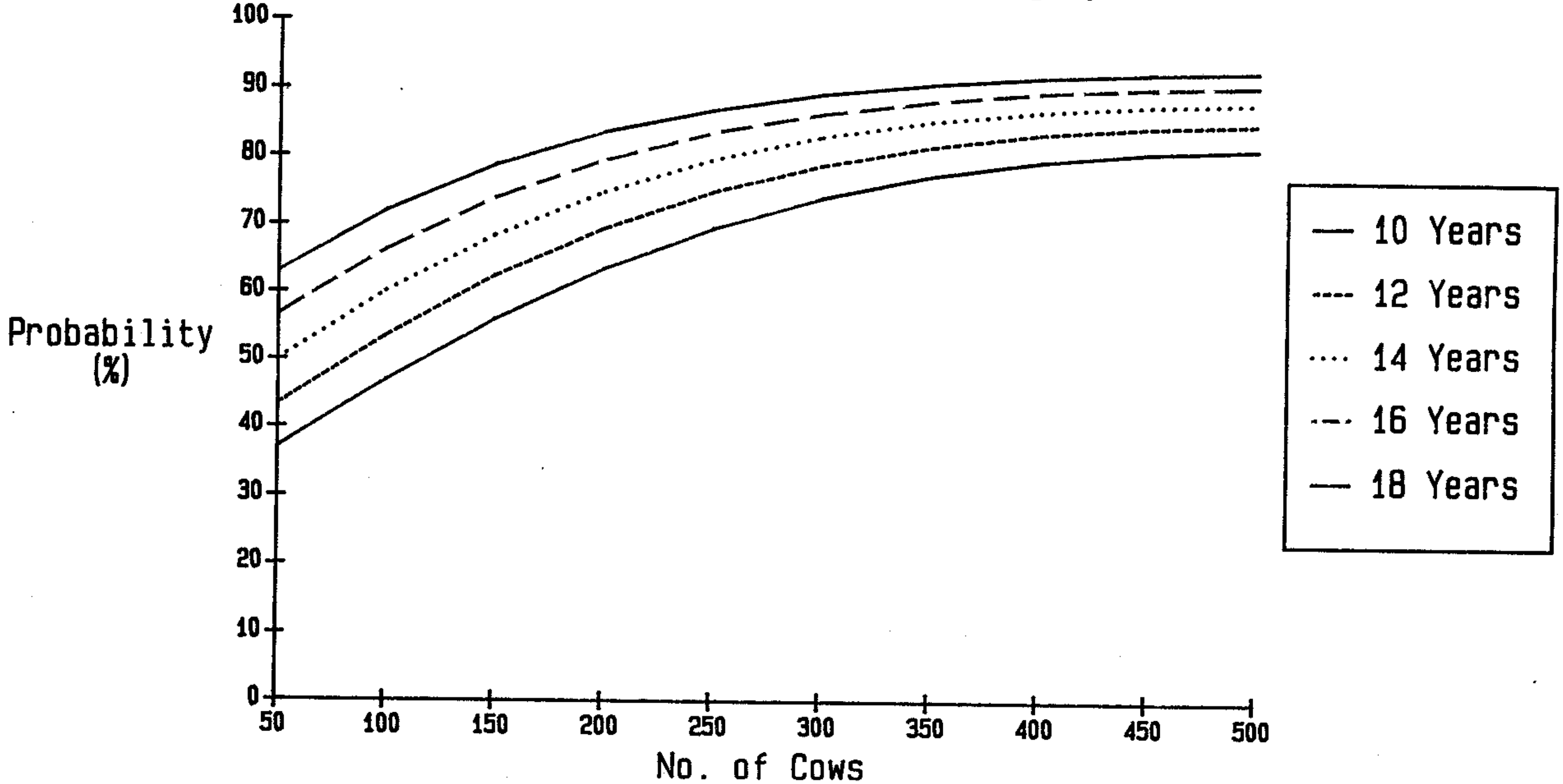


Figure 3. Effect of Education on Probability of Owning a Heat Recovery System, Parlor Milking System



Milking parlor owners with little education and a small herd only had about 35 percent chance of investing in a heat recovery system while those with college degrees and large herds had a 90 percent chance of owning a heat recovery system.

Precooler

A precooler uses cold well water to cool milk while it is being piped from the milking operation to the bulk tank. The milk passes through small tubes or channels that are surrounded by a counterflow of cold water. This process reduces the energy costs of cooling milk by reducing the temperature of the milk before it gets to the bulk tank cooling system. The warmed water resulting from this process is frequently used for washing or animal consumption.

The Model

As currently designed, this technology is possible only on farms with a milking parlor or pipeline. By selecting a milking system, farmers may simultaneously eliminate the possibility of precooler ownership. Further, practically no one familiar with dairy farm technology would suggest that parlor and pipeline ownership are randomly distributed among farms. Since we can only observe precooler ownership with farmers who have the appropriate milking system, we have self-selectivity bias (Maddala, 1987).

To adjust for this bias, the probability of having a parlor or pipeline must be incorporated in model design. This is superior to estimating a model using only the farms within a parlor or pipeline since data inherent in the self selection process would be omitted (Heckman, 1979). A bivariate probit model is used to simultaneously estimate two equations, one for ownership of a parlor or pipeline system and the second for investment in a precooler.

The two equations are specified as[4]:

(4) $Y^*_{i1} = \beta_1' X_{i1} + \epsilon_1$

(5) $Y^*_{i2} = \beta_2' X_{i2} + \epsilon_2$

Where for the i th observation (farmer)

- Y^*_1 is the estimated probit for investing in a precooler.
- Y^*_2 is the estimated probit for operating a milking parlor or around the barn pipeline milking system.
- β_1' denotes the vector of regression parameters for the investment in a precooler.
- β_2' denotes the vector of regression parameters for the operation of a milk parlor or around the barn pipeline milking system.
- X_1 denotes the vector of factors related to the investment in a precooler.
- X_2 denotes the vector of factors related to the operation of a milk parlor or around the barn pipeline milking system.

---

4 See VanDeVan and VanPraag for a similar application.

Since precoolers can only occur on farms with parlors or pipelines; (Y_{i1}, X_{i1}) is observed only when $Y_{i2}=1$.

It is assumed that: $\epsilon_j \sim N(0,1)$ with corr $(\epsilon_1, \epsilon_2) = \rho$

Because the correspondence between a zero value for an estimated probit (Y^*_{i1} or Y^*_{i2}) and the midpoint of the cumulative normal distribution, an individual farmer is expected to invest in a precooler (parlor or pipeline) if $Y^*_{i1} > 0$ ($Y^*_{i2} > 0$) i.e., $Y_{i1} = 1$ ($Y^*_{i2} = 1$) and not invest if $Y^*_{i1} < 0$ ($Y^*_{i1} < 0$), (i.e., $Y_{i1} = 0$). More succinctly:

$$Y_{ij} = 1 \text{ if } Y^*_{ij} > 0 \text{ or } Y_{ij} = 0 \text{ if } Y^*_{ij} \leq 0. \quad j = 1,2$$

The model was estimated using the bivariate probit option of LIMDEP (Green). The dependent variable for the milking system equation was one for farms with a parlor or pipeline and zero for farms without such investment, and for the precooler equation was one for farms with a precooler and zero for those without. Attempts to fit this model to the estimating sample remaining after setting aside a holdout sample of approximately 50 percent were unsuccessful. Nearly the entire sample was required for the program to obtain a solution. Thus, the model was estimated using the entire data set. The data requirements for estimation also limited the number of variables to be included in the base model.

The variables included in the milking system equation to determine the probability of a farm having a pipeline or milking parlor are conceptually similar to those important in heat recovery or precooler investment. Farm size, as measured by number of cows, is particularly important to milking parlor investment. Because of the large investment required for a parlor, it is a profitable investment only for large farms. Pipeline systems are usually installed to increase the number of cows that can be handled per milking. The average number of cows per farm participating in the Cornell Dairy Farm Business Summary in 1987 (Smith, et al) was 48 for bucket or transfer systems, 71 for pipeline systems and 157 for milking parlor systems.

Age of the operator has been shown to have a negative influence on the probability of large machinery investment (Hill and Kau, 1973). Age of the operator is, thus, expected to have a negative effect on the selection of a parlor or pipeline milking system.

Those with higher levels of education are expected to be better able to evaluate and manage the complexities of such a major technologically laden investment. Therefore, higher educated farmers are expected to more readily adopt parlor or pipeline milking systems. Regional dummies are included in the equation to represent the different soil, climatic and cost differences of the different regions of the state.

A composite cash income is included in this equation as a combination income expectations and cash flow variable. The variable is constructed as a sum of 25 percent of 1980, 50 percent of 1985 and 25 percent of 1986, net cash farm income. Since the milking systems were constructed prior to

1986, income over the 1980-85 period is likely more indicative of expectations at the time the investment was made. Those with more cash income for making payments on loans, or to make direct investments, and those with higher income expectations are expected to be more likely to invest in a modern milking system.

The precooler equation is specified exactly as the heat recovery system model except that the milking system variables are excluded. Because the precooler performs a similar function to that of the heat recovery system, the logic for inclusion of the individual variables is similar.

The Results

The results from estimating the hypothesized relationships are labeled Model 1 in Table 3. The overall model is significant in that the hypothesis that all the model coefficients are zero is rejected at the .01 level. The high level of correlation (.81) indicates that the (residuals of the) two equations are related and that fitting an equation for precooler investment using only those farmers who had a parlor or pipeline milking system would provide different results. Some factors that would appear to be influencing precooler investment would in truth be related to milking system adoption rather than precooler adoption per se.

As observed with the heat recovery model, age appears to have little relationship to this type of investment when other variables are included. In the precooler equation, the sign on age was the reverse of what would be expected and was clearly insignificant. Although age had the expected sign in the milking system equation, it was even less significant.

In Model 2, age is excluded from both equations. This omission had little effect on the overall model statistic. Model coefficients changed little except for some modest shifting of the intercept of the two equations. The coefficient for management index and cash income in the precooler equations changed but their very low level of significance implies that such instability could be expected.

Although management ability appears to be important in the milking system equation, it contributes little to the decision to invest in a precooler. As was observed for the heat recovery system, it may be that the precooler is a small enough investment or that the economics of investment are clear enough that management ability is not important in this investment decision. Alternately, education may sufficiently represent management ability for this decision. Investment must be related to managerial capacities that are better represented by education than the particular set of management practices contained in the management index. Eliminating management from the precooler equation had little effect on other model parameters or statistics (Model 3, Table 3).

In the milking system equation, however, education is consistently insignificant but management is significant. Farmers with better management ability appear more likely to believe they can correctly evaluate and manage such a system for their farm and to have the funds or credit capacity required to make the large investment. Education adds

Table 3. *Comparison of Bivariate Probit Models of Precooler Investment*

Variable	Model 1	2	3	4	5	6
Precooler Equation						
Intercept	-2.940	-2.415	-2.448	-2.460	-2.756	-2.758
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)
Cows	.807	.832	.833	.821	.909	.908
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)
Cows ^{2} (00)	-.070	-.074	-.074	-.073	-.083	-.083
	(.02)	(.02)	(.02)	(.02)	(.00)	(.00)
Education	.109	.095	.095	.097	.101	.102
	(.02)	(.02)	(.02)	(.01)	(.00)	(.01)
Cash Income	-.027	-.164	-.016	.017	-.005	
	(.86)	(.91)	(.91)	(.91)	(.97)	
Region 1	-.440	-.440	-.435	-.434		
	(.06)	(.06)	(.06)	(.06)		
Region 2	-.093	-.068	-..064	-.066		
	(.65)	(.73)	(.74)	(.74)		
Region 4	-.151	-.163	-.155	-.156		
	(.50)	(.46)	(.47)	(.47)		
Management	-.008	-.012				
	(.92)	(.89)				
Age	.007					
	(.34)					
Milking System Equation						
Intercept	-1.758	-1.936	-1.941	-1.766	-1.769	-1.769
	(.02)	(.00)	(.00)	(.00)	(.00)	(.00)
Cows	2.228	2.224	2.229	2.230	2.262	2.263
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)
Management	.258	.261	.257	.265	.267	.267
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)
Region 1	.741	.735	.735	.742	.659	.660
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)
Region 2	.525	.519	.520	.519	.515	.515
	(.01)	(.01)	(.01)	(.01)	(.00)	(.01)
Region 4	.468	.474	.475	.474	.452	.452
	(.04)	(.03)	(.03)	(.03)	(.00)	(.04)
Cash Income	.515	.509	.513	.507	.516	.515
	(.14)	(.15)	(.14)	(.14)	(.00)	(.14)
Education	.011	.014	.015			
	(.81)	(.74)	(.73)			
Age	.741					
	(.70)					
Model Statistics						
Log Likelihood	-341.1	-341.7	-341.7	-341.8	-343.8	-343.8
Correlation	-0.811	-.809	-.799	-.809	-0.753	-0.752
Significance	0.00	0.00	0.00	0.00	0.00	0.00

little to the assessment of managerial capacities indicated by the managerial index. Eliminating education from the milking system equation had little effect except to shift the intercept slightly (Model 4, Table 3).

