Entry, Exit, and Structural Change in Pennsylvania's Dairy Sector

Jeffrey R. Stokes

Data on the number of Pennsylvania dairy farms by size category are analyzed in a Markov chain setting to determine factors affecting entry, exit, expansion, and contraction within the sector. Milk prices, milk price volatility, land prices, policy, and cow productivity all impact structural change in Pennsylvania's dairy sector. Stochastic simulation analysis suggests that the number of dairy farms in Pennsylvania will likely fall by only 2.0 percent to 2.5 percent annually over the next 20 years, indicating that dairy farming in Pennsylvania is likely to be a significant enterprise for the state in the foreseeable future.

Key Words: dairy, maximum entropy, farm size, Markov chain, simulation

The structure of American agriculture continues to evolve toward fewer, larger farming units. For example, according to the U.S. Department of Agriculture, since 1940 the number of farms in the United States has declined by about 66 percent, while over the same period average farm size (in acres) increased by about 161 percent. In Pennsylvania, the number of dairy farms operating within the state was about 22,000 in 1980 and about 9,600 in 2003, representing a decline of 56 percent (USDA, various issues). The declining number of dairy farms and questions regarding what the future holds for dairying in Pennsylvania are critically important for policymakers, agribusinesses, dairy producers, and those involved in higher education through research, extension, and teaching. States like Pennsylvania perhaps have more at stake regarding the issue, given the predominance of a single commodity (i.e., dairy).

While not often thought of as an agricultural state, Pennsylvania's dairy sector is sizeable. In 2003, Pennsylvania dairy farms managed about 575,000 milk cows that produced a total of 10.3 billion pounds of milk. Both of these statistics make Pennsylvania the fourth largest dairy state in the nation with respect to cow numbers and milk production. However, dairy production in Pennsylvania oftentimes co-exists with a population that is less than enthusiastic about large dairy farms and, somewhat paradoxically, less than enthusiastic about fewer dairy farms as well.

Much of the state's dairy cows and production are in the Chesapeake Bay watershed, with high concentrations in the southeastern part of the state near large population centers. The nature of the relationship between production agriculture and the urban fringe in Pennsylvania has given rise to numerous legislative efforts to protect the environment as well as agriculture. For example, the U.S. Environmental Protection Agency's Air Quality Consent Agreement and Concentrated Animal Feeding Operation regulations are clearly designed to protect the environment. Programs such as the Pennsylvania Department of Agriculture's Agricultural, Communities and Rural Environment Initiative (ACRE), farmland preservation programs, and the Next Generation Farmer Loan Program are all designed to protect current and future generations of agricultural producers in Pennsylvania.

The future of dairy farming in Pennsylvania and Pennsylvania's role as a leading dairy state are important issues for others as well. Agribusinesses have a stake in the livelihood of the state's dairy sector through the inputs they provide to the sector such as purchased feed, pharmaceuticals, and machinery and equipment. In addition, the state has a large investment in milk processing infrastructure, which employs many of the state's rural residents. Dairy producers themselves have an interest in better understanding the drivers of structural change in the sector since many are facing critical issues such as whether or not to

Jeffrey Stokes is Associate Professor in the Department of Agricultural Economics and Rural Sociology at the Pennsylvania State University in University Park, Pennsylvania.

expand, when to retire, and how best to facilitate intergenerational transfer. Perhaps one of the most concerned entities is Pennsylvania State University's College of Agricultural Sciences, which has a commitment to conduct research and provide resident and non-resident education that is in large part related to dairy farming.

As consolidation continues, there is a need to identify and better understand the factors that influence structural change. Additionally, quantifying the extent of consolidation is a critical first step toward better understanding where the sector is headed in terms of the number and typical size of farming operations. Exploring policy options that can potentially influence the rate of consolidation in Pennsylvania's dairy sector will be easier if a better understanding of the dynamics of farm size can be obtained as a result.

Therefore, the objectives of this research are threefold. First, the primary drivers of structural change in Pennsylvania's dairy sector are identified. These are factors that influence entry into dairying, exit from dairying, and consolidation within the sector. A second objective is to investigate what the future may hold for Pennsylvania's dairy sector in terms of the geographic distribution of dairy farms within the state and the total number of dairy farms in the state. If such an analysis suggests that in all likelihood there will be few dairy farms in Pennsylvania in the future if the status quo is maintained, resources can be allocated today to increase or decrease the likelihood of such an outcome, depending on society's goals.

To accomplish these objectives, a Markov chain model of farm size is employed. However, unlike all past research that makes use of this methodology, dating to at least Padberg (1962), the Markov chain model developed here is shown to be an appropriate methodology under a minimal set of assumptions regarding the producer's objectives and sources of risk facing the business. This leads to the third objective of this research, which is to demonstrate the conditions under which the Markov chain model is appropriate for analyzing farm size adjustment over time.

Markov Processes and Farm Size

Padberg (1962) was the first to suggest the use of Markov chains as a methodology for better understanding firm size dynamics in agriculture by applying the methodology to California fluid milk processors. Since then, numerous researchers have applied the Markov chain model in various agricultural settings. Lee, Judge, and Takayama (1965), Disney, Duffy, and Hardy (1988), and Gillespie and Fulton (2001) all examine hog operations, while Hallberg (1969), Ethridge, Roy, and Myers (1985) and Kim, Lin, and Leath (1991) examine dairy processing, cotton ginning, and flour milling agribusinesses, respectively. Other uses of the Markov model include Chan (1981) for the Canadian cattle industry and Garcia, Offutt, and Sonka (1987) for Illinois crop farms. Chavas and Magrand (1988), Zepeda (1995a), Zepeda (1995b), and Rahelizatovo and Gillespie (1999) apply the methodology to dairy farms in Wisconsin, Louisiana, and other multi-state regions of the United States.

It is important to point out that in all of these previous studies, it is assumed that farm or firm size is a stochastic process that possesses the Markov property. A stationary (continuous time) Markov chain is a stochastic process $X(t)$ with the property that the conditional distribution of the future, $X(t+s)$ $s > t$, given the present, X(t), and the past, X(u) $u < t$, depends only on the present and is independent of the past. This is why the Markov property is often termed the memoryless property. Mathematically, this implies $pr\{X(t+s) = j\,|\,X(t) = i, X(u) = x(u), 0 \leq u < t\} = pr\{X(t+s) = j\,|\,X(s) = i\} = p_{ij}$, where i and j are states frequented by the process.[1] In the context of firm size, the states making up the Markov chain are size categories. For farms, acres or the number of head of livestock typically define the states so that the transition probabilities suggest the probability of a farm in size category i at time t moving to size category j at time $t+1$.

While it is natural to think of firm size as a Markov process, perhaps a more desirable setting is one in which firm size is shown to be Markovian through some properly specified structural model. Such an approach also offers the advantage of allowing the structural model to guide the researcher in terms of the identification of the appropriate economic variables that influence the state transition probabilities.

[1] Non-stationary Markov chains are also possible where oftentimes economic variables are the source of the non-stationarity.

As it turns out, a relatively simple stochastic optimal control model can be used to demonstrate the conditions under which firm size is Markovian. Such a model is presented in the Appendix. While the model is loosely cast in the context of a dairy farm, its basic structure lends itself to any type of firm that makes decisions in a dynamic framework. In addition, the model dramatically simplifies the decision environment facing an actual dairy farmer. The reason for this simplification is to identify a minimal set of assumptions that justify the Markov chain model as an appropriate approach for investigating firm size dynamics. A secondary goal of the model is to show how the framework lends insight into the relevant state variables that affect transition probabilities. This result has important empirical implications for estimating non-stationary transition probabilities, as shown in a subsequent section of the paper.

Empirical Prerequisites

If micro-level data are available, the estimation of transition probabilities for each time period for which data are available can be accomplished via the method of maximum likelihood (ML). Hallberg (1969) did this and presented a test for determining whether the transition probabilities were stationary or non-stationary. Stationary transition probabilities are simply the average of the estimated (periodic) transition probabilities. In later work, Rahelizatovo and Gillespie (1999) also used ML to estimate non-stationary transition probabilities. Hallberg's (1969) and Rahelizatovo and Gillespie's (1999) work are important because they represent the only attempt at explaining non-stationary transition probabilities by regressing the estimated probabilities on explanatory variables. Hallberg (1969) and Rahelizatovo and Gillespie (1999) accomplished this because they had access to micro-level data which permitted the estimation of the transition probabilities as a first step.

