Milked for All They Are Worth: Analyzing Livestock Mortality Costs in a Dynamic Discrete Choice Model

Jared Hutchins
Department of Agricultural and Applied Economics,
University of Wisconsin–Madison,
jhutchins@wisc.edu

Abstract

This paper examines animal replacement behavior for over one-thousand Wisconsin dairy farms during the period 2011-2014 and analyzes the rationale for high replacement rates. Dairy farmers in the United States routinely replace their cattle at around 3 years old, well before what the dairy science literature estimates is their maximum productive potential, that is 5 years in the herd. I model the replacement decision using a dynamic discrete choice model and incorporate unplanned mortality as a source of uncertainty that drives farmers to replace dairy cows before they maximize production. The empirical model incorporates cow and herd heterogeneity in mortality rates to back out the implied cost of cow mortality and the parameters of the cow's production function. Using the conditional choice probability method, I estimate the cost of mortality at 1,300-1,400 USD per death, about the average annual profit of a dairy cow. The results of the model also suggest that dairy cows maximize production at three years instead of five. Utilizing the heterogeneity in farm size, I also find that mortality costs are four times higher on small dairies than on larger ones. These results suggest that the costs incurred from increasing genetic milk production are disproportionately borne by small farms.

This is a preliminary draft. Please do not cite or distribute without permission of the author/s.
Contact author for latest draft or click here.
Draft date: September 1, 2019

1 Introduction

Dairy farms in the United States routinely cull animals before the dairy science literature says they maximize their productive potential. Measured in year-long production cycles called "lactations," a dairy cow's annual milk output is maximized at about the fifth lactation (Mellado et al., 2011; Ray et al., 1992). In practice, those animals are instead replaced at three lactations (Hadley et al., 2006; Knaus, 2009). Typical asset replacement models assuming profit maximization suggest much lower rates of culling then practiced; despite this, dairy farmers have maintained consistently high rates of replacement (De Vries, 2013; Van Arendonk, 1988). On its face, it appears that dairy farmers are leaving money on the table by culling too early. What economic rationale could there be, if any, for such a pattern of asset replacement?

I investigate whether the costs of unplanned mortality explain high replacement rates on more than one-thousand Wisconsin dairy farms with over three-hundred thousand cows using a dynamic, discrete choice model. I develop a theoretical model where the risk of unplanned mortality, an event where an animal dies or is removed from the herd unexpectedly, drives farmers to replace their cows before animals maximize production. Estimating the dynamic model on data, I calculate the cost of unplanned mortality as 1,300-1,400 USD, and the age at which cows maximize their annual production to be three lactations instead of five; this age is in line with the empirical production function calculated from the same data. Utilizing the heterogeneity in the dataset, I also find that mortality costs are almost four times higher on small dairies than on larger dairies. Finally, using the structural model, I calculated a compensating variation measure for the transition to a world without mortality risk. On average, farmers were willing to pay about 2,100 USD to eliminate unplanned mortality completely, meaning the effect of unplanned mortality on the other factors of replacement is about 700 USD.

Understanding mortality cost has become important since declining animal health is a significant factor in dairy farm decision making. While milk production of dairy cattle has increased at about 3-4% annually, so has the incidence of infertility and disease (Pryce and Veerkamp, 2001). In the UK, rates of pregnancy on the first breeding attempt have fallen from 55.6% in 1975 to 39.7% in 1998, and a global survey of Holstein cattle indicated similarly decreasing fertility worldwide (Pryce et al., 2014). Increases in yield have also been linked to a decline in cow health, resulting in a higher incidence of reproductive and metabolic diseases (Dechow et al., 2004; McConnel et al., 2008). The consequence of these trends is that dairy cows live shorter and shorter lives in the herd and are more susceptible to "unplanned mortality." Compared to dairy cows born in 1960, a dairy cow's life span in the herd has decreased by 20% (De Vries, 2013).

Further, unplanned mortality is potentially very costly to dairies. Unplanned mortality on dairy farms incurs significant costs because of lost production, the cost of treating illness, and lost salvage value when the animal is disposed of instead of sold. If mortality costs are high, dairy managers may replace early not only to acquire higher producing cows but also to salvage their older cows to avoid this cost. Early replacement then becomes a precautionary measure to protect against the high costs incurred from letting the asset "fail" in its production cycle instead of being salvaged on schedule (Burt, 1965). Such costs are a concern in times when milk profit margins are low, invalidating the economic gain from milk production. Even apart from the on-farm costs, decreasing animal lifespan has negative externalities affecting environmental quality and animal welfare, two issues of increasing importance to the general public (Garnsworthy, 2004; Oltenacu and Broom, 2010).

This paper calculates just how costly unplanned mortality is on Wisconsin dairies using empirical replacement decisions. Since the costs of mortality are hard to estimate empirically, Heikkilä et al. (2012) and De Vries (2013) use dynamic programming simulations and calculate these costs to be in the range of 500-1,000 USD per death. However, simulations are inadequate for understanding the costs that dairy farmers face. Rather than relying on purely counterfactual simulations, reliable cost estimates should be derived from an empirical model estimated on actual replacement decisions. Providing an empirical estimation of mortality cost to guide policy is the main contribution of this paper.

I also make a contribution to the literature on asset replacement and livestock mortality cost by developing an empirical model incorporating the risk of asset failure that can calculate the costs directly from data on replacement decisions. I estimate the model on dairy cow replacement decisions using the conditional choice probability (CCP) estimator pioneered by Hotz and Miller (1993) and Arcidiacono and Miller (2011), while controlling for unobserved, permanent asset heterogeneity. The CCP estimator estimates the parameters of the structural model by approximating the continuation value of the dynamic programming problem using the empirical probabilities of replacement from the data. Recent advances in CCP methods, specifically the Euler equations in conditional choice probabilities (ECCP) method of Scott (2013) and Aguirregabiria and Magesan (2013), allow estimation of dynamic discrete choice models that can incorporate fixed effects by using discrete analogs of Euler conditions. Using my structural model, the Euler condition estimates the cost of unplanned mortality, annual maintenance costs, and the parameters of the production function. Mine is also the first paper to estimate the parameters of a dairy cow production function from a behavioral model instead of from milk production data, allowing me to test whether farmers perceive a production function different than what the literature estimates.

I estimate the structural parameters of the theoretical model using an animal level dataset of over 300,000 dairy cows on over 1,000 Wisconsin dairy farms. I find that the costs of unplanned mortality explain early replacement and are variable across herds. My estimate of 1,300 - 1,400 USD aligns with the upper bound of De Vries (2013) and is the first empirical verification of these costs. The model estimates that small dairies perceive mortality costs to be as high as 2,000 USD per death, which is nearly four times higher than the cost of unplanned mortality on larger dairies. In addition to declining health, my results also show that early replacement is explained by how the manager perceives the production technology. From the manager's perspective, dairy cow milk production is maximized at age three instead of at age five, which is consistent with the empirical production function.

The policy implications of this study are that unplanned mortality costs are a key determinant of replacement patterns and disproportionately affect small dairies. Since smaller dairies experience the highest mortality costs, the costs incurred from increasing genetic milk production are disproportionately borne by small farms. In times of low milk price, as in the past few years, small farms will exit at higher rates than large dairies because of these breeding strategies.

2 Literature Review

This paper joins a very long tradition of analyzing asset replacement problems, specifically the dairy cow replacement problem. Asset replacement, a special class of the "optimal stopping problem," has been analyzed as early as 1849 when German forester Martin Faustmann developed the "Faustmann criterion" for determining the optimal harvest age of a forest (Newman, 2002). My paper makes contributions to the literature on dairy cow replacement by empirically estimating costs from data, as well as exploring the unknown causes of replacement elucidated in another empirical dairy cow replacement model, Miranda and Schnitkey (1995). This paper also uses advancements in dynamic discrete choice modeling to control for a broader range of unobserved heterogeneity than usually feasible in asset replacement models.

While the literature on optimal dairy cow replacement rules is expansive, the majority of studies use simulations to calculate costs rather than empirical models. These models represent the "normative" approach to asset replacement, which is to estimate parameters from data only to plug these in as primitives to a dynamic programming simulation. The dynamic program is then solved to recover the optimal culling rule. Attempting to estimate the optimal replacement policy for dairy cattle, dates back to Stewart et al. (1977), a paper in the Journal of Dairy Science explicitly modeling and solving the decision using dynamic programming. The state variables included the age of the cow, its body weight, its milk production, and its butterfat production. Subsequent models were more complex and gave the most attention to modeling the underlying biological processes of the dairy cow production system such as milk production (Rogers et al., 1988b; Stewart et al., 1977), fertility (Kalantari et al., 2010; Rogers et al., 1988a) and the incidence of disease (Bar et al., 2008; Heikkilä et al., 2012).

These models have been the main source of estimates of the costs of unplanned mortality and disease on dairy farms and their effect on replacement rates. Stott (1994) estimates the costs of infertility using dynamic programming models to help quantify the value of the trait in the selection index; the study arrives at about 20 GBP (about 25 USD at current conversion rates) per lactation per year as a lower bound and about 80 GBP (around 100 USD) as an upper bound. Heikkilä et al. (2012) calculates the cost of mastitis due to early exit as around 600 EUR (about 660 USD) per exit in Finland also by using dynamic programming. De Vries (2013) estimates the average cost of "involuntary disposal," which includes all of these factors, as 500-1000 USD per exit in the US when not considering lost production. Despite the complexity of these models, they often still do not rationalize the data. The majority of these models estimate that 20-30% of the herd should be culled each year, though the culling rate is usually higher than 30% (De Vries, 2013; Hadley et al., 2006).

The contribution of this paper is to investigate this discrepancy and the role of unplanned mortality costs on dairy farms using an empirical approach instead of a normative approach. The empirical approach uses an economic model to rationalize the data given a structure rather than assume all states are known, which is the most common approach in economics literature on asset replacement (Cho, 2011; Rothwell and Rust, 1997; Schiraldi, 2011). The first paper to take this approach to dairy cow replacement was Miranda and Schnitkey (1995), which found that a large component of the gain from replacement was unobserved to their model. They hypothesized that annual costs that were linear in animal age were responsible for early replacement, but for most farms this parameter was both small and statistically insignificant. The alternative specific constant for the replacement decision, which was the location parameter of the distribution of the unobserved state ϵ, was large and significant compared to other factors in the model. They theorized that this constant represented factors not explicitly modeled in their profit function, including genetic progress and unseen costs of replacement. This paper builds on their results by incorporating unplanned mortality as a risk factor in dairy cow replacement to explain this unobserved benefit to replacement. Since their model did not incorporate any mortality risk, the costs of mortality have manifested instead in the alternative specific constant.