Geographic region appears to be important in the decision to invest in a milking system but not in the precooler decision. Coefficients for two of the three regions are insignificant. Even region 1 (Southeastern New York) is significant only at the .06 level. It seems plausible that the regional cost, soil and climate differences would be important in the decision to invest in a milking parlor, but that once that decision (a parlor or pipeline system) was adopted these factors would be unimportant to the precooler decision. Since the precooler uses quantities of water, regional differences in water availability could influence investment. However, it appears that either this is not the case or that areas with limited water are not synonymous with the geographic regions selected.

Omitting the region variables from the precooler equation caused some minor shifting of the coefficients of the model (Model 5, Table 3). The intercept declined and the curvelinear nature of the cows variable became more pronounced. Overall model statistics were largely unchanged.

Cash income also appears to be important in the milking system decision, but not the precooler decision. Given the relatively small dollar outlay required to install a precooler, this result is quite logical. Eliminating the cash income variables had practically no effect on either equation (Model 6, Table 3).

The final model, Model 6, includes all of the variables of the initial model except age in one or both of the equations. All coefficients are significant at the .05 level except cash income. To test other variables that the literature has identified as important to investment behavior, the following were added to Model 6 one at a time: (1) form of business ownership, in the form of dummies for sole proprietorship, partnership or corporations, (2) distance to city of 20,000 or more, (3) degree of risk tolerance in the form of dummies for more risk averse, more risk tolerance or average risk tolerance, and (4) goals for next 10 years specified as, sell the farm, pass farm to next generation, expand size or income or improve family living. None of these variables had significant coefficients or improved the significance of the overall model.

The correct classification percentages for Model 6 are shown in the top half of Table 4. Since 22.4 percent of the sample farms had precoolers, that sample prior probability was used on the cutoff point for classification of farmers. Classification percentages are not calculated for the other models because the classification proportions are expected to move with the statistical properties of the model. In spite of the rather favorable statistical properties of the model, its classification ability is about the same as the conditional probability rate.

Consistent with the findings of the heat recovery investment analysis, the most important factors influencing investment in a precooler are herd size and education. However, investment in the milking system required for use of a precooler is also influenced by management ability, geographic region and cash income.

Table 4. *Bivariate Precooler Model 6 Correct Classification Percentages*

Total Sample	65.8
With Precooler	85.4
Without Precooler	60.2
Conditional Model Rate	65.2 [a]
Model Classification Efficiency	65.8

[a] Based on sample probability of precooler investment of 22.4 percent, (22.4 x 22.4 + 77.6 x 77.6).

A tenfold cross-validation analysis was used to estimate how error rate (correct classification rate) for the final model developed (Model 6). Cross-validation estimates the true error rate by removing each observation from the data sets used in its own predictions (Efron).

For the V-fold cross-validation procedure, all observations are randomly divided into V groups of approximately equal size. Observations in V-1 of the groups are used to estimate the model coefficient. The observations in the omitted group are then classified by the reestimated model. This procedure is repeated V times, each time with a different group left out. The average correct classification rate of V omitted samples provides a nearly unbiased estimate of the true correct classification rate (Efron; Frydman, Altman and Kao).

For this analysis, a V of 10 was used. The SAS RANUNI random number generator was used to randomly assign each observation to the 10 groups. Probit values were calculated from the precooler equation generated from each estimating sample. Probit values were converted to probabilities using the PHI cumulative standard normal distribution function from LIMDEP (Greene). A cutoff point equal to the prior probability of precooler investment (22.4) was used in developing predictions. Calculations included only those farms with a parlor or pipeline system (bucket system farms were omitted).

The results of the cross-validation analysis (Table 5) indicate that the true error rate of the model is approximately equal to or slightly greater than the conditional model rate. Although the model does not improve our probability of correctly classifying investment, it does provide some indication of the factors influencing investment. A predictive model with a lower error rate (higher correct classification rate) could be developed by modifying the cutoff probability used for classification. For example, use of a .5 probability cutoff as is used by the LIMDEP model (Green) significantly increases the correct percentages for those without a precooler (and reduces it for those with a precooler) and, thus, raises the correct classification rate (for more detail see Appendix A).

Table 5. *Comparison of 1986 Expansion Investment Models*
Estimating Sample

Variable	Model 1	2	3	4	5	6	7	8	9	10	Hold-out Sample Model 10
	---Coefficient and P Value---										
Intercept 1	-2.144 (.17)	-2.183 (.15)	-2.287 (.12)	-2.233 (.13)	-1.100 (.34)	-1.584 (.15)	-1.194 (.25)	-0.496 (.46)	.373 (.49)	.637 (.16)	1.84 (.01)
Intercept 2	-5.241 (.00)	-5.213 (.00)	-5.2778 (.00)	-5.224 (.00)	-4.078 (.00)	-4.53 (.00)	-4.138 (.00)	-3.433 (.00)	-2.538 (.00)	-2.314 (.00)	-1.617 (.00)
Gross Income	.004 (.00)	.004 (.00)	.004 (.00)	.004 (.00)	.004 (.00)	.004 (.00)	.005 (.00)	.005 (.00)	.005 (.00)	.005 (.00)	.003 (.01)
G. Income ^{2} ($1,000)	-.001 (.01)	-.001 (.01)	-.001 (.00)	-.001 (.00)	-.001 (.00)	-.001 (.00)	-.001 (.00)	-.001 (.00)	-.002 (.00)	-.002 (.00)	-.001 (.21)
Age	-.012 (.30)	-.016 (.17)	-.012 (.28)	-.012 (.28)	-.015 (.16)	-.015 (.17)	-.015 (.16)	-.017 (.09)	-.022 (.02)	-.018 (.03)	-.019 (.03)
Equity (%)	.923 (.12)	.909 (.12)	.854 (.13)	.850 (.14)	.806 (.15)	.845 (.13)	.711 (.18)	.725 (.18)	.683 (.20)		
Imprv. Fam. Liv.	.706 (.10)	.671 (.11)	.630 (.11)	.632 (.11)	.689 (.08)	.714 (.07)	.692 (.07)	.680 (.07)			
Leave and Pass Farm on	.624 (.34)	.511 (.42)	.456 (.46)	.451 (.46)	.604 (.31)	.636 (.28)	.628 (.28)	.637 (.27)			
Increase Inc.	.868 (.06)	.785 (.08)	.720 (.10)	.723 (.09)	.750 (.08)	.809 (.05)	.774 (.06)	.758 (.06)			
Education	.076 (.25)	.085 (.19)	.085 (.18)	.086 (.17)	.060 (.33)	.059 (.41)	.050 (.41)				
Management	.056 (.69)	.059 (.67)	.068 (.62)	.066 (.63)	.044 (.74)	.076 (.57)					

Table 5. (con't) *Comparison of 1986 Expansion Investment Models*
Estimating Sample

Variable	Model 1	2	3	4	5	6	7	8	9	10	Hold-out Sample Model 10
Region 1	-.570 (.10)	-.535 (.11)	-.406 (.22)	-.409 (.22)	-.424 (.20)						
Region 2	-.414 (.24)	-.491 (.14)	-.431 (.19)	-.400 (.19)	-.403 (.17)						
Region 4	-.616 (.10)	-.631 (.07)	-.498 (.15)	-.494 (.15)	-.514 (.13)						
Interest Rate	.062 (.42)	.069 (.36)	.057 (.45)	.057 (.45)							
City Distance	.001 (.95)	.002 (.77)	.002 (.78)								
Greater 1990 inc	.021 (.94)	-.007 (.98)									
Lesser 1990 inc.	-.019 (.97)	-.030 (.94)									
Poultry Farm	.063 (.94)										
Vegetable Farm	-.692 (.10)										
Grape Farm	-.327 (.71)										
Tree Fruit Farm	1.314 (.06)										
Horticulture	-.057 (.93)										

Table 5. (con't) *Comparison of 1986 Expansion Investment Models Estimating Sample*

Variable	Model 1	2	3	4	5	6	7	8	9	10	Hold-out Sample Model 10
Livestock Farm	-.410 (.56)										
Crop Farm	.207 (.78)										
Misc. Farm	-.189 (.85)										
---Model Statistics---											
Chi Square	58.6	50.9	51.0	51.0	51.4	48.0	47.8	47.8	42.76	50.73	17.70
Degrees Freedom	24	16	14	13	12	9	8	7	4	3	3
P Value	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
R	.14	.18	.20	.21	.21	.22	.23	.24	.24	.26	.13
C Statistic	.72	.70	.70	.70	.69	.69	.68	.68	.67	.68	.61

Projected Probabilities

For this model, the probabilities that a farmer will invest in a precooler is indicated by:

$$(6) \quad \hat{Y}_i = 1\text{-}F(-\beta i' X_{i1})$$

Where: $\hat{Y}_i$ = the probability that farmer i will invest in a precooler.
F = standard normal cumulative distribution function, N(0,1). That is (from Maddala):

$$(7) \quad F(-\beta_i' X_{i1}) = \int_{-\infty}^{-\beta' Xi} (2\pi)^{1/2} \ \exp\left(\frac{t2}{-2}\right) dt.$$

Calculating the investment probabilities using Model 6 (Table 3) results in Figures 4 and 5. Few farms with small herd sizes invested in precoolers (Figure 4). However, farms with 400 to 500 cows had a 80-90 percent probability of precooler investment. Similarly, only about 12 percent of the farmers with only 10th grade or less education owned a precooler (Figure 5). The probability of investment increased as education increased so that those with some college beyond the Bachelors level had about a 35 percent likelihood of investment.

Expansion Investment Models

A more basic level of investment behavior relates to whether net investment is being made. Studies of investment in individual items may include replacement investment, frequently involving the substitution of one technology for another. Investment behavior for such investments may differ from that related to expansion of a business. Studies of net investment using aggregate data are by definition using data where opposite decisions of different individuals are netted out (i.e., disinvestment by some farmers off-sets investment by others). Such analysis may obscure the multiple relationships that exist. An analysis of behavior relative to net investment by individual farms could help identify the basic individual forces that influence observed net investment. Two approaches to such an analysis are discussed below.

Surveyed farmers were asked to indicate whether they had expanded their business during the 1980-86 period (Kelleher and Bills). The character of that expansion was then determined by a series of questions on the types and amounts of investment. Farmers were also asked about their total investments for 1985 and 1986. Total investment in land, land improvements, buildings, livestock, machinery and equipment were requested. Using these data it was possible to determine which farmers made net investments, those that only replaced existing items and those who made purchases in 1985 and 1986.