The problem often faced by researchers is the lack of micro-level data. For example, to use the ML method for a farm size problem, the number of farms that moved from each state to other states, the number that entered farming, and the number that exited would all have to be known at each point in time. At best, it is typically the case that only the total number of farms in each size category is available from which the net change (typically net exit) can be determined by comparing the number of farms at time t with that of $t+1$.

To circumvent the problem created by a lack of micro-level data, aggregate share data are sometimes used to estimate transition probabilities. Estimation of transition probabilities from aggregate share data has its origins in the work by Lee, Judge, and Takayama (1965) as well as in later work by Lee, Judge, and Zellner (1973). Much of that work uncovers properties and problems associated with various parametric estimators of transition probabilities when using share data. One of the most important problems facing the applied researcher when using parametric techniques is the fact that negative degrees of freedom can be encountered even in the stationary setting if the state space of the Markov chain is large and/or the number of time-ordered observations from which estimation is conducted is small. The non-stationary case induced by adding explanatory variables into the estimation poses even more problems with degrees of freedom because parameters would presumably need to be estimated for at least the majority of the probabilities in the transition probability matrix.

One way to minimize the impact of degree of freedom problems is to use maximum entropy (ME) estimation, suggested by Golan, Judge, and Miller (1996) and Lee and Judge (1996). The ME estimation is a nonparametric technique and is useful when the data are limiting or the economic problem underlying the need for estimation is ill-posed. All that is required is that the underlying stochastic process under study be a Markov process. For the farm size problem, this assumption is plausible given the results presented in the Appendix.

Estimation of non-stationary transition probabilities is best approached from a cross-entropy formulation, which makes use of a prior as shown below (see Golan, Judge, and Miller 1996). The formulation for a given t is

(1)
$$\min_{\pi_{ijm}(t)} \Psi(\boldsymbol{\eta},\boldsymbol{\pi}) = \sum_t \sum_i \sum_j \sum_m \pi_{ijm}(t) \ln\left(\pi_{ijm}(t) \big/ \eta_{ijm} \right),$$

subject to

$$(2)\quad \begin{aligned} & y_j(t+1) - \sum_i y_i(t) \sum_m \delta_m \pi_{ijm}(t) = 0 \;\; \forall \;\; j \\ & 1 - \sum_j \sum_m \delta_m \pi_{ijm}(t) = 0 \;\; \forall \;\; i \\ & 1 - \sum_m \pi_{ijm}(t) = 0 \;\; \forall \;\; i, j \\ & \pi_{ijm}(t) \geq 0 \;\; \forall \;\; i, j, m \; . \end{aligned}$$

The matrix π (η) with elements $\pi_{ijm}(t)$ (η_{ijm}) is a matrix of non-stationary transition probabilities (prior transition probabilities) to be estimated, and $\Psi(\eta,\pi)$ is the Shannon entropy measure. The first constraint in (2) represents the Markov relation where $y_j(t)$ is the proportion of farms in the jth state of the Markov chain at time t and δ_m is a vector of parameter supports. The parameter support vector allows for the estimation of a discrete distribution of transition probability estimates (one for each parameter support value). The second constraint is the row sum condition which ensures that the underlying Markov process does not leave the discrete state space. The third constraint ensures that the discrete distribution of transition probabilities sums up to one, while the last constraint represents the non-negativity condition imposed on the probabilities. Notice that the specification does not require the addition of an error term on the Markov relation since a matrix of transition probabilities is estimated from each transition of the data.

While the cross-entropy formulation presented above could be expanded to accommodate a set of state variables that influence the dynamics of the system under study, an alternative approach is to use the entropy-generated transition probabilities in a parametric setting. In the Appendix, it is shown that the transition probabilities depend on a collection of relevant state variables, $\mathbf{s(t)}$, so that $\pi_{ij}[\mathbf{s(t)}] = \pi_{ij}(t)$, and it is the state variables that induce the nonstationarity of the transition probabilities. Therefore, the cross-entropy model in (1) subject to (2) can be used to estimate a matrix of transition probabilities for each transition of the data. These data can then be used along with data on the state variables in a parametric setting to estimate the influence that each state variable has on the associated entropy estimated transition probabilities. The advantage of this approach is that standard tests of significance can then be applied for identifying the factors affecting the transition probabilities.

As noted above, the $\boldsymbol{\eta}$ matrix is a matrix of prior probabilities. While numerous methods could be employed to uncover a prior, it would seem that one sensible method would be to estimate a prior consistent with a stationary Markov chain, that is, with all the probabilities fixed through time. An entropy formalization of the stationary Markov problem is

$$(3)\quad \min_{\eta_{ijm}, \lambda_{tjm}} \quad \Psi(\boldsymbol{\eta}) = \sum_i \sum_j \sum_m \eta_{ijm} \ln(\eta_{ijm}) + \sum_t \sum_j \sum_m \lambda_{tjm} \ln(\lambda_{tjm}),$$

subject to

$$(4)\quad \begin{aligned} & y_j(t+1) - \sum_i y_i(t) \sum_m \delta_m \eta_{ijm} + \sum_m \gamma_m \lambda_{tjm} = 0 \;\; \forall \;\; t, j \\ & 1 - \sum_j \sum_m \delta_m \eta_{ijm} = 0 \;\; \forall \;\; i \\ & 1 - \sum_m \eta_{ijm} = 0 \;\; \forall \;\; i, j \\ & 1 - \sum_m \lambda_{tjm} = 0 \;\; \forall \;\; t, j \\ & \eta_{ijm} \geq 0, \; \lambda_{tjm} \geq 0. \end{aligned}$$

In (3) subject to (4), we seek a discrete distribution of transition probabilities and allow for the possibility that there is error in the Markov relation by including an error term expressed as the product of an error support vector, γ_m, and errors, λ_{tjm}. Of course, a prior is required for this specification as well, and by default, a uniform prior is assumed.

Once the prior matrix of stationary transition probabilities is estimated, non-stationary transition probabilities can then be estimated, and it remains to empirically relate them to the vector of state variables, $\mathbf{s(t)}$, hypothesized to impact the transition probabilities over time. For this step, Zellner's (1962) Seemingly Unrelated Regression (SUR) is used as in Rahelizatovo and Gillespie (1999). Each transition probability of interest is regressed against the set of state variables. More formally,

$$(5)\quad \Phi^{-1}\left(\pi_{ij}(t)\right) = \beta_{0,ij} + \sum_{k=1}^{K} \beta_{k,ij} s_k(t) + \varepsilon_{ij}(t) \, ,$$

where it is assumed that the regression disturbances in different equations are mutually correlated. In (5), $\Phi^{-1}(\cdot)$ represents the inverse of the standard normal distribution and is used to transform the probabilities from the zero-one interval to a continuous interval. It is important to note that even with the current ME approach, degree of freedom problems can preclude investigating all the probabilities in the matrix, as shown in the next section of the paper.

Empirical Application

Empirical estimation of non-stationary transition probabilities using the cross-entropy formulation presented in the preceding section was conducted using USDA data on the number of Pennsylvania dairy farms in various size categories from 1980 to 2003. One potential problem is that the size categories used by USDA sometimes change over time. For example, prior to 1997, no data were kept on the number of dairy farms with 500 or more head. These data were simply reported as the number of dairy farms with 200 or more head. Similarly, prior to 1993, no data were kept on the number of dairy farms with 200 or more head. These data were reported as the number of dairy farms with 100 or more head. In some settings, it may be appropriate to simply combine the data in the 500+ and 200–499 head categories now reported in the 100+ head category (now 100–199 head). However, the emphasis of this research is on quantifying the progression of farm size and the identification of variables that influence it. Therefore, it would seem that better information may be obtainable by recovering the missing data rather than by limiting the number of states in the Markov chain.