Finally, this paper also contributes more generally to the literature analyzing dynamic decisions using dynamic discrete choice by using newer methods to control for unobserved asset heterogeneity. Hotz and Miller (1993) showed that continuation values in dynamic models could be approximated by empirical choice probabilities, which significantly opened up the number of ways to estimate dynamic models Another recent advancement has been Euler equation conditional choice probability methods (ECCP), which are discrete analogues to Euler equation methods that can be used to derive regression equations (Aguirregabiria and Magesan, 2013; Scott, 2013). Using this regression equation, it is computationally easier to control for unobserved heterogeneity using fixed effects. A threat to identification in these models is unobserved asset attributes that can bias parameter estimates; in the case of dairy cows, certain animals may have permanent traits that make them more prone to replacement. By using the CCP and ECCP method, I can more easily control for this time-invariant component of heterogeneity than in previous methods such as Nested Fixed Point or dynamic logit.

To accurately estimate the costs of unplanned mortality from replacement decisions using these methods, I develop a theoretical model that explicitly embeds the risk of mortality in the manager's replacement decisions. In the next section, I describe the model and how specifically I recover cost and production function parameters to investigate the causes of replacement from data.

3 Theory Model

Consider the case where a dairy farm manager makes an annual decision at time t about a dairy cow at age a_t (measured in year-long "lactations") that produces annual milk output $y(a_t)$. As production is only a function of age, this model holds all other decisions concerning annual production fixed.[1] I assume that there are only two options: keep the current cow or buy a replacement with age one. This means that every dairy cow must be replaced, so I assume that herd size is fixed:

Assumption 1. *Fixed Extensive Decision: the option to leave a stall empty for a year is always dominated by keeping a cow or replacing a cow.*

Without this assumption, we would have to consider three decisions: replace, keep, or empty. The "empty" option would in most cases be dominated by replace or keep unless a dairy farm was scaling down its operation. For simplification, I assume that dairy farms are not interested in scaling down. This also implies that for a dairy farm to maximize profit it should maximize the profits of each stall, so we can focus on profit maximization at the cow level instead of the herd level.

I assume that the manager is risk-neutral and maximizes expected profit for the next lactation. They do so by deciding between the current animal that will have progressed to $a_t + 1$ or an animal at age one. Specifically, $i_t \in \{0, 1\}$, where $i_t = 1$ is sell the current cow and buy a replacement and $i_t = 0$ is keep the current cow. The price of output is p_t and the cost of replacement is c_t. If the current animal is replaced, the expected revenue for the next year will be $p_t y(1) - c_t$.[2]

[1] This assumes there are no "inter-lactation" actions that influence production from the manager's perspective, or variable intensity of use for the asset. This is a common assumption when analyzing dairy production, as reflected in the popularity of using "income-over-feed-cost" as a measure of the profitability of milk production; this measure assumes that there are a fixed number of feed inputs that support milk production and is the profit of producing one pound of milk given milk price and feed prices.

[2] One complication in terms of expectations is that the revenue p_t and cost c_t may not occur at the same time. A further complication is that, since dairy cows produce throughout the year, no one price captures the revenue from one lactation. Also, since US dairy farmers do not seasonally calve, it is very difficult to know which month's price the farmer values next period's revenue at since a cow's lactation could span different months every lactation. For now, I assume adaptive expectations, meaning the manager values next period's revenue at the most recent, observed prices (that is the price prevailing at their last recorded lactation record). See Section 5 for more details on the prices.

3.1 The Role of Mortality

Without mortality, the payoff from deciding to continue with the current animal is just $p_t y(a_t + 1)$. However, it is common for a dairy animal to be "forced" to exit 1) before the annual return is realized (dies in calving) or 2) in the middle of its next production cycle. Dairy cows commonly die in the first one-hundred days of their cycle, when they are weakest, meaning if the animal dies then little to no revenue is realized. Instead, a new animal has to be purchased, meaning the age of the animal regenerates back to one even though they aimed to keep the current cow. An animal dying, however, incurs costs that would not have been incurred had the animal been replaced. These costs include the cost of disposing of the carcass, the costs of treating a sick animal that ultimately dies, lost production, or the costs of finding a replacement ahead of schedule (if one was not immediately available).

I model these costs from unplanned mortality as a "penalty" α, which is added to the cost of replacement when the replacement was unplanned.[3] In the case of this forced replacement, the next period's return is $p_t y(1) - c_t - \alpha$. The probability that an animal will survive to the next period is some function of age $S(a_t)$. The return from continuing with the current occupant is thus a weighted combination between these two payoffs:

$$R(a_t, p_t, c_t, i_t) = \begin{cases} p_t y(1) - c_t & i_t = 1 \\ S(a_t)\Big(p_t y(a_t + 1)\Big) + (1 - S(a_t))\Big(p_t y(1) - c_t - \alpha\Big) & i_t = 0 \end{cases}$$

This mortality cost can explain earlier than expected replacement of animals because the manager now has an incentive to replace the animal to avoid paying α. The current period return from replacement, that is $R(a_t, p_t, c_t, i_t = 1) - R(a_t, p_t, c_t i_t = 0)$, would be:

$$(1 - S(a_t))\alpha + S(a_t)\Big(p_t y(1) - p_t y(a_t + 1) - c_t\Big) \tag{1}$$

If $S(a_t) = 1$ for all ages, which is to say that animals never die, then α does not affect the decision to replace. Managers would replace when $p_t y(1) - p_t y(a_t + 1) - c_t > 0$, that is when the marginal return from resetting the age is more than the replacement cost.

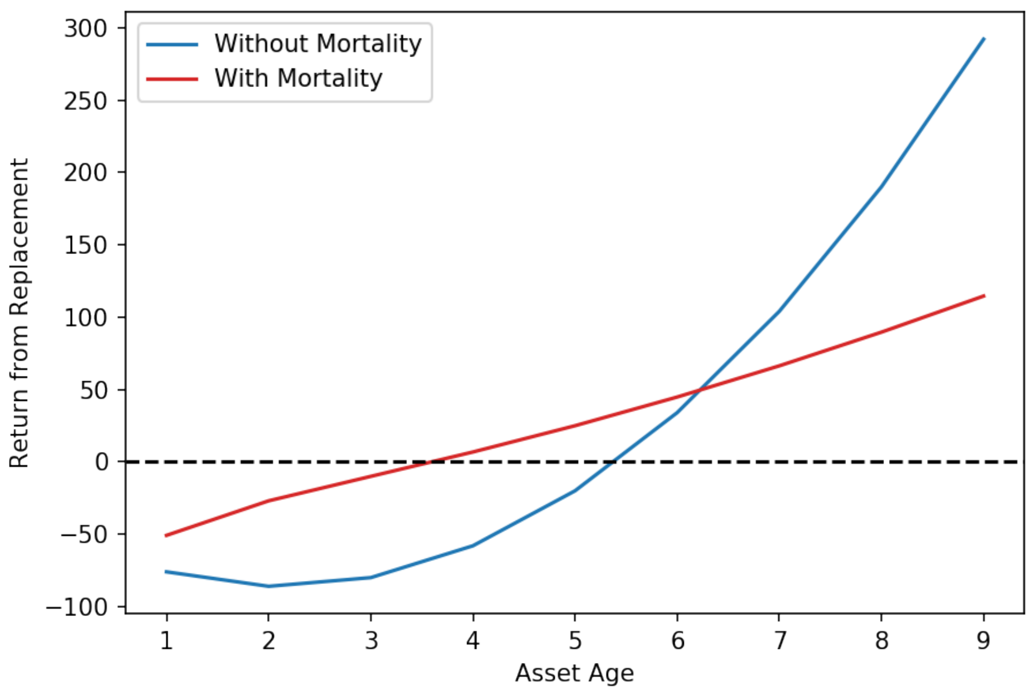
However, since quite a bit of exit is due to sickness or disease, consider the case where the probability $S(a_t)$ is decreasing in age (intuitively, older cows are more likely to have to

[3]Note that this is independent of production; there are good arguments for making the penalty term proportional to the expected output (some percentage of production is lost). I model it here more simply as independent of production while noting that the real parameter would be quite heterogeneous across cows and herds (a point I return to in Section 6).

be removed), or at least decreasing after some point. As age progresses, α will get larger and the the previous criterion will get smaller; intuitively, since the manager increasingly expects the animal to reset to age one because of mortality, the marginal return from a replacement decreases in importance as age increases. Thus the higher cost of mortality will cause higher rates of replacement than expected not considering mortality.

As an illustration, consider the parametric example in Figure 1 using the functional forms specified in Table 1. The survival probability is specified as a variant of the Weibull hazard rate function and is monotonically decreasing as age increases. The payoffs with and without asset failure are graphed in red and blue.

Figure 1: Payoffs



State Values		Parameter Values		Survival Function
p	10	(β_1, β_2)	(5,-0.8)	$S(a) = 0.8a^{-0.5}$
c	50	α	50	

The blue line shows that under no mortality the optimal policy is to replace at about age five, about two years after the production function is maximized ($\beta_1/2\beta_2 = 3.125$). However, with the penalty, the optimal replacement age is two years younger, at about three, because the risk of incurring mortality cost is too high. The only case when the assets will be replaced at the same time regardless of output price or replacement cost is when $\alpha = 0$.

3.2 Other Causes of Early Replacement

Above I demonstrate how a high cost of mortality can cause a "premature" replacement, in the sense that the asset would be replaced before it would typically be optimal. However, there are other candidate explanations for why dairy cattle may be replaced earlier than expected. I detail three of them here to explain how they are included in the model to test their relative importance in determining replacement.

First, Miranda and Schnitkey (1995) claimed that maintenance costs of aging cattle can explain early replacement. They model this by including a "maintenance cost" function that is linear in age. These costs are theoretically different than mortality costs; they represent added costs of having older animals while not affecting the transition probability. If these costs are included, age affects the current period payoff both linearly through maintenance cost and non-linearly through $S(a)$. The role of $S(a)$ is different than maintenance cost, however, for two reasons. One, because of mortality, prices should not affect the decision the same way at every age. In Figure 1, this is why the slope of the red line is different than the slope of the blue line. Two, $S(a)$ affects both the payoff and the transition probability of age; in the event of $i_t = 0$, next period the manager will have an asset with age $a_t + 1$ with probability $S(a_t)$ and an asset with age 1 with probability $1 - S(a_t)$. This means that $S(a_t)$ not only affects the current period payoff but also the continuation value (see the derivation in Section 4). To compare the linear cost function to this model, I include a linear maintenance cost function $M(a_t) = \gamma a_t$ in the payoff.