Expansion, Replacement or No Investment

Three groups of farmers who exhibited basically different investment behavior are those who; (1) expanded their businesses (net investment

Figure 4. Effect of Herd Size on the Probability of Owning a Precooler

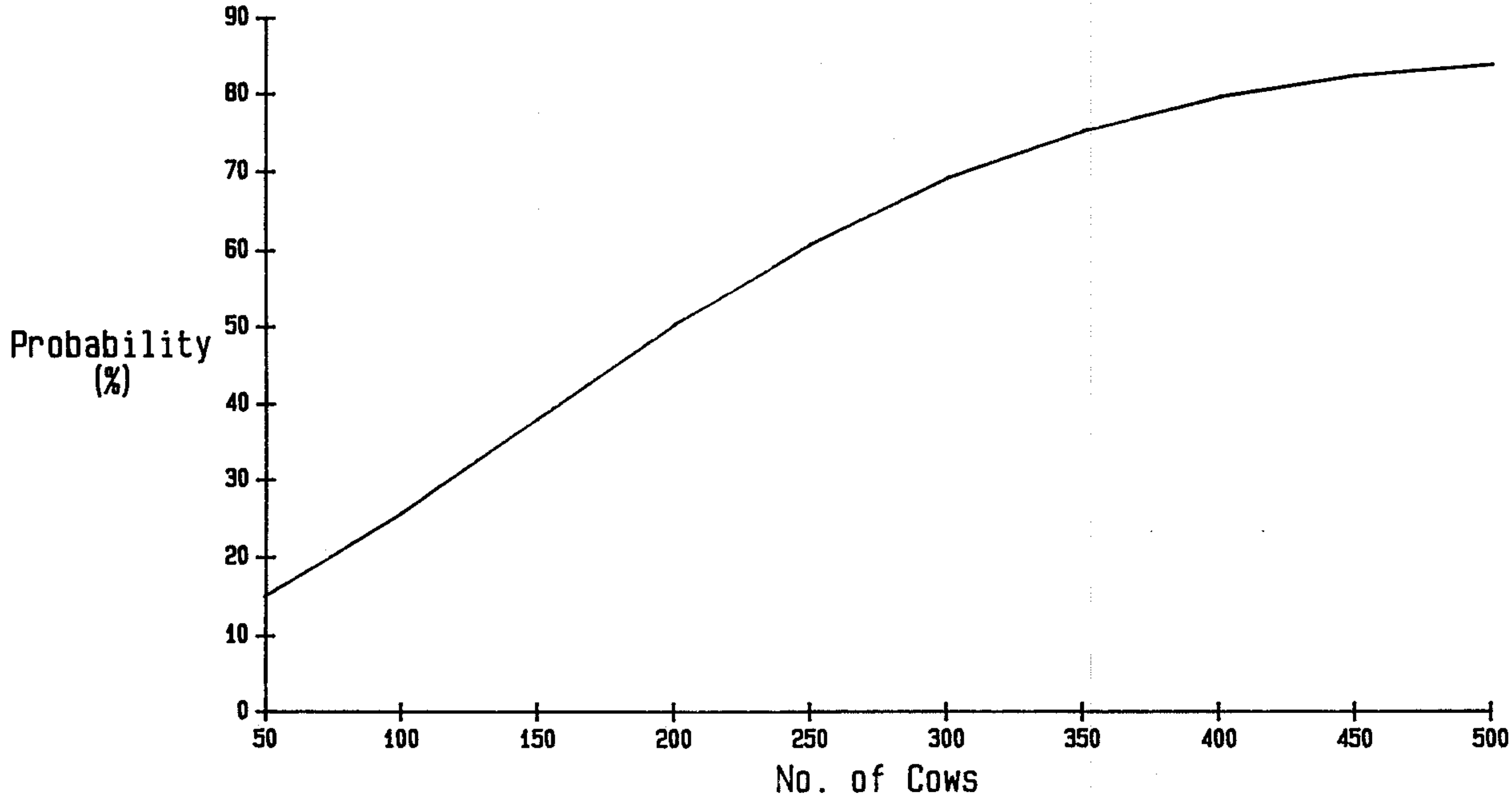
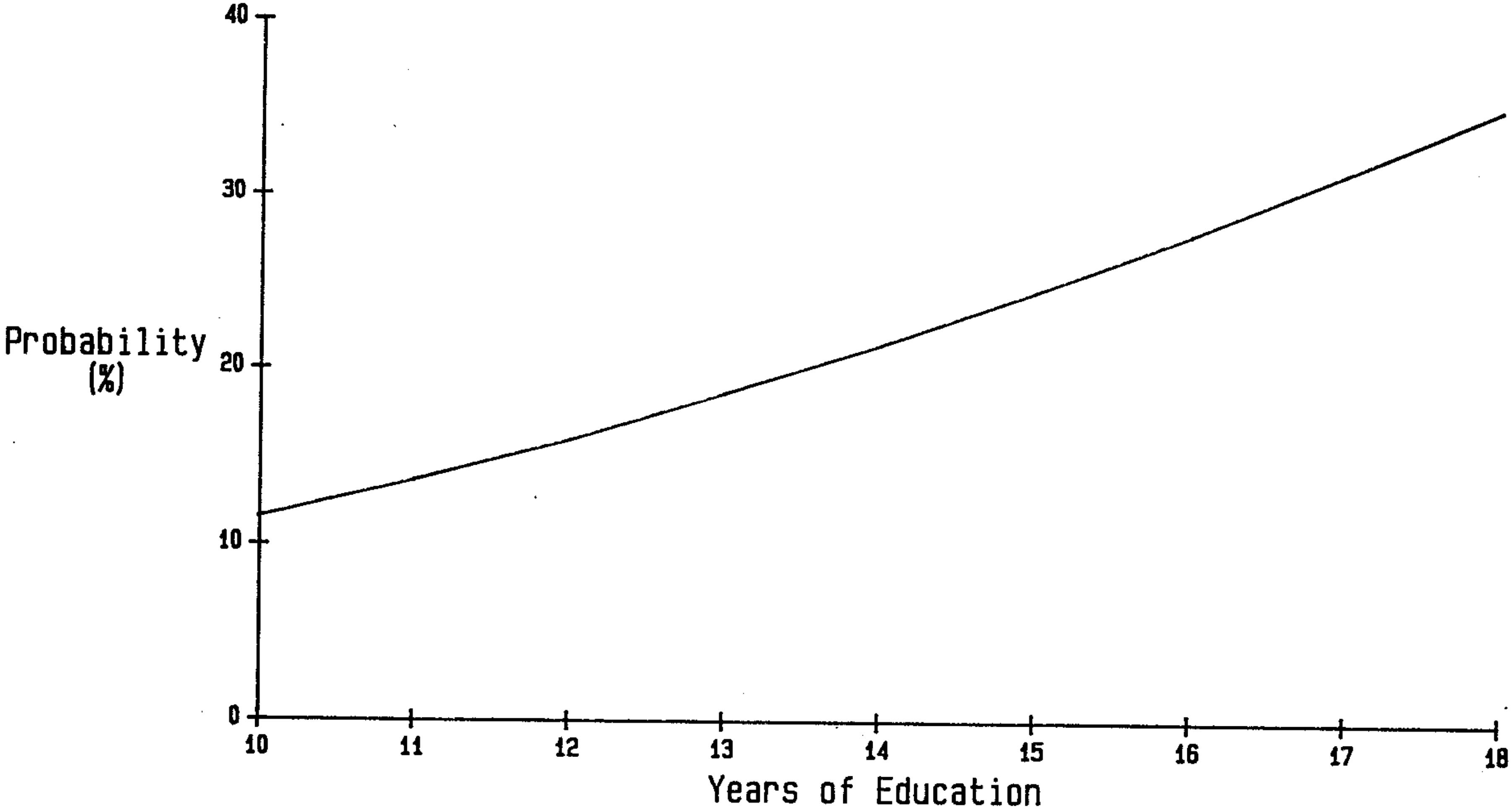


Figure 5. Effect of Education on the Probability of Owning a Precooler



occurred, (2) invested only in replacement items (no net investment occurred, and (3) made no investment (net disinvestment occurred). These three groups are ordered in the sense that to make a net investment (expand) one must first purchase sufficient assets to offset the decay of capital items and then make additional purchase to increase the existing capital stock. Alternately stated, in order for a person to move from the no investment group to the net investment group, replacement investment must be made.

The Model

For this analysis, an ordered response logit model is used. The model can be specified (Walker and Duncan) as:

$$(8) \quad P_{ik} = 1 - \left[\frac{1}{1 + \exp\ (\alpha_k + \beta' X_i)}\right]$$

Where for the i th observation:

- P_{ik} is the probability that the observation is > k.
- k denotes the levels of the observed dependent variable.
  - k_0 = no investment
  - k_1 = replacement investment
  - k_2 = expansion investment
- β' denotes the vector of regression parameters.
- X_i denotes the vector of factors related to farm investment.

Furthermore, with this model, the lowest probability level and α_o are not estimated. For this model, $P_0 = 1\text{-}P_1$. To classify observations with this model, Y^* equals the k level with the maximum calculated probability. This model makes no assumptions about the interval distance between category values.

This model was estimated using the LOGIST procedure from the Statistical Analysis Systems Institute (Harrell). Data for all farm types were used. The estimating sample contained 360 observations, the holdout sample contained 373 observations.

Models for two different dependent variables were estimated. The 1986 investment model investigated behavior relative to investment during the 1986 year. The second model considered investment in 1985 and 1986. Since the data were collected in early 1987, data are most complete and accurate for 1986. Also, data were collected for only the most recent expansion, allowing determination of replacement investment by subtracting the amount spent for expansion from total investment for the year. The same procedure was used for 1985 and 1986 except that the calculation required the assumption that farms did not expand in both 1985 and 1986. With that assumption, data on the most recent expansion allowed determination of replacement and expansion investment for both years.

Both of these models assess the probability of investment during a specific short period of time (one or two years). The model does not assess the general characteristics of those who invest at some point in time versus those who do not. It addresses whether investment will take place during the years in question.

The initial model contained 11 variables believed to influence the level of investment by farmers. Size is expected to influence investment in that larger farms frequently have a wide array of investment possibilities that may be profitable and may find it easier to generate internal or external funding for investment projects. Size was measured by gross farm income.

Given the life cycle of farm investment, younger farmers are generally investing in their businesses with middle aged farmers maintaining the status quo and older farmers disinvesting. Thus, expansion investment is less likely for older farmers.

The basic rationale for region, education, management ability, interest rate, distance to city and income expectations for 1990 are similar to those identified for heat recovery investment. The definition of these variables are the same as used with the heat recovery model. Omitted variables are Region 3 and income expectations that 1990 will be similar to 1986.

Type of farm is expected to influence the investment opportunities and the basic profitability of investment and, thus, could influence the amount of investment made. Dummy variables are included for all except dairy farms. It has long been assumed that the goals of the farmer influence his or her investment decisions in that the economic paths of the business is largely determined by the investment and disinvestment that take place. Farms were divided into one of four goal categories: (1) those who expect the current business will cease operations within the next 10 years, (2) those who plan to pass the farm on to the next generation reasonably soon, (3) those who desire to increase income or net worth, and (4) those who desire to improve family living. Dummy variables are included for all except the first goal category.

The financial position of the business influences the farmers ability to obtain funding. A high level of equity implies that there is reserve credit capacity that can be drawn upon for making investments.

1986 Expansion Results

Model statistics for the estimated 1986 investment model (Table 5, Model 1), indicate a high level of significance for the total model. The C statistic indicates that the model has discriminatory power in separating no investment, replacement or expansion. However, the low R value indicates that only a modest proportion of the variability in investment is explained by the model and a number of the included variables appear to make little contribution to the model.

Most of the farm type coefficients were not significantly different from zero, indicating that at least for 1986, type of farm was not naturally related to investment. Eliminating farm type (Model 2) improved the proportion of the variance explained and had little effect on the remaining variables.

The income expectations (expected 1990 income) variables were not significantly different from zero and the sign on the higher expectations variable was not expected. Eliminating these variables had little impact on the model except for a slight decline in the absolute value of the region variables (Model 3). Given the apparent importance of gross income in the model, it appears that investment is more closely related to current income than how that income level may change in future years.

Distance to city was consistently insignificant. Eliminating it from the model had practically no effect on model coefficients or significance levels (Model 4). Similarly, interest rate (Model 5), geographic region (Model 6), management index (Model 7) and education level (Model 8) were all insignificant and appeared to detract little from the model by their elimination.

One of the goal variables was consistently insignificant. Percent equity was nearly significant in early models but significance declined as other variables were omitted. Dropping both of these variables (Model 9 and 10) had very little effect on the remaining coefficients or overall model statistics. The R value actually improved slightly.