To this end, a simple entropy-based model was specified to let the data determine the most likely number of 500+ head dairy farms prior to 1997 and the most likely number of 200–499 head dairy farms prior to 1993 from the totals that were actually reported. The cross-entropy model specified is

(6)
$$\min_{\phi_{ij},\tilde{y}_{im}} \quad \Psi(\gamma) = \sum_i \sum_j \phi_{ij} \ln\left(\phi_{ij}\right) + \sum_{i\geq z} \sum_m \tilde{y}_{im} \ln\left(\tilde{y}_{im}\right),$$

subject to

(7)
$$y_j(t+1) - \sum_i y_i(t)\phi_{ij} - \sum_{i\geq z}\sum_m \delta_m \tilde{y}_{im}\gamma_{ij} = 0 \quad \forall \quad j$$

and a given t

$$1 - \sum_m \tilde{y}_{im} = 0 \quad \forall \quad i \geq z$$

$$1 - \sum_j \phi_{ij} = 0 \quad \forall \quad i$$

$$\phi_{ij} \geq 0, \ \tilde{y}_{im} \geq 0.$$

The model (6) subject to (7) was run one transition of the data at a time beginning with the most recent observation for which no data on 500+ cow herds were available. The notation $\tilde{y}_{im}$ represents the missing proportional data, where z is the largest state for which data are available. For example, when the model is run to determine how many 500+ head dairy farms there were in Pennsylvania in 1996, it estimates a transition probability matrix (the ϕ_{ij}'s) and, from the information available in the data, determines an entropy-maximizing way of finding what proportion of the 200–499 head data should be allocated to the 500+ head state. An identity matrix was used as the initial prior, with each estimated matrix used as the prior for each subsequent application of the model.

Shown in Table 1 are the original data as collected and reported by USDA, while presented in Table 2 are the reconstructed data after applying the entropy model in (6) and (7). Also shown in each table are the actual farm size states of the Markov chain to be modeled, namely, herd sizes of 1–29 head, 30–49 head, 50–99 head, 100–199 head, 200–499 head, and 500+ head. An additional state representing entry and exit must also be accommodated since proportional entry and exit from dairy is probably not realistic. Entry and exit are modeled as an extra state representing a pool from which entry can occur and exiting farms can go. Therefore, 6-herd–size states plus entry/exit equals a 7×7 transition probability matrix with 49 probabilities per time period. However, the row sum condition implies that only 42 of these need to be estimated at each time period.

Empirical Results

Given the magnitude of transition probabilities estimated, a complete presentation is not practical. Presented in Table 3 are the mean and standard deviation of the estimated probabilities over

Table 1. Reported Number of Pennsylvania Dairy Farms by Herd Size Category

	Number of Head						
Year	1–29	30–49	50–99	100+	200+	500+	Total
1979	11,330	6,380	3,410	880	–	–	22,000
1980	10,780	6,600	3,630	990	–	–	22,000
1981	9,345	6,825	3,885	945	–	–	21,000
1982	9,135	6,615	4,200	1,050	–	–	21,000
1983	9,135	6,300	4,515	1,050	–	–	21,000
1984	8,106	6,405	5,397	1,092	–	–	21,000
1985	8,211	6,300	5,292	1,197	–	–	21,000
1986	7,001	6,006	5,304	1,190	–	–	19,500
1987	6,290	5,902	5,106	1,203	–	–	18,500
1988	5,705	5,705	4,988	1,103	–	–	17,500
1989	5,198	5,297	4,802	1,205	–	–	16,500
1990	4,805	5,100	4,402	1,194	–	–	15,500
1991	4,292	4,800	4,205	1,204	–	–	14,500
1992	3,900	4,600	4,300	1,200	–	–	14,000
1993	3,400	4,300	4,200	930	170	–	13,000
1994	3,200	4,100	4,100	930	170	–	12,500
1995	2,800	3,900	3,900	1,000	200	–	11,800
1996	2,600	3,800	3,900	1,000	200	–	11,500
1997	2,400	3,800	3,800	1,100	190	10	11,300
1998	2,300	3,800	3,800	1,050	235	15	11,200
1999	2,300	3,700	3,800	930	255	15	11,000
2000	2,000	3,800	3,700	920	260	20	10,700
2001	1,800	3,700	3,600	910	260	30	10,300
2002	1,900	3,300	3,400	890	270	40	9,800
2003	2,000	3,000	3,300	980	280	40	9,600

Source: USDA (various issues).

the sample period. Employing a bootstrapping technique provides some guidance on the stability of the estimated transition probabilities. Transition probabilities in the shaded cells of the table are statistically significantly different from zero and, as shown, tend to be concentrated along the diagonal, the upper triangular portion, the right-most column, and the bottom row. These four areas correspond to farms remaining in their current size category, becoming larger, exiting, and entering.

As shown, most of the probability mass tends to be centered on the diagonal, indicating the high probability of farms remaining in the same size over time. The remaining probability mass tends to suggest changing farm size over time, along with a probability of exit that decreases as farm size increases. Entry tends to be a smaller farm phenomenon, while many entering the business of dairy farming in Pennsylvania exit the industry a year later (bottom row). These results are consistent with conventional wisdom and past research on a variety of farm types (e.g., Chavas and Magrand 1988, Disney, Duffy, and Hardy 1988, Garcia, Offutt, and Sonka 1987, Zepeda 1995a, and Gillespie and Fulton 2001). The fact that there can be some downward pressure on size adjustments that in many cases is stronger than the upward pressure is perhaps an indication of the potential for a bimodal distribution of farm sizes.

Seemingly Unrelated Regression Results

Degrees of freedom preclude the estimation of an equation for each probability in Table 3, so the equations in (5) were estimated for a subset of the

Table 2. Reconstructed Number of Pennsylvania Dairy Farms by Herd Size Category

	Number of Head						
Year	1–29	30–49	50–99	100–199	200–499	500+	Total
1979	11,330	6,380	3,410	813	67	0	22,000
1980	10,780	6,600	3,630	915	75	0	22,000
1981	9,345	6,825	3,885	864	81	0	21,000
1982	9,135	6,615	4,200	958	92	0	21,000
1983	9,135	6,300	4,515	946	104	0	21,000
1984	8,106	6,405	5,397	976	116	0	21,000
1985	8,211	6,300	5,292	1,069	128	0	21,000
1986	7,001	6,006	5,304	1,057	132	0	19,500
1987	6,290	5,902	5,106	1,064	138	0	18,500
1988	5,705	5,705	4,988	959	144	0	17,500
1989	5,198	5,297	4,802	1,057	147	0	16,500
1990	4,805	5,100	4,402	1,039	155	0	15,500
1991	4,292	4,800	4,205	1,044	160	0	14,500
1992	3,900	4,600	4,300	1,035	165	0	14,000
1993	3,400	4,300	4,200	930	170	0	13,000
1994	3,200	4,100	4,100	930	170	0	12,500
1995	2,800	3,900	3,900	1,000	199	1	11,800
1996	2,600	3,800	3,900	1,000	194	6	11,500
1997	2,400	3,800	3,800	1,100	190	10	11,300
1998	2,300	3,800	3,800	1,050	235	15	11,200
1999	2,300	3,700	3,800	930	255	15	11,000
2000	2,000	3,800	3,700	920	260	20	10,700
2001	1,800	3,700	3,600	910	260	30	10,300
2002	1,900	3,300	3,400	890	270	40	9,800
2003	2,000	3,000	3,300	980	280	40	9,600