Another motive for early replacement is observed asset performance. I model this by including an additional endogenous state: the production shock η_t. This state can be thought of as the deviation from the asset's expected performance, which is $y(a_t)$, which could be a deviation from a group average, for example. Since the production function is only a function of age, this state captures other aspects of productivity that are deviations from this simplistic, quadratic production function. This state, like age, is endogenous because it is influenced by the choice i_t. When the asset is not replaced, the next cycle's shock η_t is drawn from $\eta_t \sim N(\rho\eta_{t-1}, \sigma_\eta)$, where ρ is an autocorrelation coefficient. When the asset is replaced, η is expected to be zero, or $\eta_t \sim N(0, \sigma_\eta)$. When observing past performance, the manager will take into account that replacing will protect against a negative shock, and will also be more likely to keep an animal that is doing well.

Finally, the rate of genetic progress for milk production is a strong incentive to replace early. Holding an old asset in production when an even better asset is available is a significant opportunity cost. Modeling technological progress in an asset replacement model is well established in the literature (e.g. Perrin (1972) and Bethuyne (1998)), but is not the focus of this paper. We would like to incorporate the fact that the payoff will change over time, however, and that the option to replace in the year 2011 does not have the same value as 2012. I model this by including a time trend in the payoff for replacement, which allows the payoff from replacement to grow linearly over time.

3.3 The Dynamic Model

Now consider an infinite horizon, dynamic program, with discount rate $\delta \in (0, 1)$. The Bellman equation is:

$$V(x_t, z_t) = \max_{i_t \in \{0,1\}} R(x_t, z_t, i) + \epsilon(i_t) + \delta E(V(x_{t+1}, z_{t+1})|x_t, z_t, i_t) \tag{2}$$

where x_t and z_t are shorthand for endogenous states (x_t) a_t and η_t and exogenous states (z_t) p_t and c_t.[4] In addition to including the value function V in the payoff, there is also an additional state ϵ which represents the influence of states not observed in data. The two assumptions about the unobserved state ϵ to estimate the parameters in a regression model are:

Assumption 2. *Conditional Independence: The transition of states x and z are conditionally independent of ϵ.*

Assumption 3. *Additively Separable Type 1 EV: The error ϵ is additively separable in the payoff and is distributed Type 1 Extreme Value.*

The first assumption is common to models of this type, as in Rust (1987), but is also a reasonable assumption given that the decision is the manager looking forward. In this case, R represents an ex-ante payoff and ϵ expectational noise. Hotz and Miller (1993) argues that in this case, ϵ satisfies conditional independence by construction. The power of this assumption is that it frees us from having to take an integral over V with respect to ϵ but only over the states x_t and z_t. The second assumption allows for a very convenient functional form for the probabilities and allows us to estimate the parameters in a reduced form model. It also allows us to find a very convenient way to represent the differences in value functions, which is detailed in the next section.

The transition functions for the exogenous states p_t and c_t are modeled simply as normally distributed random variables that are AR(1). The shock distribution is similarly modeled as normally distributed with variance σ_η^2. The transition of age is less straightforward than in previous models due to the probability of mortality. The state a_t will always transition

[4] Assuming that the prices p_t and c_t are exogenous is equivalent to assuming that dairy farmers are price takers. This is generally true for dairy farms, especially dairy farms in Wisconsin where very few farms are above 200 cows, so market power is very dispersed.

to 1 if $i_t = 1$, but otherwise it will return to 1 with probability $1 - S(a_t)$ and transition to $a_t + 1$ with probability $S(a_t)$. This also implies that the continuation value when $i_t = 0$ is a weighted combination of $\bar{V}_1(x_t, z_t) = E(V(x_{t+1}, z_{t+1})|x_t, z_t, i_t = 1)$ and $\bar{V}_0(x_t, z_t) = E(V(x_{t+1}, z_{t+1})|x_t, z_t, i_t = 0)$. When entering the "unplanned replacement" state of nature, the value function should proceed as if a new asset was purchased.

Taking shocks and maintenance cost into mind, we can rewrite the payoff function:

$$R(x_t, z_t, i_t) = \begin{cases} \tau t + p_t y(1) - M(1) - c_t & i_t = 1 \\ S(a_t)\Big(p_t y(a_t+1) + \rho\eta_t p_t - M(a_t+1)\Big) + & i_t = 0 \\ (1 - S(a_t))\Big(p_t y(1) + \rho\eta_t p_t - M(1) - c_t\alpha\Big) & \end{cases}$$

$$\begin{aligned} \text{s.t.} \quad & M(a_t) = \gamma a_t \\ & y(a_t) = \beta_0 + \beta_1 a_t + \beta_2 a_t^2 \\ & \{x_t, z_t\} = \{a_t, \eta_t, p_t, c_t\} \end{aligned}$$

Where we now include the effect of maintenance cost, production shocks, and a time trend value τ. The shock η_t to always affect the payoff when $i_t = 0$; this is to take into account the fact that asset failure can have repercussions related to the previous cycle's performance. [5]

Taking $S(a_t) = S_t$ as essentially another state variable, this will be the difference in current period payoffs:

$$\begin{aligned} R(x_t, z_t, i_t = 1) - R(x_t, z_t, i_t = 0) =& \mu + \tau t + \alpha(1 - S_t) - \rho\eta_t p_t - S_t c_t + \gamma S_t a_t \\ & - (\beta_1 + 2\beta_2) S_t a_t p_t - \beta_2 S_t a_t^2 p_t \qquad (3) \\ R(x_t, z_t, i_t = 1) - R(x_t, i_t = 0) =& \theta X \\ \text{s.t.} \quad X =& \Big(1, t, 1 - S_t, \eta_t p_t, S_t c_t, S_t a_t, S_t a_t p_t, S_t a_t^2 p_t) \\ \theta =& \Big(\mu, \tau, \alpha, -\rho, -1, -\gamma, -(\beta_1 + 2\beta_2), -\beta_2\Big) \end{aligned}$$

where the second line is in matrix form. Our parameter vector is θ and our data vector X. Our goal is to estimate the parameter vector θ, where μ is the difference in means between $\epsilon(1)$ and $\epsilon(0)$. This was termed the "culling premium" in Miranda and Schnitkey (1995), and contains benefits to choosing to replace unexplained by the model. Now we know the

[5] To make the shock only transmit in the case of survival, we need only multiply the term $\eta_t p_t$ by the survival rate S_t in the regression equation that follows. This is left as a robustness check of the specification.

difference in current period payoffs but must now take into account the continuation value, which is the difference between two value functions. Previous methods use the nested fixed point algorithm and value function iteration to compute the continuation value; more recent methods approximate the solution of the value function iteration using basis functions. In the next section, I show how the assumptions on ϵ allow a convenient form for the continuation value which is a function of empirical replacement probabilities using the inversion theorem of Hotz and Miller (1993).

Table 1: Model Summary

Endogenous States (x_t)	a_t	Age
	η_t	Production shock
Exogenous States (z_t)	p_t	Output price
	c_t	Replacement cost
Controls	$i_t \in \{0,1\}$	Replacement decision
Technology	$y(a_t) = \beta_0 + \beta_1 a_t + \beta_2 a_t^2$	Total milk output of lactation a_t.
	$M(a_t) = \gamma a_t$	Maintenance cost function.
	$S(a_t)$	Survival rate
	$P(a_{t+1} = 1 \vert i_t) = \begin{cases} 1 & i_t = 1 \\ 1 - S(a_t) & i_t = 0 \end{cases}$	Evolution of age a
	$P(a_{t+1} = a_t + 1 \vert i_t) = \begin{cases} 0 & i_t = 1 \\ S(a_t) & i_t = 0 \end{cases}$	
Payoff, $R(x_t, z_t)$	$\mu + \tau t + p_t y(1) - \gamma - c_t$	If $i_t = 1$
	$S(a_t)\Big(p_t y(a_t+1) + \rho\eta_t p_t - \gamma(a_t+1)\Big) +$ $(1 - S(a_t))\Big(p_t y(1) + \rho\eta_t p_t - \gamma - c_t - \alpha\Big)$	If $i_t = 0$
Parameters	β_1, β_2, γ	Production and cost parameters
	$\delta \in [0,1)$	Discount factor
	ρ	Shock correlation
	τ	Time trend
	α	Cost of mortality
	μ, λ	Location and scale of error term

4 Methodology

4.1 CCP Method

When $\delta > 0$, the decision to replace also takes into account the effect that replacement has now on future decisions, which is $\Delta V(x,z) = E(V(x',z')|x,z,1) - E(V(x',z')|x,z,0)$. In Rust (1989), the solution to this problem is to solve the value function iteration problem to find $V^*(x,z)$ across all states, and calculate $\Delta V(x,z)$ and include it in the maximum likelihood estimation. Unfortunately, since $\Delta V(x,z)$ is also a function of parameters θ, in any optimization routine the value function iteration must be solved every time that the likelihood is calculated.

Instead of using the Nested Fixed Point method, I use the Conditional Choice Probability (CCP) estimator derived in Hotz and Miller (1993) and expanded on in Arcidiacono and Miller (2011). Call the probability of taking action k conditional on endogenous states x_t and exogenous states z_t, the "conditional choice probability," $P_k(x_t, z_t)$. Also denote the transition probabilities for x and z as f_x and f_z.

Finally, define the the "conditional value function," the payoff from choosing action i and acting optimally from then on, $v(x_t, z_t, i_t)$. Their recursive relationship is the following:

$$v(x_t, z_t, i_t) = R(x_t, z_t, i_t) + \delta E(\bar{V}(x_{t+1}, z_{t+1})|x_t, z_t, i_t) \tag{4}$$

where $\bar{V}$ is the "ex-ante" or "unconditional" value function when every decision after this one was made optimally, and so does not depend on i_t.

According to Lemma 1 of Arcidiacono and Miller (2011), there is a function ψ such that $\psi(x_t, z_t, i_t) = \bar{V}(x_t, z_t) - v(x_t, z_t, i_t)$ (the main result of the "Inversion Theorem" from Hotz and Miller (1993)). Now using the function ψ we can substitute $\bar{V}$ in Equation 4.

$$v(x_t, z_t, i_t) = R(x_t, z_t, i_t) + \delta E(v(x_{t+1}, z_{t+1}, k) + \psi(x_{t+1}, z_{t+1}, k)|x_t, z_t, i_t)$$

where k is an arbitrary choice. The reason k can be any given choice is that the term ψ will essentially "penalize" the returns if this is not the optimal action (Arcidiacono and Miller, 2011; Hotz and Miller, 1993).