The final model (Model 10), indicates that the factors influencing investment are gross income and age. All coefficients are significant at the .95 level except one intercept. The R value is the highest of all models considered. The C statistic was of similar magnitude for all models considered.

The holdout sample confirmed the importance of the variables included in the final model except for the squared term for gross income. The chi square value declined sharply although the P value indicated a high level of significance for the overall model. The R statistic dropped to half of the modest level achieved with the estimating sample. However, this was similar to the R value obtained with the original model (Model 1). The process of dropping a large number of the variables included the original model appears to have had little effect on the proportion of variation explained by the model. However, the holdout sample C statistic implies that the discriminatory power of the model with fewer variables is lower.

1986 Projected Probabilities

The investment probabilities obtained by using the coefficients from the final model were re-estimated using all observations (estimating and holdout sample). The result indicates the importance of farm size and operator age in explaining 1986 investment (Figure 6). Those with low gross incomes had a high probability of no investment or of purchasing only replacement items. As size increased, the likelihood that expansion investment would occur increases sharply. Very few of the large farms did not invest. Fifty to 75 percent of them expanded.

Age was strongly related to the probability of making no investment (Figure 7). About one-quarter of young operators made no investment while nearly half of older farmers made no investment. Half as many older farmers made expansions compared to younger farmers. Not surprisingly for the time period involved, about half of the farmers made only replacement investments regardless of age.

Figure 6. Probability of Expansion, Replacement or No Investment by Gross Farm Income, 1986

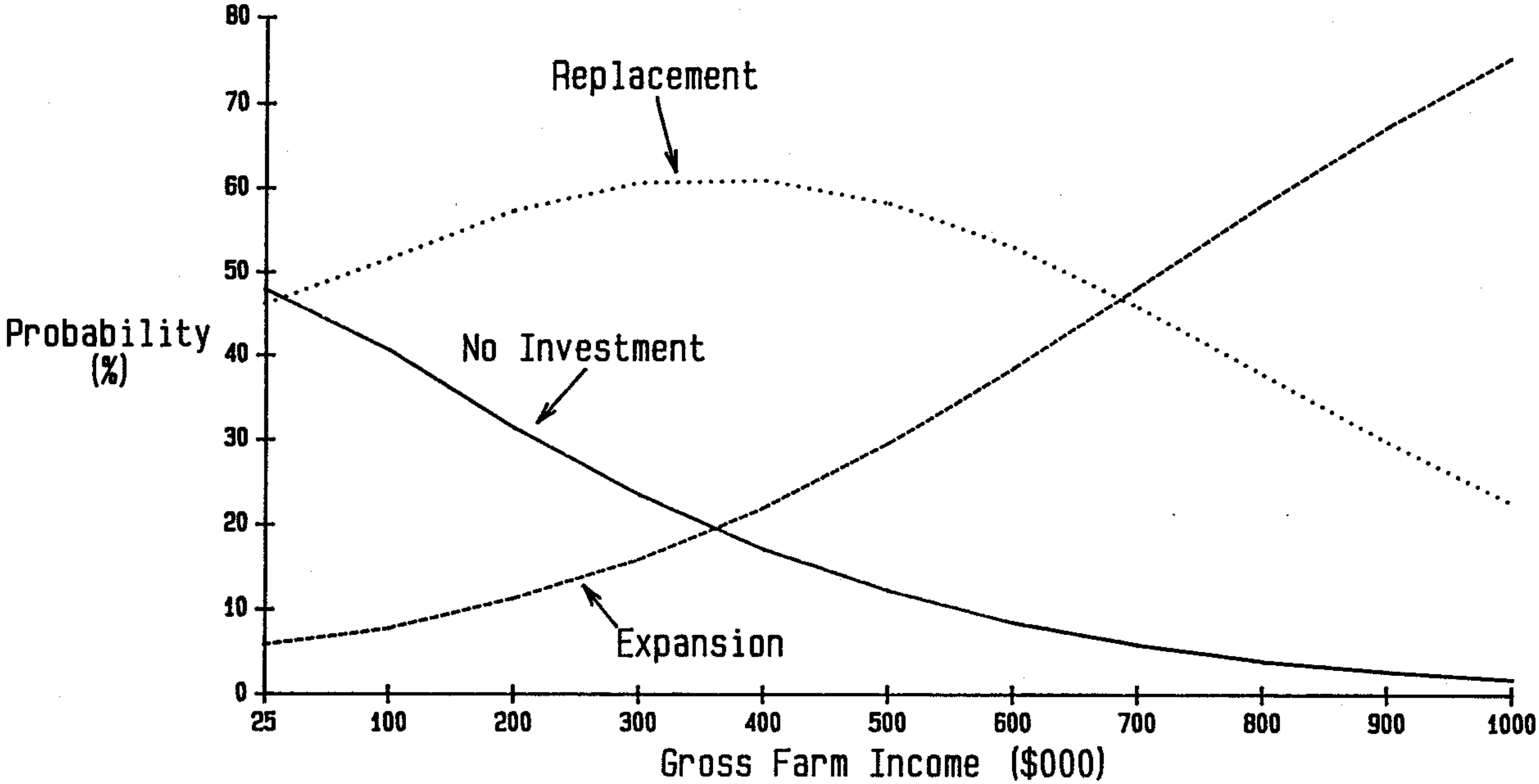
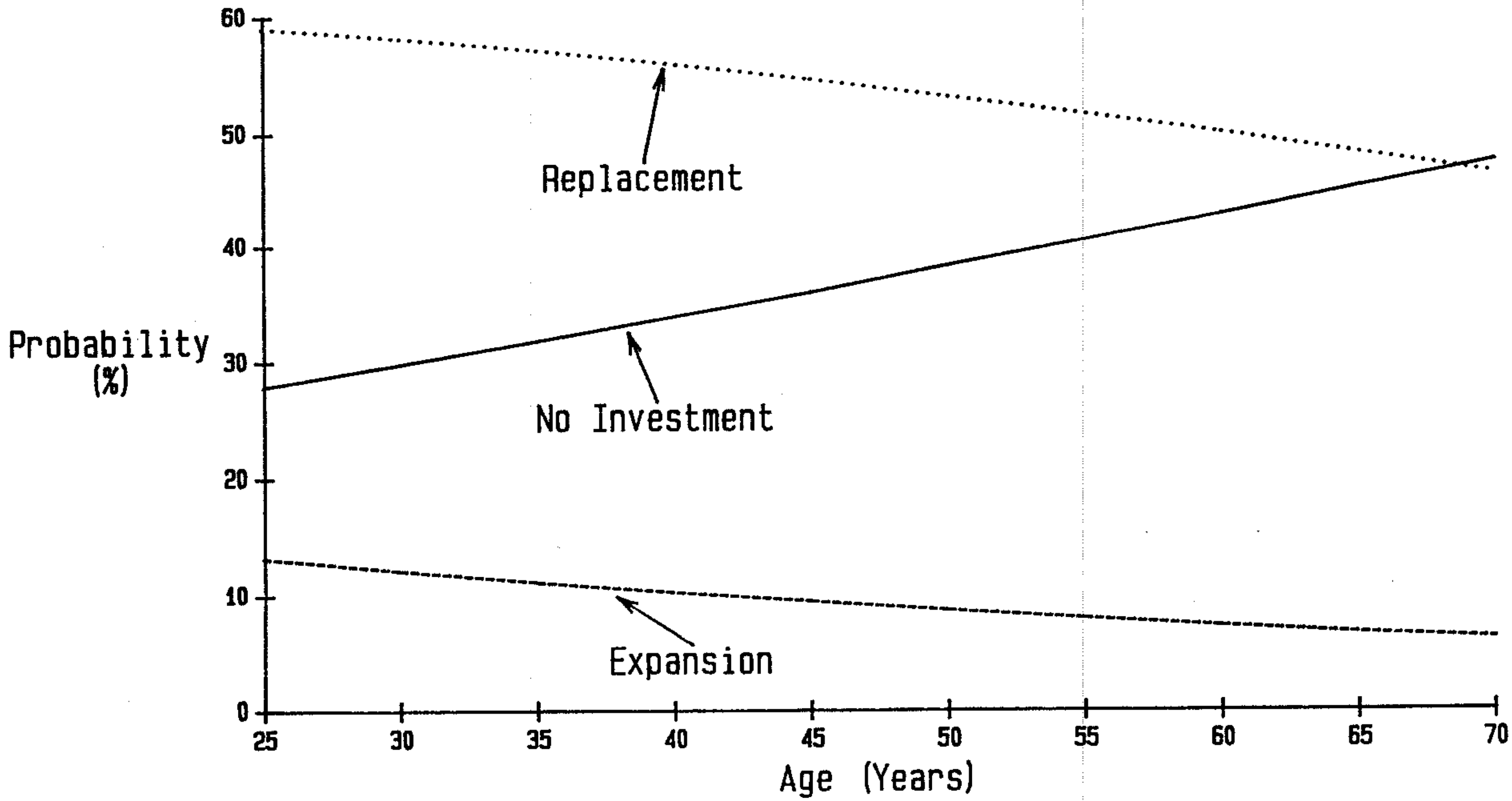


Figure 7. Probability of Expansion, Replacement and No Investment by Percent Operator Age, 1986



1985-86 Expansion Results

Statistics for the model estimated for investment in either 1985 or 1986 are similar to those found using only 1986 data (Table 6, Model 1). The P value indicates a high level of significance for the total model. The C statistic indicates some discriminatory power in separating no investment, replacement investment only and expansion investment. However, the low R value indicates that only a modest proportion of the variation in investment is explained and a number of the included variables appear to make little contribution to the model.

Interest rate was insignificant and did not have the expected sign (Model 1). Income expectation variables were highly insignificant and one had an unexpected sign (Model 2). Neither management nor education appeared to contribute to the explanatory power of the model (Models 3 and 4). Distance to city had low coefficient values and little significant (Model 5). Region variables did not contribute (Model 6). Unlike the experience with 1986 investment, goal variables were highly insignificant throughout the analysis (Model 7).

Percent equity boarders on being significant at the .9 level. Elimination of the management variable, although highly insignificant itself, appeared to lower the value of the coefficient and level of significance of equity. With the other insignificant variables omitted, equity itself becomes clearly insignificant at the .90 level. Dropping equity from the model has little effect on the coefficients of the remaining variables except that the age coefficient returns to near the level observed in prior models (Model 9).

Holdout sample results with the 1985-86 models were similar to the results with 1986 above. The overall equation is highly significant although the chi square value drops significantly. The R value drops significantly from the level achieved with the estimating sample and is similar to the original model (Model 1) results. However, with 1985-86 investment the holdout sample C statistic was similar to that obtained with the estimating sample.

1985-86 Projected Probabilities

The final model (Model 9) was re-estimated using the complete sample of farms (estimating and holdout) to obtain the best estimate of the coefficient values. Those vales were used to calculate investment probabilities. The probabilities obtained were similar to those obtained using solely 1986 data. A substantial number of businesses of all sizes except the very largest only replaced investment items during 1985-86 (Figure 8). About 40 percent of the smaller businesses made no investment whatsoever. As size increased, the proportion making no investment declined strongly and those expanding increased rapidly, especially for the very large farm.