Table 3. Mean and Standard Deviation of Transition Probability Estimates (1980–2002)[a,b]

t \ *t*+1	1 to 29	30 to 49	50 to 99	100 to 199	200 to 499	500+	Exit
1–29	**0.8051** (0.0372)	**0.0478** (0.0073)	**0.0314** (0.0091)	**0.0051** (0.0047)	**0.0014** (0.0014)	**0.0001** (0.0001)	**0.1092** (0.0261)
30–49	**0.0295** (0.0066)	**0.8312** (0.0159)	**0.0735** (0.0100)	**0.0020** (0.0033)	**0.0008** (0.0012)	**0.0002** (0.0003)	**0.0628** (0.0141)
50–99	**0.0000** (0.0000)	**0.0593** (0.0063)	**0.8696** (0.0158)	**0.0254** (0.0093)	**0.0006** (0.0011)	**0.0003** (0.0005)	**0.0449** (0.0117)
100–199	**0.0000** (0.0000)	**0.0000** (0.0000)	**0.0612** (0.0048)	**0.8975** (0.0097)	**0.0068** (0.0011)	**0.0000** (0.0000)	**0.0344** (0.0043)
200–499	**0.0000** (0.0000)	**0.0000** (0.0000)	**0.0000** (0.0000)	**0.0146** (0.0004)	**0.9853** (0.0005)	**0.0000** (0.0000)	**0.0000** (0.0000)
500+	**0.0000** (0.0000)	**0.0000** (0.0000)	**0.0000** (0.0000)	**0.0000** (0.0000)	**0.0000** (0.0000)	**0.9999** (0.0001)	**0.0000** (0.0000)
Entry	**0.2459** (0.0525)	**0.1712** (0.0434)	**0.0601** (0.0080)	**0.0003** (0.0004)	**0.0001** (0.0001)	**0.0000** (0.0000)	**0.5224** (0.0711)

[a] The top entry in each cell is the mean transition probability estimated as the average over the 1980 to 2002 period. The bottom entry is an estimate of the standard deviation.
[b] The shaded cells have mean values that are significantly different from zero.

probabilities. As Zepeda (1995a) points out, it is desirable to allow the explanatory variables to be different across equations since factors affecting entry and exit are likely different than those affecting size adjustments over time. For this reason, and because size adjustments (up or down) are of greater interest than non-changing size, equations for the probabilities along the diagonal were not specified. Of the 23 probabilities from Table 3 that are statistically significantly different from zero, ignoring the diagonal leaves 16 equations to be estimated.

The variables hypothesized as influencing all of the transition probabilities are the price of milk, the volatility of milk price, interest rates, and land values. Milk prices have been used as an explanatory variable by Rahelizatovo and Gillespie (1999) in their analysis of Louisiana dairies, while interest rates have been used as an explanatory variable by Chavas and Magrand (1988) and Zepeda (1995a) in their analyses of Wisconsin dairies.[2] High milk prices and low interest rates are hypothesized to attract entrants, foster expansion, and reduce exits from dairy.

Since the data span the time period during which the dairy termination program was in place, a dummy variable representing the program was included in the estimation of the equations describing the probability of exit from dairying. This variable is also consistent with research by Zepeda (1995a) and Rahelizatovo and Gillespie (1999).

Milk price volatility and land values have not been previously used in studies of agricultural structural change in dairy. Milk price volatility has increased in recent years and can have a pervasive influence on the decision to enter dairying as well as the decision to expand. Using real option theory, Tauer (2004) has hypothesized that the presence of milk price uncertainty implies that there is value in waiting to enter and exit dairy farming.

Zepeda (1995a) uses debt capital as a balance sheet explanatory variable. However, escalating land values in Pennsylvania are likely a more meaningful balance sheet variable for explaining entry and expansion deterrence, as well as one form of the exit decision, namely, retirement. This variable is also consistent with the structural model presented in the Appendix. Also consistent with the analytic model presented in the Appendix is the explanatory variable average milk production per cow. This variable is hypothesized to positively influence expansion of the dairy herd since, according to the Pennsylvania Agricultural Statistics Service (PASS), larger farms (i.e., more head) tend to produce, on average, more milk per cow. Zepeda (1995b) uses a variable such as this to measure the effect of technical change on farm structure, while Chavas and Magrand (1988) use it to investigate the effect of size economies on structural change. Rahelizatovo and Gillespie (1999) also use average milk/cow as a variable that describes expansion, contraction, and exit in Louisiana dairies.

Pennsylvania's annual average wholesale milk price per hundredweight reported by PASS is used as the price state variable, while the standard deviation of monthly wholesale milk prices represents annual milk price volatility. While not the price that producers actually receive, wholesale milk prices tend to track well with milk prices received by producers. Land values as reported by PASS, denoted by $v(t)$ in the structural model in the Appendix, are a state average and are in dollars per acre. Data on herd milk production, denoted by $q(t)$ in the structural model in the Appendix, was unavailable, so a proxy variable representing production per cow for each herd size i and collected by PASS was used in the equations describing farm size adjustments (up or down) only. Further, the variable used was the target level of production consistent with the size adjustment. For example, the transition from 1 to 29 head to 30 to 49 head is a function of the milk production per cow in the 30–49 head size category. Lastly, the prime rate reported by the Federal Reserve (2006) was used as a proxy for the cost of capital. While the prime rate is not the rate potentially acting as barrier to entry or expansion, it is likely correlated with the actual cost of capital facing dairy producers. Lastly, each variable is lagged one year to be consistent with the notion of entry, exit, and herd size change decisions being made with the best available information (Gillespie and Fulton 2001).

As shown in Table 4, there is generally a high degree of significance among the variables chosen to explain each of the transition probabilities. Consistent with Rahelizatovo and Gillespie (1999), milk prices negatively impact the probability of exit from dairy farming, while milk price volatility and land values both have a positive impact on the probability of exiting from dairy farming. The

[2] Chavas and Magrand (1988) and Zepeda (1995a) also use milk price as an explanatory variable but as a ratio of milk price to feed cost. In the present study, this variable had little predictive power in explaining transition probabilities.

Table 4. Seemingly Unrelated Regression Results for Selected Transition Probabilities

Dependent Variables		Explanatory Variables[a]									
$\Phi^{-1}(\pi_{ij})$[b]	Intercept	p	σ	r	dtp	v	q_1	q_2	q_3	q_4	q_5
From size i to exit											
$\Phi^{-1}(\pi_{17})$	-0.4163	-0.0993^c	0.1918^d	-0.0059	0.0698^c	0.0002^c	–	–	–	–	–
$\Phi^{-1}(\pi_{27})$	-0.6560^d	-0.0986^c	0.1311^d	0.0059	0.0355^c	0.0002^c	–	–	–	–	–
$\Phi^{-1}(\pi_{37})$	-0.5042	-0.1200^c	0.1899^c	0.0108	0.0288^d	0.0001^d	–	–	–	–	–
$\Phi^{-1}(\pi_{47})$	-1.3741^c	-0.0486^c	0.0873^c	0.0097^d	-0.0048	0.0001^d	–	–	–	–	–
From entry to size j											
$\Phi^{-1}(\pi_{71})$	-1.5992^c	0.0731	-0.2555^c	0.0041	–	0.0000	–	–	–	–	–
$\Phi^{-1}(\pi_{72})$	-1.8725^c	0.0614	0.0071	-0.0191	–	0.0001	–	–	–	–	–
$\Phi^{-1}(\pi_{73})$	-2.0176^c	0.0244^e	-0.0405	-0.0043	–	0.0001^c	–	–	–	–	–
From size i up to size j											
$\Phi^{-1}(\pi_{12})$	-1.8954^c	0.0024	0.0670	-0.0064	–	-0.0001^d	–	0.0270^c	–	–	–
$\Phi^{-1}(\pi_{13})$	-2.3366^c	0.0146	0.0062	0.0007	–	0.0001	–	–	0.0015	–	–
$\Phi^{-1}(\pi_{23})$	-2.0829^c	0.0373^d	-0.0868^d	0.0003	–	0.0000	–	–	0.0067^c	–	–
$\Phi^{-1}(\pi_{34})$	-3.5384^c	0.1041^d	-0.1906^e	-0.0359^e	–	-0.0001	–	–	–	0.0349^c	–
$\Phi^{-1}(\pi_{45})$	-2.5454^c	0.0027	0.0245	0.0037	–	-0.0000	–	–	–	–	0.0015
From size i down to size j											
$\Phi^{-1}(\pi_{21})$	-1.6570^c	-0.0149	-0.0873^e	0.0045	–	0.0002^c	-0.0287^c	–	–	–	–
$\Phi^{-1}(\pi_{32})$	-1.2443^c	-0.0230	-0.0739^c	0.0035	–	-0.0001^d	–	0.0046^c	–	–	–
$\Phi^{-1}(\pi_{43})$	-1.6498^c	-0.0135	0.0319	0.0062	–	-0.0000	–	–	0.0178^c	–	–
$\Phi^{-1}(\pi_{54})$	-2.2469^c	0.0023	-0.0041	-0.0010	–	0.0000	–	–	–	0.0022^c	–

[a] Explanatory variables are milk price (p), milk price volatility (σ), prime rate (r), dairy termination program (dtp), land value (v), and milk per cow for herd size i (q_i), all lagged. Dairy termination program is a qualitative variable.