Hotz and Miller (1993) shows that, because of Assumption 3, $\psi(x_t, z_t, k) = .577 - ln(P_k(x_t))$ (.577 being Euler's constant) where $P_k(x_t)$ is the CCP of taking action k. The most useful choice of k is $k = 1$, which is to assume that all cows are replaced next period and exploit the principle of "limited dependence." Limited dependence is a special feature of models that involve a "renewal decision," which is a decision that resets one of the states so that previous actions have no further effect on the future. In our case, replacing a cow renews the state a_t back to 1, and if the cow is replaced at $t+1$ then there will always be a new cow in $t+2$ which is unaffected by decisions in t (see the example in Aguirregabiria and Magesan (2013) for a specific application to dairy cattle replacement).

So if $k = 1$, then the difference in value functions $v(x_t, z_t, i_t = 1) - v(x_t, z_t, i_t = 0)$ is only a function of the payoffs in period t and $t+1$, since the decision is identical from $t+2$ onward. This means we now only have to estimate the payoffs in period $t+1$.

$$\begin{aligned}
v(x_t, z_t, 1) - v(x_t, z_t, 0) =& R(x_t, z_t, 1) - R(x_t, z_t, 0) \\
&+ \delta\Big(E(R(x_{t+1}, z_{t+1}, 1) - \psi(x_{t+1}, z_{t+1}, 1)|x_t, 1) \\
&- E(R(x_{t+1}, z_{t+1}, 1) - \psi(x_{t+1}, z_{t+1}, 1)|x_t, 0)\Big) \\
=& R(x_t, z_t, 1) - R(x_t, z_t, 0) \\
&+ \delta \sum_{z_{t+1}=1}^{Z} \sum_{x_{t+1}=1}^{X} \Big(R(x_{t+1}, z_{t+1}, 1) - \psi(x_{t+1}, z_{t+1}, 1)\Big) \\
&\Big(f_x(x_{t+1}|x_t, 1) - f_x(x_{t+1}|x_t, 0)\Big) f_z(z_{t+1}|z_t)
\end{aligned} \tag{5}$$

Recalling that $\psi(x_t, z_t, k) = .577 - ln(P_k(x_t))$, this reduces to:

$$\begin{aligned}
v(x_t, z_t, 1) - v(x_t, z_t, 0) =& R(x_t, z_t, 1) - R(x_t, z_t, 0) \\
&+ \delta \sum_{x_{t+1}=1}^{X} \sum_{z_{t+1}=1}^{Z} \Big(R(x_{t+1}, z_{t+1}, 1) + lnP_1(x_{t+1}, z_{t+1})\Big) \\
&\Big(f_x(x_{t+1}|x_t, 1) - f_x(x_{t+1}|x_t, 0)\Big) f_z(z_{t+1}|z_t)
\end{aligned}$$

noting that we can factor out f_z because, being exogenous states, they are not affected by the decision i.[6] So now to calculate the relative payoff from replacing, which is $v(x_t, z_t, 1) - v(x_t, z_t, 0)$, we only need to know the CCP's across different states, $P_1(x_t, z_t)$, and the difference in transition probabilities, $f(x_{t+1}|x_t, 1) - f(x_{t+1}|x_t, 0)$. For identification in dynamic discrete choice, we also have to normalize one payoff to zero (Magnac and Thesmar, 2002). In this case, I choose to normalize the payoff to replacement to zero, which is equivalent to subtracting $R(x_t, z_t, 1)$ from both payoffs. Remembering that $S(a_t) = f(a_t + 1|a_t, 0)$

[6] Also note that it is now easier to see why ψ "penalizes" the payoff when $P_1 \neq 1$; if $P_1 < 1$, then $\psi < 0$, but the payoff is unchanged if $P_1 = 1$.

and $1 - S(a_t) = f(1|a_t, 0)$, we can finally factor out the continuation value ΔV to get the following:

$$\Delta V = FV_1 + S(a_t)FV_2 \quad \text{s.t.} \tag{6}$$

$$FV_1 = \sum_{z_{t+1}=1}^{Z} \sum_{\eta_{t+1}=1}^{E} \Big(lnP_1(1, \eta_{t+1}, z_{t+1})\Big)\Big(f_\eta(\eta_{t+1}|\eta_t, 1) - f_\eta(\eta_{t+1}|\eta_t, 0)\Big) f_z(z_{t+1}|z_t)$$

$$FV_2 = \sum_{z_{t+1}=1}^{Z} \sum_{\eta_{t+1}=1}^{E} \Big(lnP_1(1, \eta_{t+1}, z_{t+1}) - lnP_1(a_t + 1, \eta_{t+1}, z_{t+1})\Big)\Big(f(\eta_{t+1}|\eta_t, 0)\Big) f_z(z_{t+1}|z_t)$$

(see Appendix A for derivation).

By having a first stage estimate of P_1, we can now include FV_1 and $S(a_t)FV_2$ as two additional regressors in the model to proxy for the continuation value if we estimate transition probabilities for p_t, c_t and η_t.

4.2 First Stage Estimation

The above is a two-step estimator: first calculate the CCP $\hat{P}$ and then estimate the regression equation. The first step, however, requires calculation of conditional choice probabilities P, both in-sample and out of sample; we do not necessarily observe all combinations of ages, production shocks, and prices to have accurate estimates of P_1. A common way to estimate P_1 is some kind of bin estimator (Scott, 2013) or a logit model with several combinations of the state variables used as predictors (Arcidiacono and Miller, 2011). The issue with the first method is having to make judgements on the size of the bins, which can be tricky when states are fully continuous (as in my case here with output price p_t and replacement cost c_t). The issue with a flexible logit is that using so many combinations of state variables is very likely to have good in sample performance but poor out of sample performance (due to overfitting).

I choose to predict P_1 using a random forest algorithm as a compromise between these two methods for the following two reasons. First, random forest prevents the econometrician from having to choose bins because it essentially selects the bins using cross validation. Many of the hyper parameters in a random forest, such as number of leaves or minimum sample on a leaf, are essentially changing the bin size. Random forest is essentially a more sophisticated bin estimator that frees the econometrician from having to choose bins for continuous variables. When a random forest is trained using brier-score loss then rather than doing classification it will deliver the desired probabilities of replacement (Boström, 2008).

A second reason is that using a method with cross-validation will prevent the model from over-fitting and causing poor out of sample performance. A logit model with many combinations and polynomial expansions of state variables as recommended in Arcidiacono and Miller (2011) is a classic example of a model that will overfit; it will produce accurate in-sample probabilities but will do poorly at predicting combinations of age, shocks, and prices that are not seen in the data. This will produce inaccurate estimates of ΔV in particular. To address this, I feel a machine learning model trained using cross-validation is a better method than logit with multiple, polynomial interactions.

Another part of the first stage estimation is the transition probabilities f_z and f_x. I estimate the transition probabilities of exogenous market states using an AR(1) regression where the error is normally distributed. I use the same regression to find flow probabilities for production shock η, except the data for η comes from data on the animal's milk production. Specifically, a milk production model from Kearney et al. (2004) is used to predict milk yield for a given animal; the residual for each lactation is my estimate for η. This proxies for the production shock in the structural model because it the production of the animal net of any observable predictor of milk production on the farm; since the milk production model uses herd fixed effects, the residual is actually its deviation from its herd mate. I argue that this is the best approximation of a "deviation" from its expected return that the manager would likely act on. More information about the milk production model is given in the Appendix B.

Finally, the survival function S_t can either be thought of as computed from data or a function of biology (and thus exogenously imposed). Due to age being a discrete variable, there is no reason for any parametric assumption. The literature on dairy cow culling calculates the probability of "involuntary exit" for each age, for example in Stott (1994) and Van Arendonk (1985). In this particular application, I assume that the functional form of S_t is exogenously imposed. Essentially S_t is an attribute of the technology of dairy cows rather than a choice variable. While it is known that management actions can have an effect on the rates of exit, it is not clear that from the manager's perspective this is actionable. The management actions that have an effect on cow death and infertility are broad structural changes which cannot be changed in the short run. Still, this particular functional form might not be the one that dairy farmers expect. We can allow *shifts* in the level of S_t while still imposing a curvature common to all farms, which can be controlled for using fixed effects. The shape of S_t is estimated from the data based on the percentage of cows at each age that exit the herd in the first 120 days of their laction, which is most often an unplanned exit.

4.3 Second Stage Estimation

Logit Method With the above derivation, we can estimate the equation with logit by simply including an estimate of the continuation value ΔV as an additional regressor. Assuming the parameter δ as the coefficient on ΔV gives us enough degrees of freedom to estimate the parameters of our model (see Arcidiacono and Miller (2011) for an explanation of when δ is identified). Given an estimate of the survival function S, we have the following reduced form logit model that maps to our structural coefficients.

$$P(i_{jt} = 1|x_{jt}, z_{jt}) = \frac{e^{\lambda\theta X+\delta FV}}{1 + e^{\lambda\theta X+\delta FV}} \tag{7}$$

for herd-animal j at time t. The parameter λ is the scale parameter of the distribution. In order to interpret the coefficients in dollar terms, we must divide through by λ, which we can estimate as the coefficient on the term $S_t c_t$.

The reduced form coefficients conditions, after we divide through by λ, are:

$$X = \Big(1, 1 - S_{jt}, \eta_{jt}p_t, S_{jt}c_t, S_{jt}a_{jt}, S_{jt}a_{jt}p_t, S_{jt}a_{jt}^2 p_t\Big)$$
$$\theta/\lambda = \Big(\mu, \alpha, -\rho, -1, \gamma, -(\beta_1 + 2\beta_2), -\beta_2\Big)$$

$$\begin{array}{lll} \theta_0 = \mu & \theta_1 = \alpha & \theta_2 = -\rho \\ \theta_4 = \gamma & \theta_5 = -\beta_1 - 2\beta_2 & \theta_6 = -\beta_2 \end{array}$$

So we can recover the structural parameters:

$$\begin{array}{lll} \mu = \theta_0 & \alpha = \theta_1 & \rho = -\theta_2 \\ \gamma = \theta_4 & \beta_1 = \theta_5 - 2\theta_6 & \beta_2 = -\theta_6 \end{array}$$

Note that here θ_1 is essentially estimating the "willingness-to-pay" for a lower mortality rate, $1 - S_t$. In this structural model, this is equal to the cost of mortality, α.