As age increased, the likelihood that no investment would be made increased from about 20 percent to about 40 percent and the likelihood of expansion declined (Figure 9). The proportion who only purchased

Table 6. *Comparison of Models for 1985 or 1986 Expansion Estimating Sample*

Variable	Model 1	2	3	4	5	6	7	8	9	Hold-out Sample Model 9
	---Coefficient and P Value---									
Intercept 1	-0.946 (.53)	-.126 (.92)	-.331 (.78)	.134 (.87)	.476 (.52)	.605 (.39)	.391 (.56)	.923 (.09)	1.229 (.01)	1.690 (.00)
Intercept 2	-3.777 (.01)	-2.964 (.01)	-3.15 (.01)	-2.67 (.00)	-2.30 (.00)	-2.17 (.00)	-2.37 (.00)	-1.81 (.00)	-1.53 (.00)	-1.0027 (.03)
Gross Income	.003 (.01)	.003 (.01)	.003 (.01)	.004 (.00)	.004 (.00)	.004 (.00)	.004 (.00)	.004 (.00)	.004 (.00)	.003 (.01)
Gross Income ^{2} ($000)	-.001 (.10)	-.001 (.08)	-.001 (.06)	-.001 (.01)	-.001 (.01)	-.001 (.01)	-.001 (.01)	-.001 (.00)	-.001 (.00)	-.001 (.15)
Age	-.0153 (.17)	-.017 (.12)	-.014 (.18)	-.017 (.11)	-.017 (.11)	-.017 (.11)	-.016 (.12)	-.022 (.03)	-.018 (.04)	-.021 (.02)
Equity	1.01 (.08)	.974 (.08)	.961 (.08)	.958 (.08)	.835 (.12)	.844 (.12)	.828 (.12)	.782 (.14)		
Impr. Fam. Liv.	.162 (.68)	.232 (.56)	.147 (.69)	.130 (.73)	.182 (.61)	.187 (.61)	.191 (.60)			
Incr. Income	.619 (.16)	.627 (.15)	.493 (.23)	.447 (.28)	.482 (.23)	.485 (.20)	.496 (.20)			
Pass on Farm	-.002 (.99)	.079 (.90)	-.014 (.98)	-.014 (.98)	.044 (.94)	.032 (.96)	.031 (.96)			
Region 1	-.468 (.16)	-.514 (.12)	-.392 (.23)	-.357 (.27)	-.36 (.27)	-.369 (.25)				
Region 2	-.216 (.51)	-.392 (.22)	-.278 (.37)	-.251 (.41)	-.21 (.48)	-.146 (.61)				
Region 4	-.331 (.34)	-.466 (.17)	-.376 (.26)	-.359 (.28)	-.358 (.27)	-.348 (.28)				

Table 6. (con't) *Comparison of Models for 1985 or 1986 Expansion Estimating Sample*

Variable	Model 1	2	3	4	5	6	7	8	9	Hold-out Sample Model 9
City Distance	.008 (.29)	.01 (.19)	.01 (.22)	.01 .(23)	.01 (.48)					
Management	.132 (.34)	.110 (.41)	.113 (.39)	.108 (.41)						
Education	.062 (.33)	.027 (.65)	.029 (.62)							
Greater 1990 Income	-.020 (.93)	-.102 (.69)								
Lesser 1990 Income	-.134 (.75)	.01 (.95)								
Interest Rate	0.18 (.80)									
---Model Statistics---										
Chi Square	43.65	43.30	41.90	41.62	40.20	39.70	37.84	34.48	42.87	20.58
Degrees Fre.	16	15	13	12	11	10	7	4	3	3
P Value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R	.14	.15	.16	.16	.17	.18	.19	.20	.23	.14
C Statistic	.67	.67	.66	.65	.65	.65	.65	.64	.64	.63

Figure 8. Probability of Expansion, Replacement or No Investment by Gross Farm Income Upstate New York Farmers, 1985-86

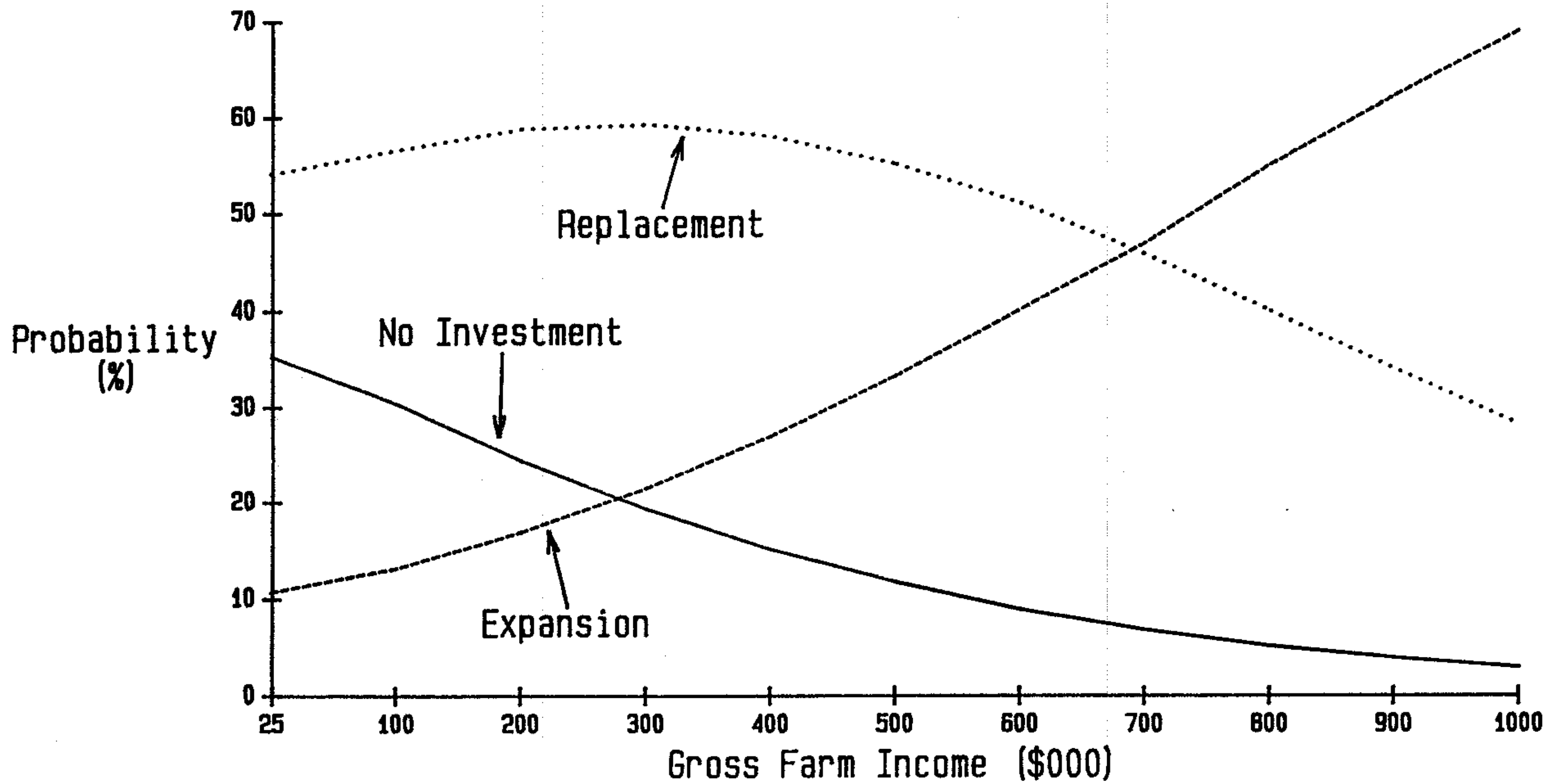
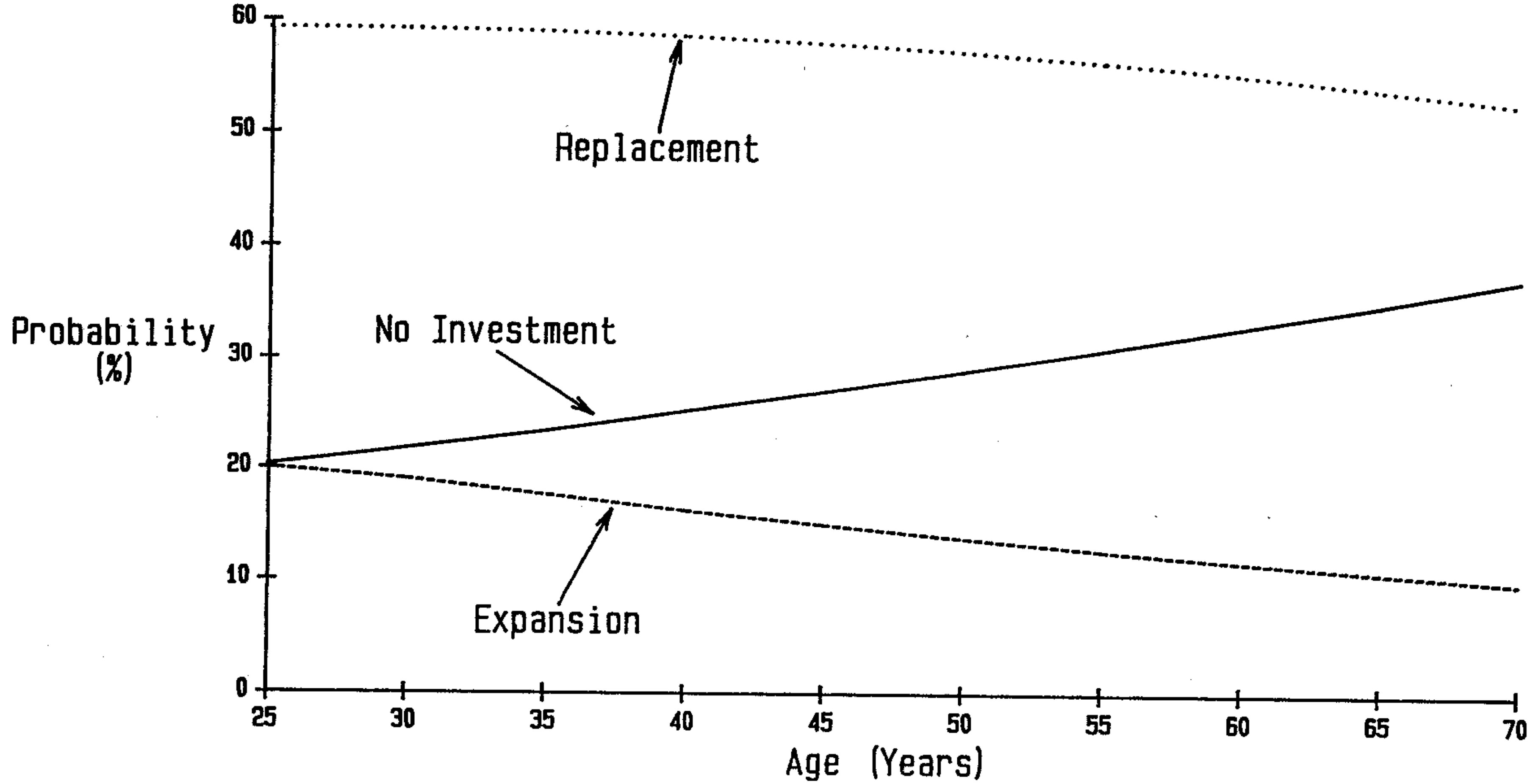


Figure 9. Probability of Expansion, Replacement or No Investment by Age Upstate New York Farmers, 1985-86



replacement items remained relatively constant with only a very slight decrease as age increased.

Expansion vs. No Expansion, 1980-86

At its most basic level investment behavior revolves around whether net investment takes place over a period of years. This avoids the "specific year" problem where investment may not take place on a particular farm in a specific year for a wide variety of reasons not related to basic investment incentives. Personal health of the operator(s), weather, family vacation and graduation plans and the location of concerted sales efforts by manufacturers and dealers are examples. Focusing on net investment regardless of the particular assets purchased also abstracts from the specific characteristic of the investment resulting in a study of the characteristics of those who invest versus those who do not invest.

The Model

This basic level of investment behavior was investigated using the logit model described in equations (2) and (3). The model was estimated using the supplemental logist procedure from the Statistical Analysis Systems Institute (Harrell). The dependent variable was one (1) for those farmers who expanded their business by at least 20 percent sometime during the 1980-86 period, and zero (0) for those who did not expand during that period.

The variables included in the basic model and the rationale for inclusion are similar to those used in the previously discussed models. Percent equity was not included because a preinvestment equity level could not be calculated. Data were available only on the most recent expansion and total expansion investment during the seven year period was not collected. The 1986 equity level was expected to be a poor indicator of preinvestment equity. The three dummy variables for risk were consolidated into two: risk tolerant versus all others.

The Results

The initial model chi square and P values indicate a high level of significance. The C statistic indicates some discriminatory power. The R value is modest, even for individual farm data, indicating that much of the variability in investment is not explained by the model.