[b] Subscripts on the π_{ij} reference states of the Markov chain where 1 refers to 1–29 head, 2 refers to 30–49 head, 3 refers to 50–99 head, 4 refers to 100–199 head, 5 refers to 200–499 head, 6 refers to 500 head or more, and 7 refers to the *entry* state when it is the first subscript and the *exit* state when it is the second subscript.

[c,d,e] Superscripts denote significance levels where the superscript c denotes significance at the 1 percent level in a two-tailed test that the estimated parameter in question is equal to zero. Superscript d (e) denotes significance at the 5 percent (10 percent) level.

dairy termination program also positively impacted exit from dairy in Pennsylvania, which is consistent with Rahelizatovo and Gillespie's (1999) finding for Louisiana dairies but not consistent with Zepeda's (1995a) for Wisconsin dairies. Interest rates appear to have no effect on the probability of exiting dairy farming, which is not consistent with Rahelizatovo and Gillespie's (1999) or Zepeda's (1995a) findings.

By contrast, few of the structural variables affect entry into dairying in any meaningful way. For relatively small startup dairies, milk price volatility has a negative impact on the probability of entry, while milk prices and land values tend to raise the probability of entry. This last result, while somewhat counterintuitive, is not completely unexpected as land prices have generally risen over the sample period, as have entrants into the smallest herd size state.

In terms of herd size adjustment, milk prices tend to play a role only on the probability of expanding the size of the dairy farm from a small- to medium-sized operation. Rahelizatovo and Gillespie (1999) found that higher milk prices increased the probability of small and large dairy farms remaining in their size categories. However, the probability of expanding or contracting the herd size is negatively impacted by milk price volatility for small- and medium-sized dairy farms. Also, small dairy farms likely face a significant barrier to expansion in land values, as evidenced by the negative coefficient estimated for the land price variable on the probability of expanding herd size from one to 29 head to 30 to 49 head. Small farms in Southeastern Pennsylvania, for example, are often faced with significant land values that tend to curtail expansion.

Lastly, the probability of adjusting herd size is typically positively impacted by the milk production per cow typical of the herd after expansion. This result is consistent with past research, such as Rahelizatovo and Gillespie (1999) as well as Chavas and Magrand (1988), which indicates that per cow productivity and economies of size play an important positive role in dairy farm size structural change. Zepeda (1995b) found that on Wisconsin dairy farms, technical change had no measurable impact on the probability of adjusting herd size.

Changing Share of Pennsylvania Dairy Farms

To get an idea of how the concentration of dairy farming has changed throughout the state of Pennsylvania over time, transition probabilities were also estimated for the proportion of dairy farms attributable to each crop-reporting district. In this setting the crop-reporting districts represent the states of the Markov chain and the transition probabilities measure the likelihood that the share of dairy farms in crop-reporting district i remain intact or are lost through exit. Transition probabilities were estimated using the stationary model presented in (3) subject to (4), with additional constraints to ensure that the transition probabilities capture the fact that farms do not generally move from one region to another. One exception is the migration of Amish farms from the Southeastern and South Central crop-reporting districts to districts further west in the state.[3] An identity matrix was used as a prior in conducting the estimation. An identity matrix prior and stationary estimation impose the condition that the proportion of dairy farms in each crop-reporting district is constant over time unless the data suggest otherwise.

Table 5 presents stationary transition probability estimates that suggest that the proportion of dairy farms does generally remain constant, with each crop-reporting district's probability of maintaining its current proportion of Pennsylvania dairy farms greater than 90 percent in all cases. Entry occurs for each region, with the East Central region experiencing the lowest number of new entrants and the Southeastern and Central regions experiencing the highest number of entrants. Exit also occurs for each region, and is highest for the Southwestern region while lowest for the Northwest region. Net entry (entry minus exit), while positive for four regions, is really significantly positive only for two regions, namely, the Southeastern (about 10 percent) and Central regions (about 3 percent). In fact, there is a small probability (about 1 percent) that Southeastern dairy farms relocate to the Central region. This last finding is consistent with recent events in

[3] Many Amish-run dairy farms exit out of dairy in Southeastern Pennsylvania and have relocated in states like Indiana, New York, and Ohio.

Table 5. Stationary Transition Probability Estimates for the Proportion of Dairy Farms in Each Crop-Reporting District

	Northwest	North Central	Northeastern	West Central	Central
Northwest	97.2%	–	–	–	–
North Central	–	90.3%	–	–	–
Northeastern	–	–	90.3%	–	–
West Central	–	–	–	94.5%	–
Central	–	–	–	–	91.8%
East Central	–	–	–	–	–
Southwestern	–	–	–	–	–
South Central	–	–	–	–	–
Southeastern	–	–	–	–	0.7%
Entry	3.0%	9.1%	4.4%	3.2%	11.3%

	East Central	Southwestern	South Central	Southeastern	Exit
Northwest	–	–	–	–	2.8%
North Central	–	–	–	–	9.7%
Northeastern	–	–	–	–	9.7%
West Central	–	–	–	–	5.5%
Central	–	–	–	–	8.2%
East Central	92.1%	–	–	–	7.9%
Southwestern	–	90.0%	–	–	10.0%
South Central	–	–	90.7%	–	9.3%
Southeastern	–	–	–	94.7%	4.6%
Entry	1.1%	5.1%	10.3%	14.9%	37.6%

Pennsylvania where dairy farms have been liquidated in high land value areas such as Lancaster County and subsequently relocated to Central Pennsylvania.

These results suggest that, in general, most regions in Pennsylvania have more exits than entrants. However, at least two regions have an increasing share of the total number of dairy farms, indicating that the decline in the number of dairy farms within the entire state is likely not uniform, with some areas harder hit than others. The results also suggest that dairy farms in the Southeastern region, the traditional Pennsylvania dairying region, have the most net entry and have had some farmers relocate their dairies to Central Pennsylvania, where land values are considerably lower.

Stochastic Simulation Results

To get an idea of what the future might hold for Pennsylvania dairy farming in terms of the number of dairy farms in each size category, a stochastic simulation model was developed. Matrices of transition probabilities were simulated by first simulating correlated state variable values out over a horizon of 20 years and then simulating correlated errors for each of the SUR equations over a 20-year period. Empirical distributions of the number of farms in each size category were then constructed and are presented in Figure 1. Not surprisingly, the distribution of the total number of dairy farms shifts leftward and widens as the forecast period lengthens. This implies fewer dairy farms in the future and more uncertainty about the exact number.

Presented in Table 6 are summary statistics for the simulated distributions appearing in Figure 1. As shown, the number of dairy farms is expected to fall from its 2003 level of about 9,600 to, on average, about 8,600 over the next 5 years, for a total reduction of just over 10 percent. All of the reduction is expected to come from farms with