In contrast to previous work, specifically Miranda and Schnitkey (1995) and Aguirregabiria and Magesan (2013), I do not estimate the parameters of the production function from outside the structural model. Instead, the production function parameters β_1 and β_2 are identified off of interactions between age, survival rate, and the output price. Were the paremeters to be estimated with milk production data and then plugged into the model, this would be assuming that the econometric estimates are the parameters the manager assumes. Unfortunately, this ignores the fact that a cow's milk production curve may be perceived differently by the manager than what could be discovered from a regression. For example, the farmer may have information about their production curve under their own management that would not be uncovered with an econometric regression. The manager may also have a different notion of when an animal's milk production is maximized, which would cause them to replace animals differently than expected. This approach allows any of these things to be true, but changes the interpretation of β_1 and β_2: they are no longer the parameters of the "empirical" production function but rather the parameters of the "perceived" production function from the perspective of the manager. These need not be the same as the estimates of the production function from the dairy science literature.

To understand the "perceived" production function, I also calculate the age of maximum production (a^*) and the "age of free replacement," (a^{free}) that is where $y(1) - y(a+1) = 0$:

$$a^* = -\tfrac{\beta_1}{2\beta_2} = \tfrac{\theta_5 - 2\theta_6}{\theta_6} \quad a^{\text{free}} = -\tfrac{\beta_1 + \beta_2}{\beta_2} = \tfrac{\theta_5 - \theta_6}{\theta_6}$$

If a^* differs from the estimates from dairy science, this is evidence that replacements may happen earlier than expected because managers do not think their animal's production function is the same as what the literature says.

ECCP Method An alternative estimator of $\hat{\theta}$ and arguably more flexible method has been implemented using GMM (Aguirregabiria and Magesan, 2013) or OLS (Scott, 2013) by utilizing moment conditions. Taking logs of both sides of Equation 7 gives the following moment condition:

$$X\theta - \frac{1}{\lambda}(\delta FV + \Delta\nu) = 0 \tag{8}$$

where $\Delta\nu = ln\left(\frac{P(i_t=1|x_t,z_t)}{P(i_t=0|x_t,z_t)}\right)$. Note that for a myopic decision maker ($\delta = 0$), the optimal condition would be the difference in current period profits equal to $\Delta\nu$. Essentially, $\Delta\nu$ is the "cutoff" that the current period relative payoff need to be bigger than in order to make the manager choose to replace. For a non-myopic decision maker, they also weight the continuation value ΔV.

Aguirregabiria and Magesan (2013) uses GMM to estimate the parameter vector after including an estimate of ΔV and $\Delta\nu$. This opens up any GMM method for use in estimating θ. Scott (2013), however, rearranges the above moment condition to get a regression equation, which increases the number of methods that can be used (expecially fixed effects estimators). I derive a regression equation using the moment condition the following way:

$$
\begin{aligned}
X\theta - \frac{1}{\lambda}(\delta FV + \Delta\nu) = 0 \\
\delta FV + \Delta\nu = \lambda X\theta \\
\delta FV + \Delta\nu = \lambda X\theta + \tilde{\xi} \\
\tilde{Y} = \lambda\theta X + \tilde{\xi} \\
\text{s.t.} \quad \tilde{Y} = \delta FV + \Delta\nu
\end{aligned} \tag{9}
$$

Unobserved benefits to replacement now manifest in a regression error $\tilde{\xi}$. Scott (2013) argues that $\tilde{\xi}$ is actually a compound error term, one part expectational error and the other "unobservable shock"; the first of these is arguably uncorrelated with the information the manager, having to do with the evolution of exogenous market variables. The second component is likely not exogenous to payoffs, however; there may be unobserved components of performance that make the replacement decision attractive that are not observed in data. This method, however, significantly opens up the number of analysis options, and in this case it allows me to robust check the logit results with fixed effects more efficiently than using conditional maximium likelihood.

4.4 Endogeneity and Identification

A correct estimate of α depends on there being no unobserved variables correlated with both the failure probability $1 - S_t$ and exit. In Equation 10, α is identified when $E(S\tilde{\xi}) = 0$. There are two main sources of endogeneity concern: herd environment and unobserved health information.

Thomsen and Houe (2006) shows that herd environment affects the survival rate; more intensive dairies, for example, generally have higher rates of exit. Since management intensity, for example number of times milked or type of feed used, is not in the model, this can confound α. The survival rate is also usually affected by large scale decisions, for instance decisions on housing and bedding. It is reasonable to think, however, that herd level factors are fixed in the short run, and can be conditioned out. To eliminate such variation I also need to assume that the survival rate is fixed in its curvature:

Assumption 4. *Exogenous Survival Rate: The curvature of the survival rate S is fixed from the perspective of the manager and the same across all farms.*

I assume that the only heterogeneity in the perception of survival rate across farms is linear shifts in the curve at the farm level. This means that farms can have differing survival rates provided they are linear shifts in the rate *at every age*. Only then do herd fixed effects meaningfully address enodgeneity. There is no evidence one way or the other in the literature about how dairy farmers perceive the survival rate, but it seems likely that dairy farmers would not have radically different ideas of how survival decreases with age. A combination of general knowledge of biology and exposure to extension material explaining cow health means dairy farmers would likely converge roughly on a survival rate curvature that is similar across farms.[7]

Unobserved health information is also a source of endogeneity, but instead this manifests at the animal level. A farmer may observe something about one animal that updates the probability of survival while at the same time affecting replacement. There is no explicit "health state" in the model that captures this information; instead that state would show up in the unobserved state ϵ. If the health states are invariant across animals, we can go one level deeper and use cow level fixed effects. This changes the interpretation of the model, however, because it excludes cows from the analysis with only one observed lactation. In partciular, it ommitts all cows that do not surive past their first lactation. This is not necessarily a bad thing, as it is very unlikely that dairy cows replaced in their first lactation represent the model well. These cows are usually sold in their first lactation alive to generate income or because of a untreatable health problem; the above theoretical model does not fit either of these exits, as neither are initiated with a replacement in mind. In the analysis that follows, I default to cow-level fixed effects to robust check results.

Finally, neither of the above fixed effects methods deal with health shocks, which clearly affect both survival and replacement. One example is the onset of a disease that is unobserved in the data but observed by the manager. One proxy measure of such an event found in the data is somatic cell count (SCC), which is a measure of the bacteria count in the milk. SCC is usually monitored closely by managers because a high SCC indicates the onset of mastitis, the most common lactating dairy cow disease. SCC is available from the data, but is not currently modeled here; arguably this measure produces just as many endogeneity problems, because SCC is also correlated with certain management practices. A more sophisticated model would have to include an indicator like SCC as a dynamic, evolving state that would need to be included in the continuation value FV.

[7] The above assumption is one of simplicity to achieve identification in the model, however. A more realistic model would have a fully endogenous survival rate, which would be a function of farm characteristics. Unfortunately, we do not observe many farm characteristics in this data. Having a more sophisticated model of the survival rate is a subject of future work.

5 Data

My sample is for DHI herds in Wisconsin served by one Dairy Records Processing Center, which covers about 90% of the DHI herds in the state. It covers the period June 2011 to January 2015, which has about 1,500 herds and almost 500,000 cows, bringing a total of 866 thousand lactation records. I look specifically at lactation level records, which record the total fat and protein at the end of the lactation (though test day information is also available). The raw data contained many more herds than 1,500, but were dropped from the analysis based on three criteria. One, herds had to have at least 40 milking cows at any given point in the data. Two, herds had to have been observed from June 2011 up until December 2014 (making a balanced panel). Three, herds that do not have wild fluctuations in herd size (in accordance with Assumption 1), in this case meaning I drop herds whose herd size has a coefficient of variation more than one.[8] For animal level records, lactations above five or six are routinely omitted from analysis of these data (Pinedo et al., 2014; Weigel et al., 2003) because of survival bias; animals that live to be that long are usually extraordinarily good at producing milk and do not represent a typical sample. Including these records causes issues with studying replacement, because the rate of culling for those animals is usually either zero or one. I use animals up to lactation eight, which cuts out only about 1% of the data.

5.1 Exit Rates

A dairy cow "exit" is when a dairy cow leaves that data set. In this data, if a dairy cow leaves the data set less than 6 months before the end of the sample time frame, the cow is considered right censored rather than an exit. Figure 3 shows the rate of exit for cows that are uncensored in the data on average and also by herd size. Exit rates are very high for dairy cows in this sample; around 50% of dairy cows end up leaving the data in their first lactation. Fewer than 30% of dairy cattle make it to their output maximizing age of five. There are not significant differences between herds, however; a slightly smaller percentage of cows exit at the first lactation on smaller farms, though cows at older ages are more likely to be kept on those small farms.

In this data set, we also know whether a dairy cow was bred. Breeding is an important decision to look at because it is, in many cases, a signal of intention for the cow to be kept in the herd. This is not always the case, however. The cow may be bred to be sold pregnant to another farm, for example. It may also be that the manager receives information about

[8] A certain amount of leeway is allowed in the herd size because herd size can fluctuate even when dairy farms are not actively scaling up or down. Herd size can fluctuate temporarily, for example, because a replacement is being purchased or is not quite ready from the replacement herd.

	Record Counts
Herds	1,519
Cows	492,998
Cows with Full History	312,636
Lactation Records	866,011

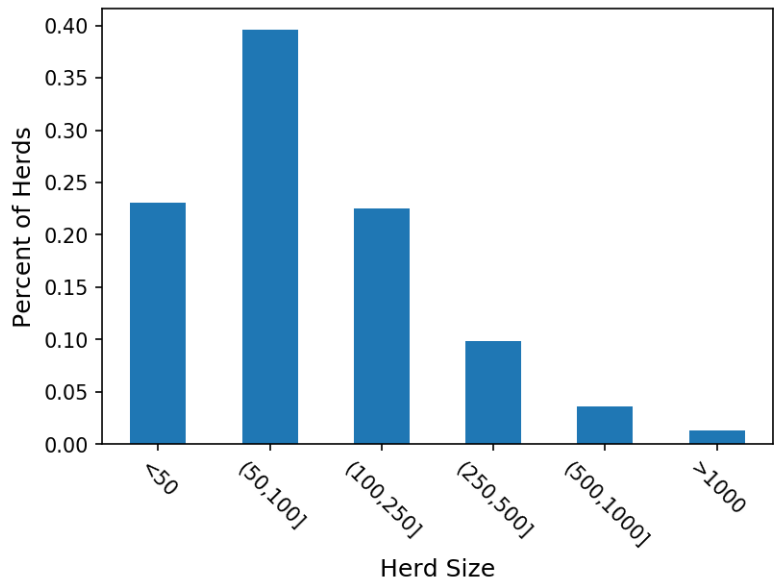


Figure 2: Record Counts and Herd Sizes

the cow later that makes it more worthwhile to sell the cow even if breeding was attempted. While breeding does not indicate the cow was *not* replaced, purposefully *not* breeding is equivalent to taking the cow out of production. Since dairy cows can only produce if they are bred roughly annually, a cow that remains unbred will inevitably stop producing milk. Whether or not this cow is sold or slaughtered, it must be replaced in those herd sizes that are maintaining a roughly constant herd size. Figure 4 shows both the rate at which cows are bred, and also the percentage of them that are pregnant at the time of their exit. The top figure tells us how agressive farms are at attempting breeding while the bottom figure tells us how common it is for a cow to exit pregnant. A cow exiting pregnant is an indication of either a) an agressive replacement policy or b) an unplanned mortality event.