The interest rate coefficient was small in absolute value and highly insignificant (Table 7, Model 1). It appears that the differences in interest rates paid by farmers were not sufficient to decide whether expansion was undertaken by individual farmers. Also some farmers with the lowest interest rates (i.e., Farmers Home Administration borrowers), were least able to expand for reasons not connected to rates.

Distance to city has little effect on investment and is not statistically significant. Omitting the variables had little effect on model values (Model 3). Only one of the regions showed any tendency to be significant. Eliminating the regions had little effect (Model 4) except to cause a shift in the magnitude of the coefficients for education and

Table 7. *Comparison of Models for Expansion Investment in 1980-86*

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Holdout Sample Model 6	Entire Sample
	---Coefficient and P Value---							
Intercept	-.869 (.57)	-1.234 (.32)	-1.502 (.22)	-1.11 (.34)	-.319 (.76)	-.516 (.61)	-1.506 (.11)	-1.152 (.09)
Age	-.029 (.02)	-.026 (.03)	-.026 (.03)	-.031 (.01)	-.035 (.00)	-.034 (.00)	-.015 (.11)	-.022 (.00)
Education	.105 (.11)	.127 (.05)	.126 (.05)	.106 (.09)	.102 (.10)	.109 (.07)	.027 (.64)	.076 (.07)
Management	.169 (.18)	.162 (.19)	.165 (.18)	.194 (.10)	.216 (.06)	.264 (.02)	.547 (.00)	.381 (.00)
Risk	1.001 (.01)	1.089 (.00)	1.076 (.00)	1.020 (.00)	.938 (.01)	.989 (.00)	.124 (.68)	.497 (.02)
1980 Net Income ($000)	.004 (.11)	.004 (.13)	.004 (.12)	.004 (.10)	.004 (.14)			
Family Living	.556 (.19)	.517 (.21)	.526 (.21)	.554 (.18)				
Incr. Income	.825 (.07)	.825 (.06)	.842 (.06)	.743 (.09)				
Pass on Farm	-.341 (.60)	.439 (.48)	.481 (.44)	.576 (.35)				

Table 7. (con't) *Comparison of Models for Expansion Investment in 1980-86*

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Holdout Sample Model 6	Entire Sample
Region 1	-.678 (.06)	-.656 (.07)	-.643 (.07)					
Region 2	-.140 (.67)	-.004 (.99)	-.072 (.82)					
Region 4	.302 (.40)	.403 (.26)	.391 (.27)					
City Distance	-.008 (.30)	-.009 (.26)						
Interest Rate	.002 (.98)							
---Model Statistics---								
Chi Square	51.84	52.88	51.62	45.30	40.33	41.94	32.53	63.72
Degrees Freedom	13	12	11	8	5	4	4	4
P Value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R	.24	.25	.25	.26	.25	.27	.22	.24
C Statistic	.72	.72	.72	.71	.70	.70	.67	.67
---Correct Classification Percentage---								
Total	64.8	66.9	65.3	64.1	63.1	64.1	62.6	62.9
Expanded	78.2	81.2	78.8	78.2	76.5	77.0	56.1	67.2
No Expansion	50.6	51.6	50.9	49.1	49.1	50.9	66.8	59.4

management. This shift towards increased importance for management at the expense of the education variable continued as other insignificant variables were dropped.

Only one of the goal variables was significant in Models 1 through 4. Dropping the set of goals had little impact on other model variables (Model 5).

Net income was marginally not significant in Models 1 through 5. Omitting net income materially affected only the intercept (Model 6). The final model (Model 6) statistics for the estimating sample indicate a high level of significance for the equation. The C statistic indicates that the model has some discriminatory power. However, the R of .27 indicates that a modest amount of the variability in the probability of investment is explained by the model. The model correctly classified 64 percent of the farms indicating some classification ability compared to the conditional model rate of 53 percent.[5]

Holdout sample results for the overall model were similar to those found for the estimating sample. The chi square declined significantly but continued to indicate a high level of significance for the entire equation. The R, C and total correct classification percentages declined modestly. However, the degree of significance of some of the variables included in the equation deteriorated significantly in moving from the estimating to the holdout sample. Age, which was highly significant with the estimating sample, became marginally insignificant. Education and risk which were significant in the estimating sample were highly insignificant in the holdout sample.

The reason for the large difference in the performance of the estimating and holdout samples is difficult to explain. The normal expectation, that this results from overfitting to the estimating sample, appears unwarranted. Estimating Model 1 using the holdout sample (Table 8) indicates that many of the differences were inherent in the samples drawn. The random procedure used to assign the observations resulted in two samples that were significantly different. While some of the differences in outcomes are likely due to overfitting, much appears to result from basic differences in the samples. Age, education and risk are all highly insignificant in the holdout sample.

An alternate explanation is that the low level of variability explained by the model could be explained by a variety of combinations of variables depending on relatively minor differences in the composition of observations included in the data set used to estimate the model. Either of these explanations imply that the results of the model should be used with care.

Projected Probabilities

The final model (Model 6) was re-estimated using the entire (estimating plus holdout) sample to obtain the best estimate of coefficient

---

5 Thirty-eight percent of all farms expanded (.38 x.38 + .62 x.62 = .529).

Table 8. *Comparison of Model 1 Results for Estimating and Holdout Samples 1980-86 Expansion Investment*

Variable	Estimating Sample	Holdout Sample
Intercept	-.896	-2.54
	(.57)	(.06)
1980 Net Cash Income ($000)	.004	.0003
	(.11)	(.88)
Age	-.029	-.016
	(.02)	(.14)
Region 1	-.678	-.494
	(.06)	(.15)
Region 2	-.140	-.248
	(.67)	(.43)
Region 4	.302	-.413
	(.40)	(.22)
Education	.105	-.024
	(.11)	(.70)
Management	.169	.472
	(.18)	(.00)
Interest Rate	.002	.171
	(.98)	(.02)
Distance to City	-.008	-.004
	(.30)	(.63)
Risk	1.001	-.028
	(.01)	(.93)
<u>Goals</u>		
Improve Family Living	.556	.341
	(.19)	(.39)
Increase Income	.825	.532
	(.07)	(.20)
Pass on Farm	.347	-.135
	(.60)	(.82)
Percent Equity		
	---Model Statistics---	
Chi Square	51.84	38.31
Degrees Freedom	13	13
P Value	0.00	0.00
R	.24	.16
C Statistic	.72	.68
	---Correct Classification---	
Total	64.8	63.1
Expansion	78.2	53.9
No Expansion	50.6	69.2

values. Those values were used to calculate expansion investment probabilities. The probability of investment increased modestly with operator age, increasing from 37 percent for those with 10 years of education to 52 percent for those with some graduate school level training (Figure 10).

Management index had a strong impact on investment probabilities (Figure 11). Those with the lowest index of management practices had a 30 percent probability of expanding while those with the highest value had a 52 percent likelihood of expansion investment. Farmers with greater risk tolerance were more likely to expand (52 percent) than farmers with average or above average risk aversion (40 percent, Figure 12).

The variable most strongly related to expansion investment is age (Figure 13). About 55 percent of the youngest farmers expanded their businesses, while only 31 percent of farmers who were 70 years of age would have been expected to expand during the prior seven year period.

Summary and Conclusions

Four basic models of investment behavior were developed to assess the importance of variables identified in the literature as influencing investment decisions of farmers. The basic models were:

1. A binary logit model of investment in a dairy farm heat recovery system.

2. A bivariate probit model of investment in a dairy farm precooler with equations for precooler investment and milking system selection.

3. An ordinal logit model to identify farmers who:

    a. Made no investment,
    b. Made replacement investments only, or
    c. Made net (expansion) investments during 1986, or 1985 and 1986.

4. A binary logit model of expansion versus no expansion during the 1980-86 period.

For each basic model, a number of alternate specifications were studied in an effort to determine which variables did and which did not influence investment behavior. When possible, holdout samples were used to identify the statistical properties of final model specifications.

Herd size, the existence of a parlor or pipeline system and education were contained in the final model specification for heat recovery. Holdout model statistics indicated a R value of .39 and a C statistic of .77 with correct classification rates of 67 percent.

The final version of the bivariate probit model of precooler investment contained herd size, management index, geographic region and cash income in the milking system equation and herd size and education in the precooler equation. Correct classification rates were 66 percent for

Figure 10. Probability of 1980-86 Expansion by Years of Education Upstate New York Farmers

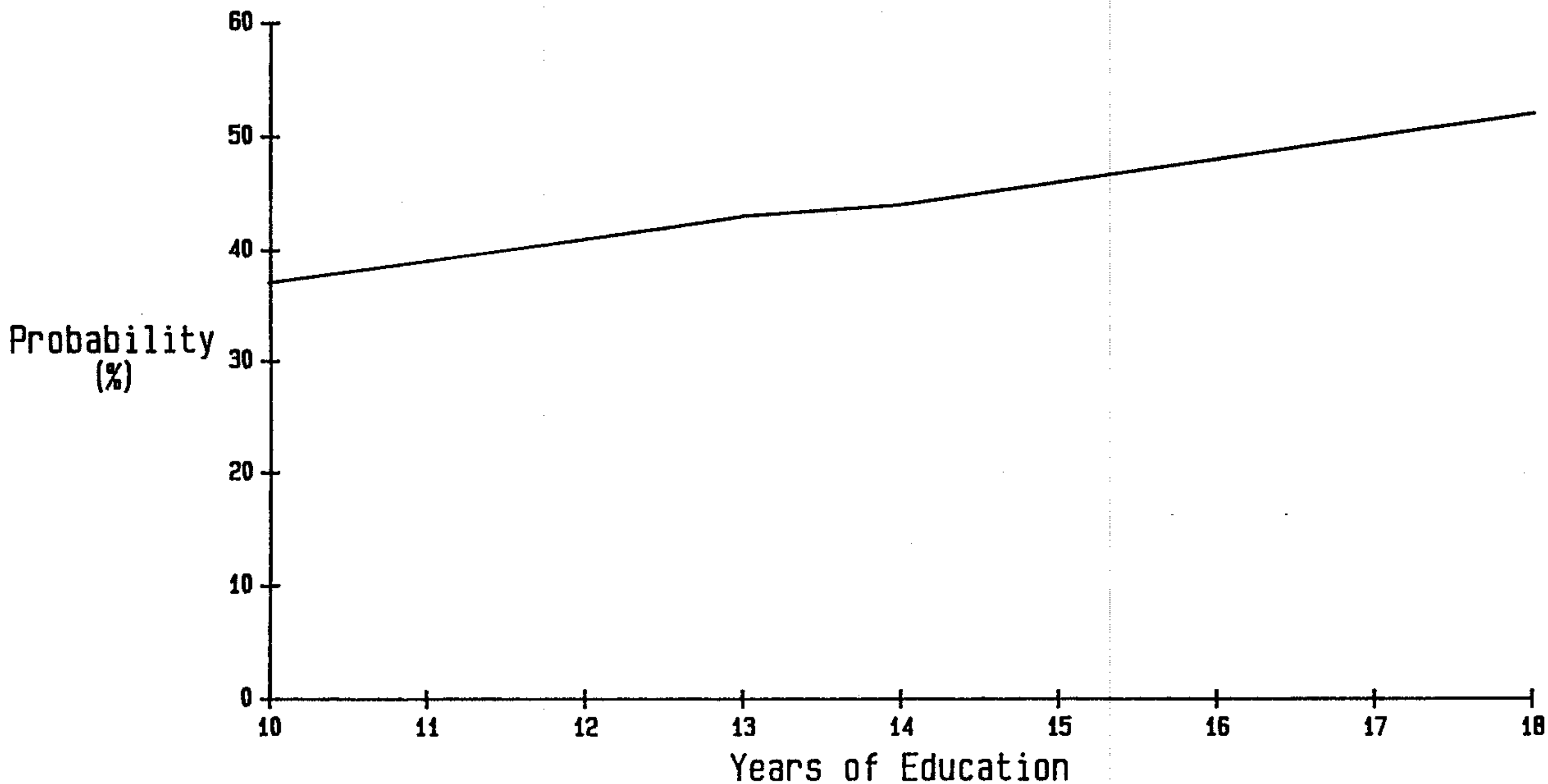


Figure 11. Probability of 1980-86 Expansion by Management Index Upstate New York Farmers