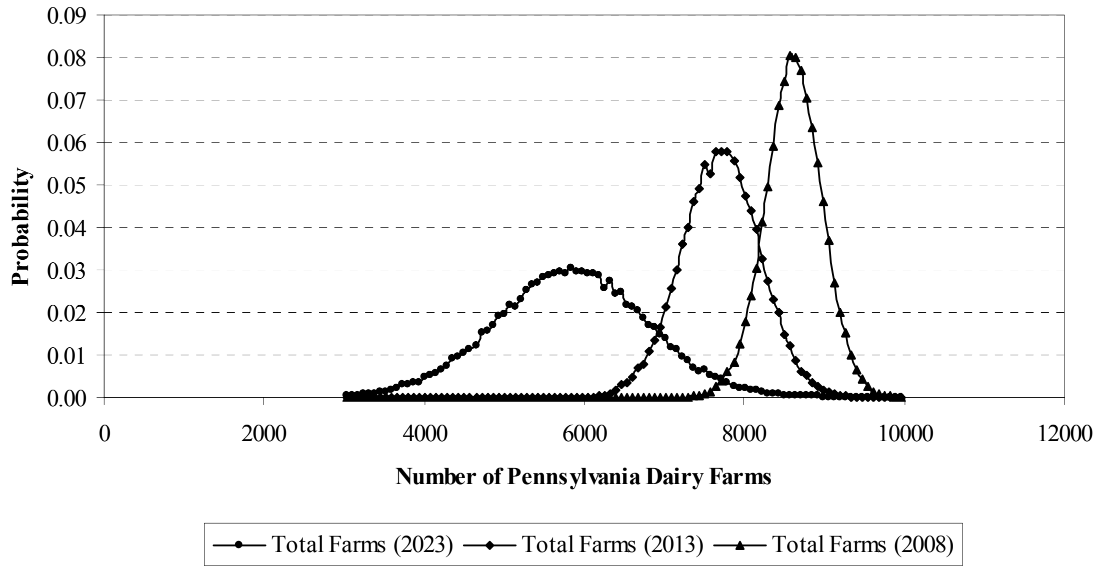

Figure 1. Simulated Distributions of the Total Number of Pennsylvania Dairy Farms for Various Years

less than 100 head, as farms with more than 100 head are expected to increase in number from the 2003 levels.

The results are similar for more distant years as the mean number of farms is expected to fall to about 7,700 by 2013 (just under a 20 percent reduction) and 5,800 by 2023 (just under a 40 percent reduction). These numbers are suggestive of an average annual loss of about 2.0 percent to 2.5 percent. The largest losses are the small dairy farms, while larger farms will continue to grow in number. Interestingly, the 100–199 cow herds appear to be a size category that is right at the break point between expansion and contraction. The average number of 100–199 head dairy farms is expected to increase from 980 in 2003 to 1,138 in 2008, to a high of 1,176 in 2013, and then fall off to 1,098 by 2023. This result is potentially consistent with the idea of a bimodal distribution of dairy farms in the future. Such a finding would be inconsistent with empirical evidence for U.S. dairies (Wolf and Sumner 2001), but may be a possibility in Pennsylvania. Even so, the most important feature of the simulations is that if the status quo is maintained, Pennsylvania will likely have a significant number of dairy farms for at least the next 20 years. In all likelihood, there may always be a significant number of dairy farms in Pennsylvania to serve the state's fresh milk markets.

Summary and Conclusions

The primary purpose of this research has been to identify factors affecting structural change in Pennsylvania's dairy sector, including growth and contraction as well as entry and exit from dairying. Milk prices, price volatility, land values, and the dairy termination program have all had pervasive effects on exit from dairying in Pennsylvania. Similarly, milk price volatility is generally a deterrent to entry for small dairies, the size that new dairies entering the sector are most likely to be. Dairy farm size growth is inhibited by milk price volatility and land values, but responds positively to higher milk prices. Lastly, growth and contraction are positively related to productivity through milk production per cow.

From a policy perspective, it is clear that if the state of Pennsylvania desires to be known as a dairy state in the future, strong, stable milk prices are an important factor. While strong and stable milk prices have traditionally attracted new entrants to dairy, they have also provided signals to producers to expand their operations. Over time,

Table 6. Selected Statistics for Future Distribution of Farm Numbers by Size

Year	Farm Size (no. of head)	Mean	Standard Deviation	Min.	Max.	%Δ from 2003	p^a
2008	1–29	1,307	167	709	2,123	-34.7%	95.43%
	30–49	2,623	146	1,931	3,265	-12.6%	5.84%
	50–99	3,174	162	2,527	3,829	-3.8%	0.51%
	100–199	1,138	104	825	1,661	16.1%	33.41%
	200–499	324	3	312	339	15.7%	0.86%
	500+	48	0	48	49	20.0%	90.83%
	Total	**8,615**	**355**	**7,152**	**10,235**	**-10.3%**	**0.56%**
2013	1–29	1,059	169	479	1,813	-47.1%	99.14%
	30–49	2,148	184	1,372	2,898	-28.4%	91.42%
	50–99	2,905	219	2,049	3,971	-12.0%	10.93%
	100–199	1,176	148	707	2,052	20.0%	46.82%
	200–499	362	6	338	394	29.3%	100.0%
	500+	56	0	54	58	40.0%	100.0%
	Total	**7,705**	**486**	**5,720**	**9,889**	**-19.7%**	**46.86%**
2023	1–29	811	321	0	3,324	-59.5%	98.23%
	30–49	1,359	292	265	2,475	-54.7%	98.82%
	50–99	2,091	356	184	3,497	-36.6%	94.93%
	100–199	1,098	253	458	2,776	12.0%	32.94%
	200–499	417	14	370	484	48.9%	100.0%
	500+	67	1	62	73	67.5%	100.0%
	Total	**5,843**	**945**	**1,914**	**10,674**	**-39.1%**	**97.31%**

[a] For 1–29 head, 30–49 head, 50–99 head, and total farms, the value reported is the probability of observing a number of farms less than 20 percent of 2003 level. For 100–199 head, 200–499 head, and 500+ head, the value reported is the probability of observing a number of farms in excess of a 20 percent increase from the 2003 level.

it will become increasingly difficult for new entrants to secure enough capital to begin dairying, and state programs like the Next Generation Farmer Loan Program could provide a necessary tool for dealing with the issue. In recent years, milk prices have been more volatile, which has likely contributed to some of the exodus from dairy farming within the state. Efforts to stabilize prices could have a reversing effect.

Perhaps the most important factor contributing to the decline in the number of dairy farms is the value of agricultural land in Pennsylvania, especially in areas where there is significant dairy production. High land values act as a barrier to entry and expansion, but also provide exit incentives for older farmers seeking to retire. The problems created by high land values for agricultural producers will undoubtedly become more pronounced in the future. As a result, Pennsylvania's farmland preservation programs will likely become an increasingly important policy mechanism for dealing with the problem.

A secondary contribution of the research has been to suggest what the future might hold for the number of dairy farms in Pennsylvania. It is clear that if the status quo is maintained, there will be fewer, larger dairy farms, with the rate of decline estimated to be about 2.0 percent to 2.5 percent annually over the next 20 years. Exiting dairy farms will be primarily smaller operations. However, negative growth rates such as these hardly suggest that there is no future in dairy in Pennsylvania. It is likely that in 20 years, Pennsylvania will still have a significant number of dairy farms, although it remains unclear whether or not the state will still be producing a significant portion of the nation's milk supply.

Lastly, a significant innovation of the research is the presentation of the linkage between an analytic model of the firm and the Markov chain model used so often to conduct structural change studies. As long as the decision maker makes decisions consistent with a dynamic planning horizon and faces uncertainty that is Markovian, the size of the firm will also be Markovian, having inherited the property from the underlying sources of uncertainty through the optimization process. This puts the Markov chain methodology on a more solid foundation for analyzing firm size dynamics.

References

Arnold, L. 1974. *Stochastic Differential Equations: Theory and Applications*. New York: John Wiley and Sons.

Chan, M. 1981. "A Markovian Approach to the Study of the Canadian Cattle Industry." *The Review of Economics and Statistics* 63(1): 107–116.

Chavas, J., and G. Magrand. 1988. "A Dynamic Analysis of the Size Distribution of Firms: The Case of the U.S. Dairy Industry." *Agribusiness* 4(4): 315–329.

Disney, W., P. Duffy, and W. Hardy. 1988. "A Markov Chain Analysis of Pork Farm Size Distributions in the South." *Southern Journal of Agricultural Economics* 20(2): 57–64.

Ethridge, D., S. Roy, and D. Myers. 1985. "A Markov Chain Analysis of Structural Changes in the Texas High Plains Cotton Ginning Industry." *Southern Journal of Agricultural Economics* 17(2): 11–20.

Federal Reserve. 2006. *Statistical Release H.15 Selected Interest Rates* (1980–2003). Washington, D.C.: Federal Reserve. Available at www.federalreserve.gov (accessed June 6, 2006).

Garcia, P., S. Offutt, and S. Sonka. 1987. "Size Distribution and Growth in a Sample of Illinois Cash Grain Farms." *American Journal of Agricultural Economics* 69(2): 471–476.