The rates of breeding are quite heterogeneous across herd size. Larger farms are most likely to attempt breeding at the first lactation, leaving only 10% of the herd which they decide to leave unbred. Smaller farms appear to leave more lactation one cows unbred, consistent with the fact that their exit rate in Figure 3 is higher at the first lactation. At all levels they breed less cows. Larger farms may have an advantage in the fact that they can afford to inseminate in more cases at their scale without worrying about the cost; smaller farms may have to be more prudent in their decisions.

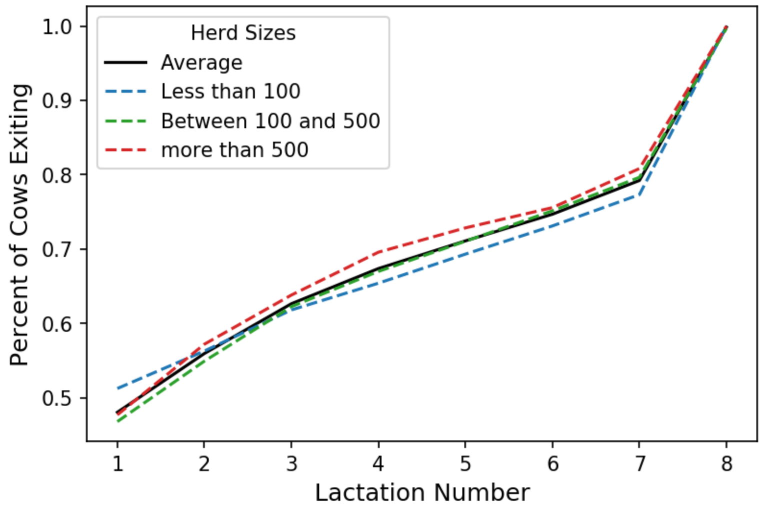


Figure 3: Percent Exiting at Each Lactation
(Uncensored Cows Only)

Unfortunately, herd testing data often does not have enough information about *why* the cows exited. What percentage of the exit rates above are planned replacements versus unplanned mortality? Most importantly for the model, how do we go about ascribing economic motives to each of these exits? Which exit is one which we should study with our model?

Fetrow et al. (2006) outlines three categories of exit for dairy cows: sold alive, sold to slaughter, and died on farm. The first category is generally not analyzed as "culling" because a dairy animal is generally sold alive to generate income separate from the milking operation (for embryos, for calves, ect.). In other words, dairy cows sold alive are sold without considering the productivity of a replacement, and so this decision should not be considered "culling" (Hadley et al., 2006). Death on the farm is similarly not considered "culling," as the manager did not plan the cow to exit.

This leaves the second category: sold for slaughter. What is tricky is that not all of these exits should be studied using a replacement model. In the theoretical model, we made a key distinction between when the cow is planned to be replaced ($i_t = 1$) and when unplanned mortality forces the cow to be replaced (an event happening with probability $S(a_t)$). In reality, however, such a dichotomy is unrealistic. While death is a clear unplanned cull and culling for low production is a clear planned cull, there are a continuoum of circumstances between these that are a mix of planned and unplanned. When a cow gets a treatable disease, for example, if the cow is culled it is not clear whether this is a planned cull because of a drop in productivity or an unplanned cull because of a sudded health event.

As little information is available about why the cow exited (or the economic motive behind it), I analyze replacement using the exit rate as the dependent variable (i_t).This

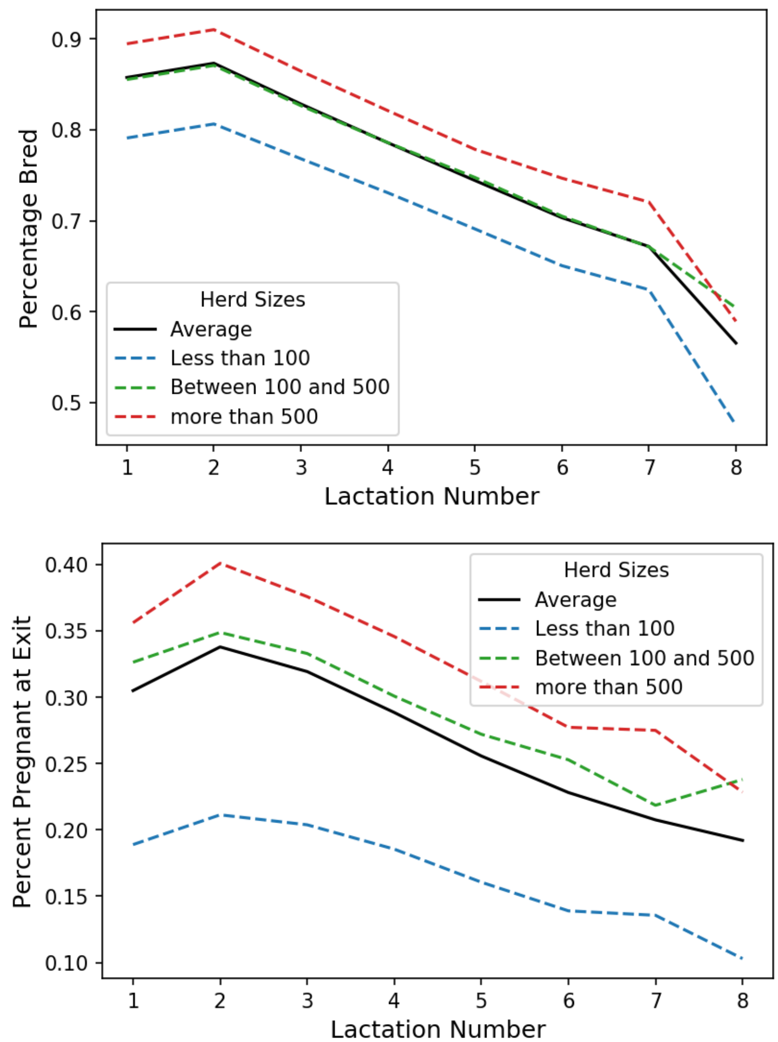


Figure 4: Breeding Attempted and Pregnancy Rate
(Uncensored Cows Only)

strategy particularly relies on anything pushing the exit rate up to be both cow-specific and time-invariant. For example, since more intensive operations can push the death rate higher, this is still allowed if the increase in death rate is a linear shift in the exit rate. If this is the case, then fixed effects will soak up this confounding variation and we can still examine replacement decisions. This is an attractive option because it does not require us to make any major judgements on what sorts of exits are made with economic rationale in mind. It is also a naive assumption, however, that should be relaxed in future work.

5.2 Survival Rate

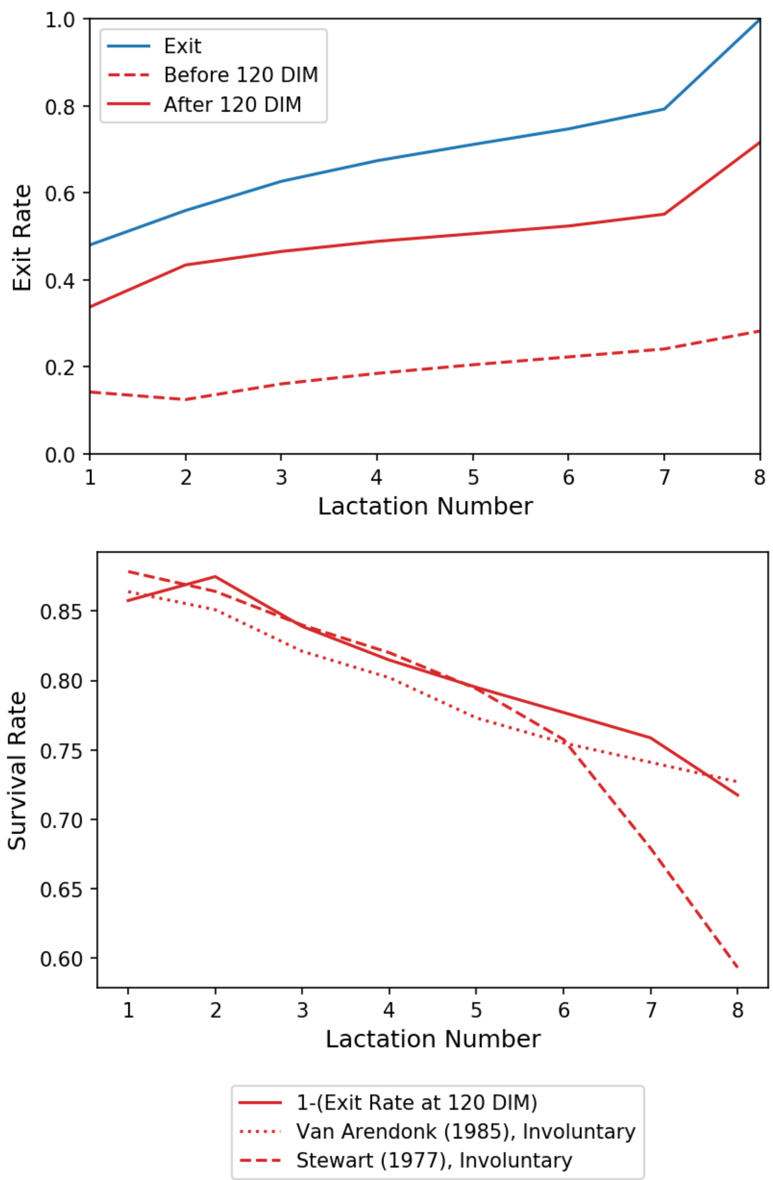
Finally, we also need an estimate of the survival rate. To get an estimate from the data, one thing to keep in mind is that dairy cows are weakest early in their lactation. For this reason, exits in the first four months are most likely to be deaths rather than planned exits. Figure 5 shows what happens when we decompose exit rates by either being at less than or more than 120 days in milk (DIM), which is the number of days into the lactation). The exit rate before 120 DIM has a bathtub shape, which is to be expected with a death rate on dairy farms; cows in their first lactation are at higher risk than older cows. When compared to other rates of “involuntary culling” from the literature, our constructed rate matches more or less shows the same trend; however, rates from the literature are missing the bathtub shape. For the rest of the analysis, I use the estimation of the survival rate from the data since it appears to approximate the survival rates from the literature.