Figure 12. Probability of 1980-86 Expansion
by Risk Tolerance
Upstate New York State Farmers

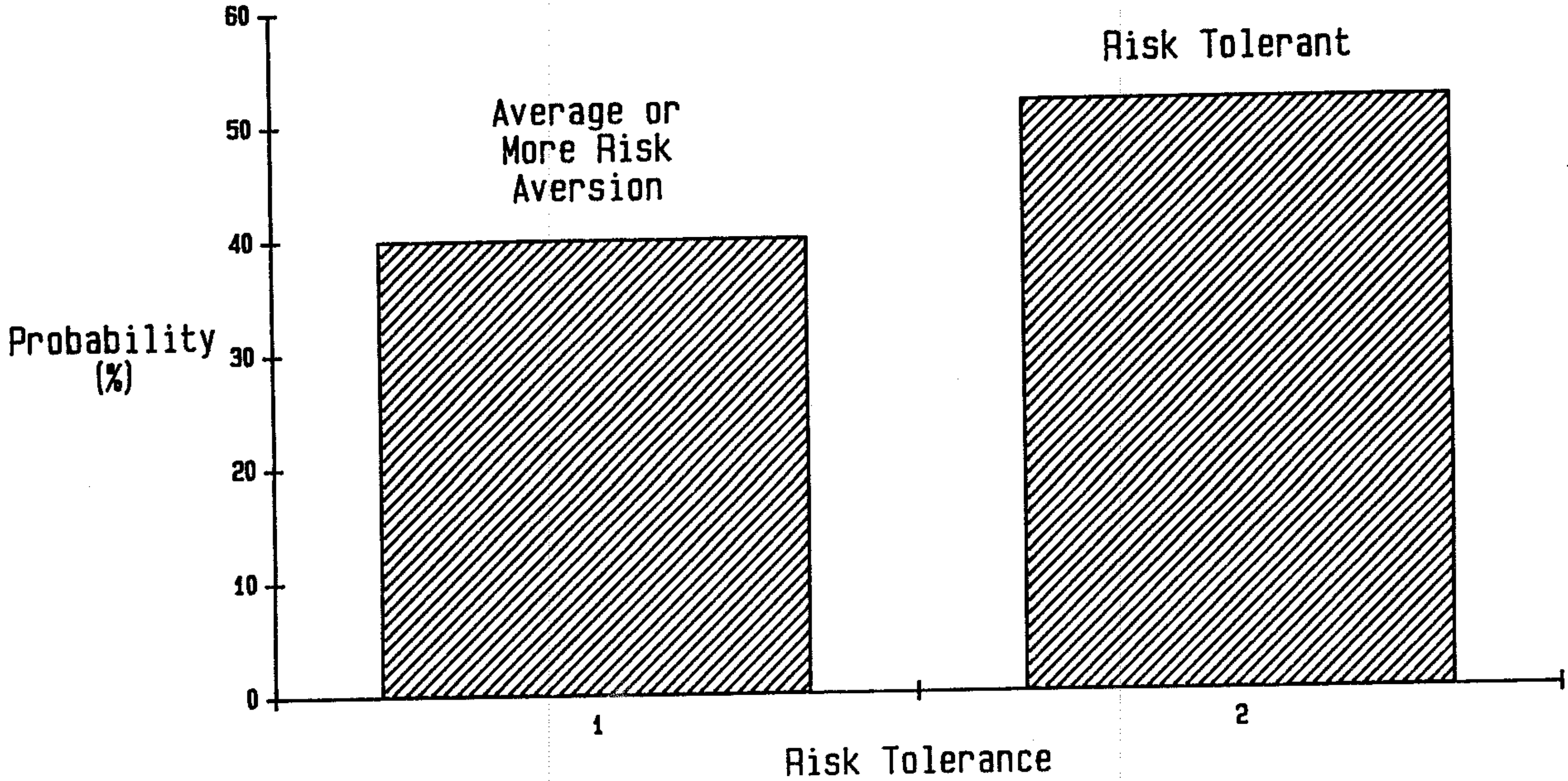
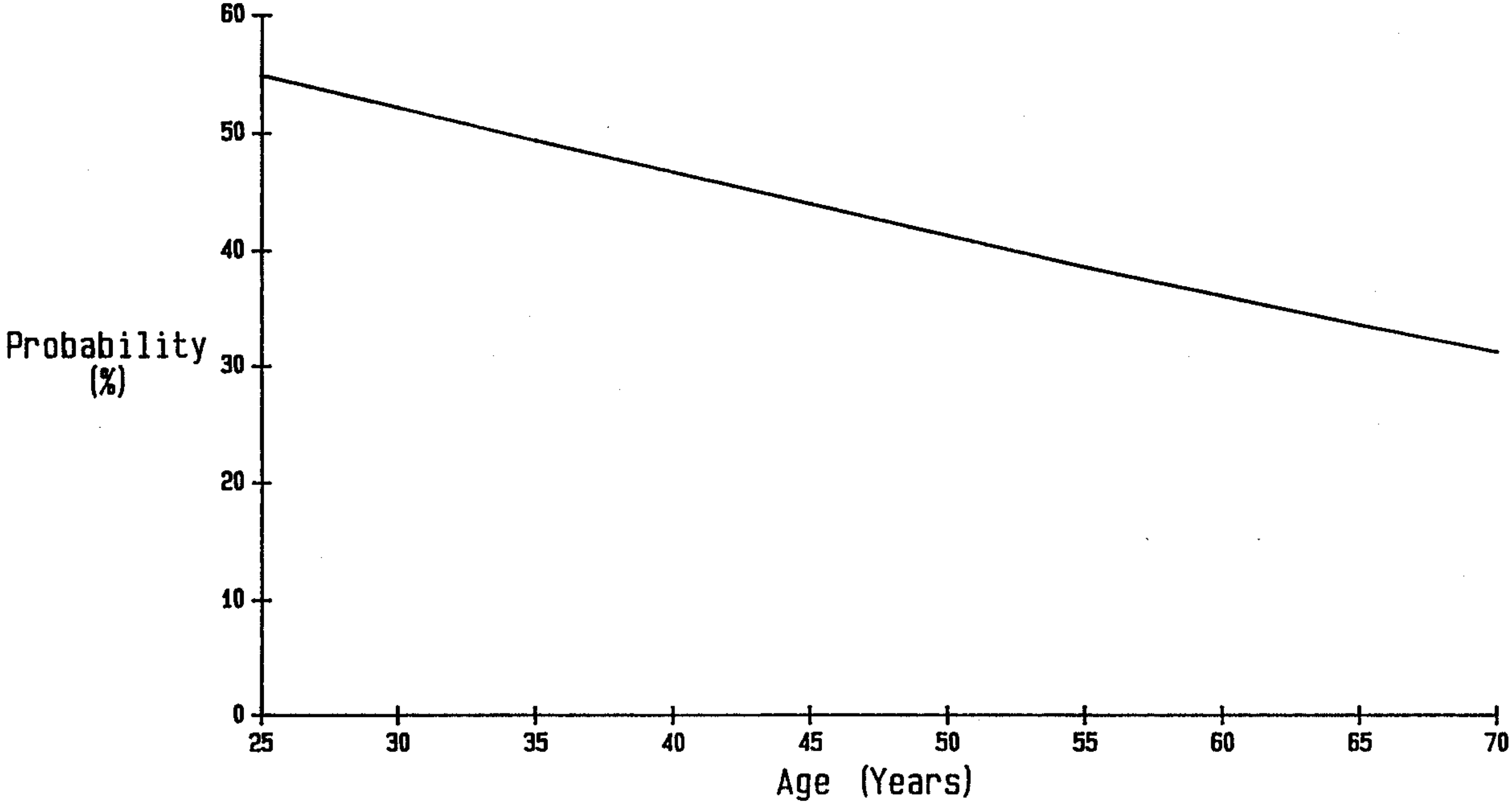


Figure 13. Probability of 1980-86 Expansion by Age
Upstate New York Farmers



the overall model, 85 percent for those with a precooler and 60 percent for those without.

The ordinal logit models contained gross income and operator age in the final equations for both 1986 investment and 1985 and 1986 investment. Although overall model significance was high, holdout sample R values were .13 and .14 and C statistics were .61 and .63.

Age, education, management index and risk were included in the final binary logit model of investment versus no investment during 1980-86. Holdout sample statistics included an R of .24 and a C of .67. The overall correct classification percentage was 63 percent.

Interest rate did not significantly influence investment behavior in any of the models developed. This is consistent with the findings of Gustafson et al. For the specific investment items, heat recovery systems and precoolers, the insignificance likely resulted from the low initial investment required and, thus, the relatively low effect of the interest rate on investment profitability. Interest rate is also confounded by the existence of Farmers Home Administration limited resource programs which provide low interest rate loans to farmers who are frequently limited in the amount of further investment they can make. Also, major investments may force farmers further out on the credit supply curve resulting in more dealer or other higher cost credit. The insignificance of interest rate for expansion investments is also influenced by the combined effect of the long term nature of expansion investments and variable interest rates. The rate of interest a farmer faces at the time of investment is not nearly as important for farmers charged variable rates as is the expected average rate over the life of the investment.

Distance to a city of 20,000 or more population was not related to individual item nor general expansion investment. Possible explanations for this include: (1) only investment very close to the city (i.e., within a mile or two), is affected and this is not of sufficient magnitude to support a general relationship, (2) proximity results in investment substitution (i.e., more cows and equipment and fewer buildings), rather than reduced investment, and (3) reduced investment results from ownership discontinuity rather than lower levels of investment by individual farmers. That is, retiring farmers sell to developers or speculators who rent out the land and buildings to existing farmers who expand their home operations but make little investment in the rented property. Such land could also be used as starter units for beginning farmers who invest in livestock and machinery (investment substitution).

Goals were not related to investment behavior in any of the models developed. This may have resulted from the lack of relationship between goals and investment or inability of the data collected to properly elicit goals. Investment may be one approach to achieving a wide variety of goals. For example, those whose primary goal is to increase leisure time may expand the business to bring in a partner who can handle the business while the investor is away from the farm. Similarly, a focus on improved family living may require investment to increase income to allow increased leisure.

Data about goals correspond to the period after the investment, if investment took place. Farmers were interviewed in 1987 about investment prior to that time and were asked about their goals for the next 10 years. Further, farmers who planned to exit farming in the next 10 years were not asked their goals but were ascribed the goal of leaving farming. In effect this implied that farmers who plan to leave farming during the next 10 to 16 years will not invest. Given the lives of most farm assets, this is clearly inappropriate. Future research should redefine the goal categories. Although widely used in prior research, the categories are not clearly mutually exclusive.

Income expectations defined as the expected direction of change in income was consistently unimportant in investment behavior. It may be that the direction of change in income is just so much less important than the level of income that it has little impact on investment. However, the level of net income was also generally unrelated to investment. It was an element in the parlor/pipeline equation for precooler investment, but not important in precooler investment per se or in any other equations. It appears that low income may necessitate investment to improve economic well being frequently enough to offset the facilitative effects of cash flow on investment ability and the profit incentive of making added investments in a profitable businesses. Inaccurate recall may contribute noise to the analysis, since farmers were asked to indicate their 1980 net income (prior to investment) in relation to their 1986 net income.

Region was generally unrelated to investment behavior. It was significant only in the pipeline/parlor equation of the precooler model. Although the level of investment differs by region, the effect of any soil, climate or other characteristics that differ by region were apparently sufficiently represented by other variables.

The degree of risk aversion was not related to investment in specific items such as the heat recovery or precooler. However, it may be related to the more general question of whether expansion investment occurs.

Education was important in determining investment in heat recovery or precooler systems but not important in whether expansion investment is made. Thus, it appears that education does contribute to investment in new technology where understanding of technical or complicated factors may be important, but is not important in determining added investment in items with which all farmers could be expected to be familiar.

The management index was important only in the parlor/pipeline equation of the precooler model and the 1980-86 investment model. As indicated by the two precooler model equations, there is likely some interaction or trade-off between education and management. The measure of management ability used here focused on record keeping, buying decision processes and technology adoption. Development of improved measures of management ability might change the results obtained.