Gillespie, J., and J. Fulton. 2001. "A Markov Chain Analysis of the Size of Hog Production Firms in the United States." *Agribusiness* 17(4): 557–570.

Golan, A., G. Judge, and D. Miller. 1996. *Maximum Entropy Econometrics: Robust Estimation with Limited Data*. New York: John Wiley and Sons.

Hallberg. M. 1969. "Projecting the Size Distribution of Agricultural Firms: Application of a Markov Process with Non-stationary Transition Probabilities." *American Journal of Agricultural Economics* 51(2): 289–302.

Kim, C., W. Lin, and M. Leath. 1991. "The Changing Structure of the U.S. Flour Milling Industry." *The Journal of Agricultural Economics Research* 43(3): 18–25.

Lee, T., and G. Judge. 1996. "Entropy and Cross-Entropy Procedures for Estimating Transition Probabilities from Aggregate Data." In D. Berry, K. Chaloner, and J. Geweke, eds., *Bayesian Analysis in Statistics and Econometrics*. New York: John Wiley and Sons.

Lee, T., G. Judge, and T. Takayama. 1965. "On Estimating the Transition Probabilities of a Markov Process." *Journal of Farm Economics* 47(3): 742–762.

Lee. T., G. Judge, and A. Zellner. 1973. *Estimating Transition Probabilities of the Markov Chain Model from Aggregate Share Data.* Amsterdam*:* North-Holland.

Malliaris, A. and W. Brock. 1991. *Stochastic Methods in Economics and Finance*. New York: North-Holland.

Padberg, D. 1962. "The Use of Markov Processes in Measuring Changes in Market Structure." *Journal of Farm Economics* 44(1): 189–199.

Rahelizatovo, N., and J. Gillespie. 1999. "Dairy Farm Size, Entry, and Exit in a Declining Production Region." *Journal of Agricultural and Applied Economics* 31(2): 333–347.

Tauer, L. 2004. "When to Get In and Out of Dairy Farming: A Real Option Analysis." Working Paper No. 2004-12, Department of Applied Economics and Management, Cornell University, Ithaca, NY.

U.S. Department of Agriculture. Various issues. "Milk Production." USDA, Washington, D.C. (February issues, 1980–2004).

Wolf, C., and D. Sumner. 2001. "Are Farm Size Distributions Bimodal? Evidence from Kernel Density Estimates of Dairy Farm Size Distributions." *American Journal of Agricultural Economics* 83(1): 77–88.

Zellner, A. 1962. "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests of Aggregation Bias." *Journal of the American Statistical Association* 57(298): 348–368.

Zepeda, L. 1995a. "Asymmetry and Nonstationarity in the Farm Size Distribution of Wisconsin Milk Producers: An Aggregate Analysis." *American Journal of Agricultural Economics* 77(4): 837–852.

____. 1995b. "Technical Change and the Structure of Production: A Non-Stationary Markov Analysis." *European Review of Agricultural Economics* 22(1): 41–60.

APPENDIX

Let $q(t)$ be a state variable representing time t milk production (say, in cwt) for a dairy herd with n cows. More clearly, milk production is a state variable with $q(t) = q[\mathbf{x}(\mathbf{t}),n(t)]$, where $\mathbf{x}(\mathbf{t})$ is a vector input [the kth of which is $x_k(t)$], and $n(t)$ is the size of the herd (i.e., the number of cows to be milked). In this setting, $\mathbf{x}(\mathbf{t})$ and $n(t)$ are decision variables that the farmer has control over. It is assumed that $q(t)$ is a well-behaved production function.[4]

The objective of the producer is to maximize the discounted flow of profit from dairying plus

[4] Namely, $\partial q(t)/\partial x_k(t) \geq 0$ for all $k = 1,\ldots,K$ and $\partial q(t)/\partial n(t) \geq 0$.

the terminal value of the farm's land, $v(T)$, over a finite planning horizon.[5] Let ρ be the producer's discount rate and let $J(q,v,t)$ represent the time t value of the dairy farm conditional on the optimal use of inputs. The objective of the producer is then

(A1)
$$J(q,v,t) = \max_{\mathbf{x},n} \quad E_t \int_t^T e^{-\rho s} \left[pq(\mathbf{x},n) - \mathbf{wx} - w_n n \right] ds + e^{-\rho T} v(T),$$

where time dependence has been intentionally suppressed where no confusion can arise. The first term in (A1) represents the discounted flow of profit from dairying (revenue less input and cow costs), while the second term represents the present value of the farm's terminal stock of land. The E_t in (A1) is a time t expectation operator.

In this setting, each state variable must have an equation of motion that describes how the variables evolve over time. The following set of first-order, stochastic differential equations (SDEs) describes this evolution:

(A2)
$$dq = \sigma_q(q,n,t) \quad dz_q$$
$$dv = \mu(v,t) \quad dt + \sigma_v(v,t) \quad dz_v.$$

As shown, milk production is a driftless stochastic process with an instantaneous expected value equal to zero and instantaneous variance equal to

$$\sigma_q^2(q,n,t) \quad dt.$$

Notice that the volatility of the change in milk production depends on how much is being produced as well as the decision regarding the number of cows to milk. In this setting, there is likely an optimal number of head of cows to milk given the farmer's managerial ability.[6]

[5] A finite planning horizon and the inclusion of a terminal value are unnecessary to show that farm size is Markovian. However, as shown below, such a specification does demonstrate how such state variables come to influence transition probabilities.

[6] For example, $\partial\sigma_q/\partial n$ would be negative (positive) if managerial ability is under- (over-) capitalized.

Land values are assumed to increase over time with an instantaneous expected value equal to $\mu(v,t)\ dt$ and variance equal to

$$\sigma_v^2(v,t) \quad dt,$$

although their moments are unrelated to the control variables. Like the production function in (A1), the functional forms of the drift and diffusion terms in (A2) are left unspecified to promote generality. However, an important feature of the equations in (A2) for what follows is the assumption that each SDE possesses the Markov property. This is a minor assumption in that all first-order SDEs with an initial condition and solution possess the Markov property (Arnold 1974).

Without functional forms for the production function in (A1) and the drift and diffusion terms in (A2), an explicit solution to the system is not possible. However, to show that maximizing a system like (A1) subject to (A2) results in optimal policies that support the Markov chain model of farm size is relatively straightforward, although tedious. Of specific interest is the herd size control variable. If herd size can be shown to be Markovian, then the Markov chain model of farm size is an appropriate methodology for examining structural change and its effect on farm size.

To begin, the Hamilton-Jacobi-Bellman (HJB) equation of stochastic control for the system (A1) subject to (A2) is

(A3)
$$0 = \max_{\mathbf{x},n} \quad E_t \left\{ \left[pq(\mathbf{x},n) - \mathbf{wx} - w_n n \right] e^{-\rho t} dt + \xi(J) \right\},$$

where $\xi(J)$ is a differential operator representing Ito's lemma applied to J. $\xi(J)$ is given by

(A4)
$$\xi(J) = \frac{\partial J}{\partial t} dt + \frac{\partial J}{\partial q} dq + \frac{\partial J}{\partial v} dv + \frac{1}{2} \left(\frac{\partial^2 J}{\partial q^2} dq^2 + 2 \frac{\partial^2 J}{\partial q \partial v} dq dv + \frac{\partial^2 J}{\partial v^2} dv^2 \right).$$

Substituting the SDEs in (A2) into (A4) where appropriate, substituting the result into (A3), passing the expectation operator through the resulting expression, and canceling dts gives the following HJB equation:

(A5)

$$-\frac{\partial J}{\partial t} = \max_{\mathbf{x},n} \left\{ \begin{array}{l} e^{-\rho t}\left[pq(\mathbf{x},n) - \mathbf{wx} - w_N n\right] + \left(\frac{\partial J}{\partial v}\right)\mu(v,t) \\ +\frac{1}{2}\left[\begin{array}{l} \left(\frac{\partial^2 J}{\partial q^2}\right)\sigma_q^2(q,n,t) \\ +2\left(\frac{\partial^2 J}{\partial q \partial v}\right)\sigma_q(q,n,t)\ \sigma_v(v,t)\ r_{qv} \\ +\left(\frac{\partial^2 J}{\partial v^2}\right)\sigma_v^2(v,t) \end{array}\right] \end{array} \right\}$$

In (A5), r_{qv} is the correlation coefficient between the milk production and land value state variables.