5.3 Market Prices and Shocks

In our model, expected revenue R is a function up of a “profit-margin” state p_t, replacement cost c_t, and “revenue shock” $\eta_t p_t$. For these two prices, I use the income-over-feed-cost (IOFC), a measurement of the profit margin from producing one pound of milk “at test.”[9] The replacement cost is calculated as the salvage value of a 1,400 pound dairy cow minus the market price for a new heifer. Prices are deseasonalized, which makes the assumption that dairy farmers adjust their expectations about price based on seasons. What is particularly difficult about this problem is that dairy cow replacement is not seasonal in Wisconsin. For this reason, it is very difficult to understand at which time dairy farmers pay attention to prices, leading to no obvious choice of what price to use (contrast this with land use examples, which usually use price at the time of harvest or planting to construct expected

[9] The IOFC measurement is the return from producing one pound of milk with “average component values” for a given area (in this case Wisconsin). It includes feed cost, labor cost, and capital cost, and generally reflects “average profitability” of producing one pound of milk.

Figure 5: Exit Rate Decomposed and Compared to Literature Involuntary Rates



revenue e.g. Scott (2013)). As a simple solution, I make the simplifying assumption of adapative expectation only for the term R, which means that managers value next period's expected revenue at the most recent prices. For the expection to calculate FV, I assume the probability of next period prices is derived as rational expectations, meaning from an AR(1) regression. The transition path at the monthly level is estimated using the deasonalized data, which takes into account the seasonal patterns of prices.

I calculate the production shock η from a milk production model, described in Appendix B. The objective of the model is to calculate the performance of the cow relative to its herd mates taking into account, lactation number, lactation length, and milking intensity. The "shock" portion of the production function is calculated as the residual from this regression model. Since the covariate in the model is actually $p\eta$, I calculate the residual for both fat and protein, multiply them by that months Class III component prices, and sum them together.

6 Results

Below I estimate the structural model using logit and the ECCP method. Standard errors in all models are estimated as the standard deviation of the 1000 bootstrap replications. Because estimates of FV are the most inaccurate when they are combinations of states not often seen, regression weights are used in all calculations. The regression weights specifically weight observations less if the particular combination of states is not often seen in the data. In all models, the discount rate is fixed at .99 unless otherwise specified. The tables present the structural parameters, which are non-linear combinations of the reduced form parameters. To be in dollar terms, they are divided by the scale λ, which represents the "marginal utility of money" (since we normalized the coefficient of $S_t c_t$ to be -1). The estimate of λ is presented in all tables at the bottom. To take into account the large amount of heterogeneity in herd size in the data, I estimate the model on different categories of herd size.

6.1 First Stage Estimation

The first stage estimation of FV proceeds in three steps. First, the transition probabilities f_x and f_z are calculated from data on prices and shocks. Second, the CCP P_1 is calculated on every combination of states after being trained on the sample. Finally, FV is derived by taking the expectation over P_1 using the derived distribution of p, c, and η. A random forest algorithm is used to prevent overfitting and assure good out-of-sample properties for the prediction of P_1.

State Transitions I used the following equations to estimate the distributions of p_t, c_t and η_t assuming that the error term is normally distributed:

$$p_t = \mu_p/(1-\rho_p) + \rho_p p_{t-1} + \upsilon^p, \quad \upsilon^p \sim N(0, \sigma_p^2)$$

$$c_t = \mu_c/(1-\rho_c) + \rho_c c_{t-1} + \upsilon^c, \quad \upsilon^c \sim N(0, \sigma_c^2)$$

$$\eta_t = \mu_\eta/(1-\rho_\eta) + \rho_\eta \eta_{t-1} + \upsilon^c, \quad \upsilon^\eta \sim N(0, \sigma_\eta^2)$$

The prediction was done using monthly data that was deseasonalized and CPI adjusted.[10] Table 2 presents the results of these regressions, as well as the results of an AR(1) regression to estimate the initial value of ρ to use in the state transitions for η (which could be different or the same as the ρ calculated in the behavioral model).

Table 2: AR(1) Regressions for State Transitions

	η	p	c
μ	43.087	8.030	291.336
σ	647.430	0.939	54.730
ρ	0.300	0.942	0.791

Conditional Choice Probabilities The random forest algorithm was trained using brier score loss and 10-fold cross validation. Figure 6 shows the performance of the random forest estimator in-sample as compared the empirical probabilities, which are the average exit rates at each lactation number. The biggest deviation in performance is at higher ages given the small number of observations at that level (less than 1% of the data).

6.2 Logit Model Estimates

In Table 3, the first column estimates a model very similar to Miranda and Schnitkey (1995) in order to compare the results from their data. The main difference between their model and mine is that all costs due to aging assets affect payoff linearly and do not impact the continuation value. Unlike in their model, the constant term μ, which they call the

[10]Since it is an annual decision, it may not be technically correct to assume that the next value of p or c is perceived using the monthly pattern of prices. This implies that I am assuming dairy farmers only look at monthly changes in price and then project forward to the entire year of production. Whether this is a realistic assumption is very hard to test, but could be a direction of future research.

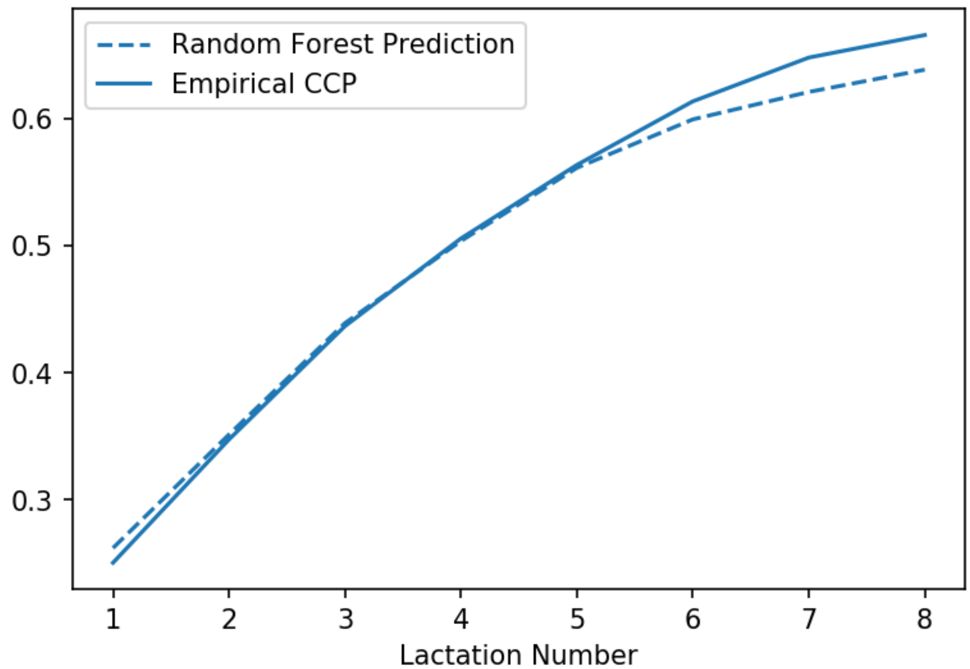


Figure 6: CCP In-Sample Predictions

"premium," is negative. The marginal cost parameter γ is also much larger than in their model. Shock correlation ρ is in the direction as expected, but is quite small, noticably smaller than the .3 found in the regression model (see Table 2). The implied production function parameters suggest that the age at which production is maximized is about three lactations, which is two years ealier than indicated by the dairy science literature. This age, however, is in line with both the findings of Miranda and Schnitkey (1995) and our own milk production model (the results of which are in Appendix B). While we would expect the time-trend to be positive, indicating the increasing benefit from adopting new cows, in this specification it is either near zero or negative. It is possible that our time trend, in this case measured at the month level, is not adequately picking up the effect of genetic progress (or more structure on the term is needed). It could also be the case that farmers do not necessarily think new cows are more attractive.

In the second column is my model specification, which incorporates the survival rate and hazard rate as covariates. The coefficient on the hazard rate, which is α, is positive and significant. The results suggest the cost of unplanned mortality is between 3,000 and 4,000 dollars. The marginal cost of an extra lactation is in the range of 1,300 dollars. Looking at the AIC, it also appears that this model fits the data marginally better than the "no mortality" model of Miranda and Schnitkey (1995) does.

Table 3: Logit Model Estimates

		No Mortality	Mortality Risk, $\delta = .99$	Mortality Risk, $\delta = .95$
Premium	μ	-7981.58	-7535.62	-7141.18
		(1940.3374)	(1799.4385)	(1731.1562)
Time Trend	τ	0.7718	-1.1878	-2.1419
		(2.1086)	(1.6805)	(1.4655)
Penalty	α		4130.92	2999.85
			(1636.1057)	(1437.9537)
MC	γ	1200.58	1354.06	1365.56
		(285.6288)	(323.1812)	(334.1948)
Shock Correlation	ρ	0.124	0.1125	0.1091
		(0.0282)	(0.0253)	(0.0252)
Age of Max	$-\frac{\beta_1}{2\beta_2}$	3.2706	3.2018	3.1544
		(0.0544)	(0.0668)	(0.0662)
Age of Free	$-\frac{\beta_1+\beta_2}{\beta_1}$	5.5412	5.4037	5.3087
		(0.1089)	(0.1337)	(0.1324)
Scale	λ	0.0002	0.0002	0.0003
		(0.00004)	(0.00004)	(0.00005)
Observations		865,990	865,990	865,990
AIC		14.000	16.000	16.000

Bootstrapped "standard errors" in parentheses

6.3 ECCP Model Estimates

Table 4 presents the ECCP model estimates, which control for cow-level fixed characteristics. Additionally, I estimate the model with a zero discount rate (a myopic decision maker). As in the logit model, the model without the survival rate fits the data poorly. Parameters are extraordinarily large, and in the wrong direction; for example, ρ is neither positive nor between zero and one. The production function parameters also give puzzling results. While the myopic model fits the data a little better, the parameters still do not agree; specifically, the production function parameters are negative and too large. The dynamic model results have parameters that are more in line with expectations.