Age was not important in determining investment in the relatively low cost new technology represented by heat recovery or precooler systems. However, it was one of the most important variables in determining whether expansion investment occurred. This appears to support the life cycle hypothesis or the contention of Gustafson et al., that farmers seek an even pace of investment.

Size of farm, measured by herd size or gross income, was most consistently related to investment. Larger business appear to provide the base for expansion and are operated by managers with interest in and confidence in their ability to manage expansion.

Appendix A

Details of Precooler Model Cross-Validation Analysis

The tenfold cross-validation analyses run for the final precooler model (Model 6) with a probability cutoff of .224 are presented in Table A1.

Viewed as a predictive model, a cutoff probability other than .224 could be used if it improved the correct classification rate. The correct classification rate is much higher with a cutoff of .50 (Table A2). However, the correct classification rate for those with a precooler is very low indicating a low rate of success for the primary use of the model. An optimum cutoff rate could be determined that either maximized the classification efficiency or resulted in balanced error rates for those with or without precoolers.

Table A1. Results from Cross-Validation Analyses of Precooler Model

Variable (P Value)	Sample 1	2	3	4	5	6	7	8	9	10	Ave.
Precooler Equation:											
Intercept	-2.925	-2.769	-3.349	-2.521	-2.777	-3.049	-2.516	-2.30	-2.71	-2.691	
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	
Cows	.929	.939	1.240	.898	.866	.931	.694	.904	.881	.731	
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.20)	(.00)	(.00)	(.00)	
Cows^2	-.085	-.087	-.116	-.088	-.079	-.086	-.029	-.084	-.080	-.064	
	(.00)	(.00)	(.00)	(.01)	(.01)	(.00)	(.82)	(.01)	(.01)	(.02)	
Education	.110	.102	.113	.088	.105	.119	.099	.664	.10	.110	
	(.01)	(.01)	(.01)	(.03)	(.01)	(.00)	(.02)	(.11)	(.01)	(.01)	
Milking System Equation:											
Intercept	-1.751	-1.665	-1.682	-1.681	-2.168	-1.799	-1.861	-1.740	-1.769	-1.633	
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	
Cows	2.317	2.279	2.016	2.114	2.659	2.221	2.394	2.250	2.258	2.13	
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	
Region 1	.70	.617	.747	.560	.726	.755	.562	.627	.791	.549	
	(.00)	(.01)	(.00)	(.02)	(.00)	(.00)	(.02)	(.01)	(.00)	(.02)	
Region 2	.609	.504	.529	.463	.615	.529	.451	.510	.491	.488	
	(.00)	(.01)	(.01)	(.02)	(.00)	(.01)	(.03)	(.01)	(.01)	(.01)	
Region 4	.439	.429	.503	.395	.519	.503	.356	.491	.477	.436	
	(.06)	(.07)	(.04)	(.09)	(.02)	(.02)	(.15)	(.03)	(.04)	(.06)	
Management	.237	.226	.244	.280	.318	.275	.274	.262	.278	.262	
	(.01)	(.01)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)	
Cash Income (100,000)	.528	.455	.581	.529	.712	.472	.575	.381	.459	.525	
	(.15)	(.20)	(.11)	(.15)	(.10)	(.19)	(.11)	(.28)	(.22)	(.14)	
Rho	-.703	-.737	-.438	-.761	-.804	-.734	-.807	-.813	-.744	-.849	
Log Likelihood	-302.4	-316.5	-310.7	-312.5	-295.4	-315.6	-309.5	-306.2	-310.6	-308.5	
No. of Observations	356	363	359	353	352	366	356	356	361	360	

Table A1.(con't) Results from Cross-Validation Analyses of Precooler Model

Variable (P Value)	Sample 1	2	3	4	5	6	7	8	9	10	Ave.
	---Correct Classification Percentages---										
Estimating Sample:											
Total	69.7	65.7	75.2	64.0	66.5	67.8	60.1	66.9	65.7	43.6	64.5
With Precooler	81.3	86.6	81.3	86.6	85.5	84.0	89.5	89.5	85.0	96.1	86.5
Without Precooler	66.3	59.6	73.7	57.2	61.2	63.2	50.7	60.7	60.1	29.6	58.2
Model Efficiency	69.7	65.7	65.4	63.8	66.5	67.8	59.5	67.1	65.6	44.5	63.6
Holdout Sample:											
Total	64.3	60.0	69.2	62.2	63.0	68.8	28.6	69.0	64.9	63.2	61.3
With Precooler	88.9	100	37.5	100	84.6	62.5	37.5	69.2	100	84.6	76.5
Without Precooler	57.6	51.7	77.4	55.3	54.5	70.8	26.5	69.0	53.6	52.0	56.8
Class Efficiency	64.6	62.5	68.5	62.2	61.3	69.0	29.0	69.1	64.0	59.4	61.0
No. of Observations	42	35	39	45	46	32	42	42	37	38	

[a] P Values are in parentheses.

Table A2. Cross-Validation Correct Classification Percentages with Cutoff of .5
Precooler Investment Model 6

Sample No.	Estimating Sample			Hold-Out Sample		
	With Precooler	Without Precooler	Model Efficiency	With Precooler	Without Precooler	Model Efficiency
1	35.0	93.8	80.6	22.2	97.0	80.2
2	32.5	95.0	80.7	66.7	86.2	81.8
3	37.0	93.9	81.1	25.0	87.1	73.1
4	35.4	94.1	80.5	14.3	100	80.1
5	31.6	94.9	81.3	21.6	78.4	83.6
6	35.8	94.4	81.4	12.5	100	80.1
7	36.0	94.8	80.6	46.2	100	87.0
8	31.6	94.3	80.9	46.2	100	88.5
9	33.8	94.7	81.2	33.3	92.9	79.7
10	34.2	95.1	82.2	30.8	100	85.4
Average	34.3	94.5	81.1	31.9	94.2	82.0

References

Beaver, W.H. "Financial Ratios as Predictors of Failure." Journal of Accounting Research Supplement. 4(1966):71-127.

Bischoff, C.W. "Business Investment in the 1970s: A Comparison Models." Brookings Papers on Economic Activity. (1971) 1:13-58.

Brase, B.T. and E.L. LaDue. "Farmer Investment Behavior: A Review of Literature." Department of Agricultural Economics Research Bulletin 89-5, Cornell University, March 1988.

Clark, P.K. "Investment in the 1970s: Theory, Performance and prediction." Brookings Papers on Economic Activity. (1979) 73-113.

Dixon, B.L., L.D. Hill and M.S. Saffell. "The Relationship of Farmers' Intentions to Purchase and Their Actual Purchases of Metal Bin Grain Storages." North Central Journal of Agricultural Economics. (1980) 2(1):61-68.

Efron, B. "Estimating the Error Rate of a Prediction Rule: Improvement on Cross-Validation." Journal of the American Statistical Association. Vol. 78, No. 382, June 1983.

Fisher, B.S. "A Quarterly Model of Agricultural Investment in Australia." Australian Journal of Agricultural Economics. (1974) 18(1):21-31.

Frydman, H., E.I. Altman and D.L. Kao. "Introducing Recursive Partitioning for Financial Classification: The Case of Financial Distress." The Journal of Finance. Vol. XL, No. 1, pp. 269-90, March 1985.

Funk, T.F. "Farmer Buying Behavior: An Integrated Review of Literature." Working Paper AE-72-16, University of Guelph, 1972.

Greene, W.H. "LIMDEP, Version 5." New York: William H. Greene 1988, 312 pp.

Gustafson, C.R., P.J. Barry and S.T. Sonka. "Machinery Investment Decisions by Cash Grain Farmers in Illinois." Paper presented to NC-161 Regional Research Project, "Evaluating Financial Markets for Agriculture." St. Paul, MN. October 7-8, 1986.

Hanley, J.A. and B.J. McNeil. "The Meaning and Use of the Area Under the Receiver Operating Characteristic Curve." Diagnostic Radiology. Vol. 143, No. 1, pp. 29-36, April 1982.

Harrell, F.E., Jr. "The Logist Procedure." Chapter 23 in SAS Institute, Inc. SUGI Supplemental Library User's Guide, Version 5 Edition. Cary, NC: SAS Institute, Inc. 1986.

Heckman, J.J. "Sample Selection Bias as a Specification Error." Econometrica. Vol. 47, No. 1, pp. 153-61, January 1979.

Hill, L.D. and P. Kau. "Application of Multivariate Probit to a Threshold Model of Grain Dryer Purchasing Decisions." American Journal of Agricultural Economics. (1973) 65(3):553-57.

Johnson, T.G., W.J. Brown and K. O'Grady. "A Multivariate Analysis of Factors Influencing Farm Machinery Purchase Decisions." Western Journal of Agricultural Economics. (1985) 10(2):294-306.

Jorgenson, D. "The Theory of Investment Behavior." Determinants of Investment Behavior. R. Ferber, Ed. Columbia University Press, New York 1967.

Kelleher, M.J. and N.L. Bills. "Statistical Summary of the 1987 Farm Management and Energy Survey." Department of Agricultural Economics, Cornell University, A.E. Res. 89-3, February 1989.

LaDue, E.L. and J.H. Kwiatkowski. "An Analysis of the Investment Related Characteristics of New York Farmers." Department of Agricultural Economics, Cornell University, A.E. Res. 89-18, September 1989.

Lamm, R.M., Jr. "Investment in Agriculture: An Empirical Analysis." Agricultural Finance Review. (1982) 42:16-23.

LeBlanc, M. and J. Hrubovcak. "Tax Policy Aggregate Investment." American Journal of Agricultural Economics. (1986) 68(4):767-77.

Maddala, G.S. "Limited-Dependent and Quoliterative Variables in Econometrics." Econometric Society Monographs, No. 3, Cambridge University Press, Cambridge, 1983.

Penson, J.B., Jr., R.F.J. Romain and D.W. Hughes. "Net Investment in Farm Tractors: An Econometric Analysis." American Journal of Agricultural Economics. (1981) 63(4):629-35.

Picard, R.R. and R.D. Cook. "Cross-Validation of Regression Models." Journal of the America Statistical Association. Vol. 79, No. 387, pp. 575-83, September 1984.

Smith, S.F., W.A. Knoblauch and L.D. Putnam. "Dairy Farm Business Summary, New York." Department of Agricultural Economics, Cornell University, A.E. Res. 88-8, July 1988 and prior issues.

VanDeVen, W.P.P.M. and M.S.B. VanPraag. "The Demand for Deductibles in Private Health Insurance, A Probit Model with Sample Selection." Journal of Econometrics. 17(1981):229-52.

Walker, S.H. and D.B. Duncan." Estimation of the Probability of an Event as a Function of Several Independent Variables." Biometrika. (1967)54:1 and 2, pp. 167-78.

Other Agricultural Economics Working Papers

No. 88-13	The Economics of Sustainable Agricultural Systems in Developing Countries	Randolph Barker Duane Chapman
No. 89-1	Resource Economics: Five Easy Pieces	Jon M. Conrad
No. 89-2	Optimal Dairy Policy with bovine Somatotropin	Loren W. Tauer Harry M. Kaiser
No. 89-3	Environmental Standards and International Trade in Automobiles and Copper: The Case for a Social Tariff	Duane Chapman
No. 89-4	Optimal Fluid Milk Advertising in New York State: A Control Model	Donald J. Liu Olan D. Forker
No. 89-5	A Bioeconomic Model of the Harp Seal in the Northwest Atlantic	Jon Conrad Trond Bjorndal
No. 89-6	A Money Metric Measure of Welfare Change From Multiple Price and Income Changes	Jesus C. Dumagan
No. 89-7	The Strategic Role of Supermarket Buyer Intermediaries in New Product Selection: Implications for Systemwide Efficiency	Edward W. McLaughlin Vithala R. Rao
No. 89-8	Estimating Parameters in a Dynamic Open-Access Model With an Application to the Flemish Cap Ground Fishery	Nick Gomersall
No. 89-9	The Dynamic Effects of Agricultural Subsidy Programs in the United States	Harry de Gorter Eric O'N. Fisher
No. 89-10	Estimating Sources of Fluctuations in the Australian Wool Market: An Application of VAR Methods	Robert J. Myers R. R. Piggott William G. Tomek