Differentiating the right-hand side of (A5) with respect to the control variables gives the $K + 1$ first-order conditions:

(A6)

$$\frac{\partial(\cdot)}{\partial x_k} \equiv e^{-\rho t}\left[p\left(\frac{\partial q}{\partial x_k}\right) - w_k\right] = 0 \quad \forall \quad k = 1,...,K$$

$$\frac{\partial(\cdot)}{\partial n} \equiv e^{-\rho t}\left[p\left(\frac{\partial q}{\partial n}\right) - w_n\right] + \frac{1}{2}\left(\frac{\partial^2 J}{\partial q^2}\right)\left(\frac{\partial \sigma_q^2(q,n,t)}{\partial n}\right)$$

$$+2\left(\frac{\partial^2 J}{\partial q \partial v}\right)\left(\frac{\partial \sigma_q(q,n,t)}{\partial n}\right)\ \sigma_v(v,t)\ \ r_{qv} = 0.$$

The solution to the first-order conditions in (A6) can be symbolically expressed as

(A7)

$$x_k = x_k\left(x_1,...x_{k-1}, x_{k+1},..., x_K, n; p, w_k, \boldsymbol{\alpha}\right),$$

$$\forall \quad k = 1,...K \text{ and } n = n\left(\mathbf{x}, \frac{\partial^2 J}{\partial q^2}, t; \rho, p, w_n, \boldsymbol{\beta}\right),$$

where α represents a vector of parameters arising from the technical coefficients of the production function and β represents a vector of parameters arising from the production function and volatility of milk production.[7] A set of simultaneous solutions to (A7) can be symbolically written as

(A8) $$x_k^* = x_k^*\left(\frac{\partial^2 J}{\partial q^2}, t; \boldsymbol{\theta}\right), \ \forall \ k = 1,...K$$

$$\text{and } n^* = n^*\left(\frac{\partial^2 J}{\partial q^2}, t; \boldsymbol{\theta}\right),$$

which shows that the optimal policies depend on a second partial derivative of the value functional, time, and a vector of parameters consisting of the producer's discount rate, output price, input prices, technical coefficients from the production function, and the volatility of milk production.

Substituting (A8) into (A5) implies that the HJB equation (A5) may be rewritten as

(A9)

$$-\frac{\partial J}{\partial t} = e^{-\rho t}\left[pq\left(\mathbf{x}^*, n^*\right) - \mathbf{wx}^* - w_n n^*\right]$$

$$+\left(\frac{\partial J}{\partial v}\right)\mu(v,t)$$

$$+\frac{1}{2}\left[\left(\frac{\partial^2 J}{\partial q^2}\right)\sigma_q^2\left(q, n^*, t\right) + \left(\frac{\partial^2 J}{\partial v^2}\right)\sigma_v^2(v,t)\right].$$

which is a second-order, nonlinear,[8] partial differential equation (PDE). The boundary condition for (A9) is $J[q(T), v(T), T] = e^{-\rho T}v(T)$, and the solution can be denoted symbolically as $J^*(q,v,t)$.

Upon arriving at a solution to the PDE (A9) subject to the boundary condition, partial derivatives can be computed and substituted where necessary in the solutions to the first-order conditions given in (A8). More clearly,

(A10)

$$x_k^* = x_k^*\left(\frac{\partial^2 J^*}{\partial q^2}, t; \boldsymbol{\theta}\right) = x_k^*(q,v,t;\boldsymbol{\theta}), \ \forall \ \ k = 1,...,K$$

$$\text{and } n^* = n^*\left(\frac{\partial^2 J^*}{\partial q^2}, t; \boldsymbol{\theta}\right) = n^*(q,v,t;\boldsymbol{\theta}).$$

---

[7] From this point forward, it is assumed that $r_{qv} = 0$. Milk production and land values are likely not correlated, and the assumption simplifies the math to some degree.

[8] The equation is nonlinear because σ_q^2 is a function of n^* which itself is a function of $\partial^2 J/\partial q^2$. This term is multiplied by $\partial^2 J/\partial q^2$ in the last term of (9), implying that $(\partial^2 J/\partial q^2)^2$ appears in (9).

This result follows from the fact that the optimal solutions depend upon the partial derivatives of the value function, time, and a vector of parameters. Since the optimal $J^*(\cdot)$ can depend on only the state variables and time, partial derivatives of $J^*(\cdot)$ can at most depend on only the state variables and time. Therefore, the optimal policies depend at most on the two state variables, time, and the vector of parameters.

The most important implication of the result presented in (A10) is the following. The dynamics of herd size can be derived by applying Ito's lemma to n^* in (A10), giving

(A11)
$$dn^* = \left(\frac{\partial n^*}{\partial t}\right) dt + \left(\frac{\partial n^*}{\partial q}\right) dq + \left(\frac{\partial n^*}{\partial v}\right) dv + \frac{1}{2}\left[\left(\frac{\partial^2 n^*}{\partial q^2}\right) dq^2 + 2\left(\frac{\partial^2 n^*}{\partial q \partial v}\right) dq dv + \left(\frac{\partial^2 n^*}{\partial v^2}\right) dv^2\right].$$

Substituting n^* from (A10) into the SDE describing milk production evolution in (A2), and substituting the result into (A11), implies that the dynamics of herd (farm) size are characterized by the first-order SDE,

(A12)
$$dn^* = \mu^*(q,v,t;\boldsymbol{\beta})\ dt + \left(\frac{\partial n^*}{\partial q}\right) \sigma_q(q,t)\ dz_q + \left(\frac{\partial n^*}{\partial v}\right) \sigma_v(v,t)\ dz_v,$$

where

(A13)
$$\mu^*(q,v,t;\boldsymbol{\beta}) = \frac{\partial n^*}{\partial t} + \frac{\partial n^*}{\partial v}\mu(v,t) + \frac{1}{2}\left[\left(\frac{\partial^2 n^*}{\partial q^2}\right)\sigma_q^2(q,v,t;\boldsymbol{\beta}) + \left(\frac{\partial^2 n^*}{\partial v^2}\right)\sigma_v^2(v,t)\right].$$

A similar representation can be accomplished for each of the optimal inputs. However, our primary interest is with regard to n^*, so no attempt is made to derive the SDEs that characterize optimal input use dynamics. Arnold (1974) has shown that when SDEs like those in (A2) possess the Markov property, an equation like (A12) will also possess the property. This is because n^* depends on state variables that are Markovian, and therefore the dynamics of farm size (as measured by number of head) will inherit the Markov property from those state variables.

What this all means is that if a producer can be characterized as behaving in a manner consistent with the stochastic optimal control problem presented in (A1) subject to (A2), namely, chooses inputs and the number of cows to milk consistent with the objective of maximizing the dynamic value of the farm (and optionally the terminal value of farm's land) and faces uncertainty that is Markovian, then the size of the farm, as measured by the number of head of cows to milk, is Markovian as well.

In this setting, state transition probabilities are necessarily dependent on the state variables describing the underlying stochastic optimal control problem faced by the decision maker.[9] Let $\pi_{ij}(t)$ denote nonstationary transition probabilities of a Markov chain that quantify the time t probability that a dairy farm with a herd size of i transitions to a herd size of j over one period. In the context of the previous control problem, it must be the case that $\pi_{ij}(t) = \pi_{ij}[q(t), v(t)]$, which shows that the transition probabilities depend on the two state variables describing milk production and land values. More generally, $\pi_{ij}(t) = \pi_{ij}[\mathbf{s}(\mathbf{t})]$, where $\mathbf{s}(\mathbf{t})$ is a vector of state variables that describe the firm's dynamic decision problem. How each of the state variables affects the transition probabilities is critical to understanding how structural change influences the composition of dairy farms, both in terms of the number and size of farms.

[9] See, for example, Malliaris and Brock (1991) for a more rigorous discussion on this point.