The estimates of unplanned mortality cost and marginal cost are much smaller in this model, however. Unplanned mortality cost is in the range of 1,400 dollars, while marginal cost is around 200 per lactation. As a benchmark, in the year 2011 the average market price for a dairy cow was on average 1,400 USD (USDA, USDA). It is also 400 USD higher than the upper range of the estimates that De Vries (2013) gives for the cost of "involuntary culling." Conditioning out cow specific factors appears to have given us estimates of mortality cost that agree with the existing literature and also with the general market price of dairy cows. Recall that the replacement cost c_t is the price of a replacement cow minus the average salvage value, and α is the added on cost due to unplanned mortality. Essentially, this result is consistent with the idea that dairy farmers expect the costs of unplanned mortality to be the cost of replacing a dairy cow without any salvage value.

6.4 Heterogeneity Across Herds

The meaning of this estimate is still hard to pin down, but more can be learned at looking at how heterogeneous the unplanned mortality cost is across herds. In the descriptive statistics, there was significant heterogeneity in exit and breeding rates across herd size. It is likely that different sizes of dairy farm face different costs, or even different production functions. Table 5 estimates the model on different categories of herd size to determine whether the model's results change significantly across herd type.

In general, there is no significant heterogeneity in λ or ρ across farm size. There appear to be slight different in the age at which production is maximized, with bigger farms expecting cows to maximize slightly sooner than small farms. These differences are not economically significant, however, as farms with between 250 and 1000 cows only differ from small farms by about 0.3 in the age at which cows maximize production.

The most heterogeneous estimates are the unplanned mortality cost α; farms with less than 250 dairy cows pay a mortality cost about four times higher than farms with between

Table 4: ECCP Model Estimates

		No Mortality	Mortality Risk $\delta = 0$	Mortality Risk $\delta = .95$	Mortality Risk $\delta = .99$
Time Trend	τ	-11968.8	-19.9571	-18.2073	-18.1317
		(403.47755)	(0.37001)	(0.38223)	(0.35785)
Penalty	α		-382.708	1380.53	1453.03
			(85.11278)	(84.60353)	(84.95581)
MC	γ	-860.718	400.758	241.184	233.957
		(15.28434)	(8.38322)	(6.91008)	(6.35147)
Shock Correlation	ρ	-6727.46	0.0249	0.0237	0.02362
		(96.40752)	(0.0005)	(0.00052)	(0.00052)
Age of Max	$-\frac{\beta_1}{2\beta_2}$	1.06924	-21.3925	3.04568	3.13374
		(0.00219)	(93.93215)	(0.0506)	(0.05198)
Age of Free	$-\frac{\beta_1+\beta_2}{\beta_1}$	1.13847	-43.785	5.09135	5.26749
		(0.00438)	(187.86429)	(0.1012)	(0.10397)
Scale	λ	0.0015	0.0019	0.0018	0.0018
		(0.000017)	(0.00003)	(0.00003)	(0.00003)
Observations		865,990	865,991	865,990	865,990
Adjusted R^2		0.126	0.526	0.452	0.449

Bootstrapped "standard errors" in parentheses

500 and 1000 cows. Farms with between 250 and 500 dairy cows pay about 1,000 whereas the largest dairy farms, which make up a very small percentage of herds in this data, pay about 1,200. Taking 1,400 as an average market rate for a dairy cow, large farms appear to pay less than the market price for a dairy heifer whereas small farms pay more. Assuming that the cost of unplanned mortality here is related to the price of a replacement, this suggests that large farms can purchase or raise replacements cheaper than small farms can. This is a sensible explanation since large farms can operate at a scale where their costs to raise heifers is cheaper than small farms simple because of economies of scale. These costs are not necessarily related to the price of a replacement, however. Instead, small dairies may also disproportionately pay more for disposal of cows that die on farm or have higher costs for treating diseases.

7 Welfare Analysis

As a structural parameter, α is the added cost to replacement induced by an unplanned mortality event. This does not necessarily calculate a comprehensive cost of unplanned replacement, however; it does not take into account, for example, the effect that $S(a_t)$ has

Table 5: Model Results Across Herd Size

		Less than 100	100 to 250	250 to 500	500 to 1000	More than 1000
Time Trend	τ	-16.44	-19.0298	-19.305	-16.2811	-16.9584
		(0.7209)	(0.7704)	(0.8519)	(0.9637)	(0.8428)
Penalty	α	2072.00	2313.28	1067.77	566.106	1214.3
		(182.9017)	(202.6049)	(184.7803)	(232.904)	(194.496)
MC	γ	215.171	245.002	256.533	229.746	177.749
		(13.1448)	(14.2751)	(15.0209)	(17.8267)	(13.7098)
Shock Correlation	ρ	0.0205	0.0243	0.0247	0.025	0.0202
		(0.0012)	(0.0013)	(0.0012)	(0.0014)	(0.001)
Age of Max	$-\frac{\beta_1}{2\beta_2}$	3.2705	3.031	2.9809	2.9913	3.4733
		(0.0969)	(0.1205)	(0.1141)	(0.1447)	(0.1057)
Age of Free	$-\frac{\beta_1+\beta_2}{\beta_1}$	5.5411	5.062	4.9618	4.9826	5.9467
		(0.1938)	(0.2411)	(0.2281)	(0.2895)	(0.2114)
Scale	λ	0.0016	0.0017	0.0018	0.0018	0.0023
		(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Observations		193,912	204,677	201,008	142,448	123,945
Adjusted R^2		0.443	0.458	0.472	0.468	0.413

Bootstrapped "standard errors" in parentheses
Discount rate set to .99

on the rest of the states. Specifically it does not take into account the lessening importance of $p_t(y(1)-y(a_t+1))-c_t$, which involves the marginal benefit of replacing given the production technology.

Earlier, we saw how including $S(a_t)$ is the distinguishing factor from Miranda and Schnitkey (1995). The payoff with and without the probability of unplanned mortality is:

$$\begin{aligned}\theta^0 X_{jt}^0 &= \mu + \tau t + \alpha(1-S_{jt}) - \rho\eta_{jt}p_t - S_{jt}c_t + \gamma S_{jt}a_{jt} \\ &\quad - (\beta_1 + 2\beta_2)S_{jt}a_{jt}p_t - \beta_2 S_{jt}ta_{jt}^2 p_t + \delta FV_{jt}^1 + \delta FV_{jt}^2 S_{jt} \\ \theta^1 X_{jt}^1 &= \mu + \tau t + \rho\eta_{jt}p_t - c_t + \gamma a_{jt} \\ &\quad - (\beta_1 + 2\beta_2)a_{jt}p_t - \beta_2 a_{jt}^2 p_t + \delta FV_{jt}^1 + \delta FV_{jt}^2\end{aligned}$$

Since $\theta^1 X_{jt}^1$ is the payoff a manager receives when every cow will survive to the next lactation if it is not replaced ($S(a_{jt}) = 1 \forall a_{jt}$), a simple counterfactual to calculate is how much we have to pay farmers to be indifferent between these two states of the world. Essentially, how much money would a farmer pay to completely eliminate mortality risk?

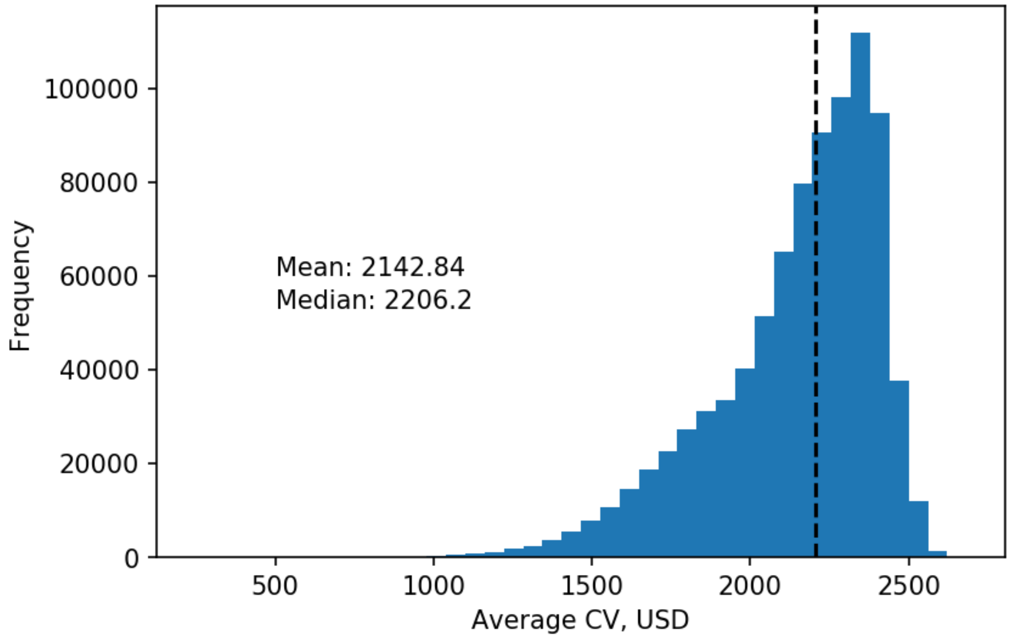
Given Assumption 3, there is a closed form for Compensating Variation (CV) in this model to calculate how much would have to be taken away from farmers with payoff $\theta^0 X_{jt}^0$ to make them indifferent to transitioning to $\theta^1 X_{jt}^1$. Assuming that the value of λ is the same for every farmer, the average CV for transitioning to this payoff function is:

$$E(CV(X^1, X^0, \theta^1, \theta^0)) = \frac{1}{\lambda}\Big(ln(1 + e^{\theta^1 X_{jt}^1}) - ln(1 + e^{\theta^0 X_{jt}^0})\Big)$$

(see Small and Rosen (1981) for derivation). For all of the CV calculations, I fix the replacement cost at the average, which is \$316.80.

On average, farmers are willing to pay about 2,100 per cow to elimate all risk of unplanned mortality. For the parameters in Table 4, the average, expected CV in the data is 2,142.84 USD, about 700 more than the penalty α. The distribution of CV in the data, shown in Figure 7 is assymetric, and is truncated above 2,500 with a long, left tail. Recall from the model that the effect of unplanned mortality is adds to the replacement cost but also decreases the importance of the cow's production function; specifically, it makes the importance of the increase in production less important to the decision. Taken together with the estimate of α from Table 4, farmers would pay about 700 USD to eliminate this effect independent of the penalty. Of the total cost of unplanned mortality to farmers, fully one-third comes from the limiting effect unplanned mortality has on the return from a replacement.

Figure 7: Compensating Variation



8 Discussion and Conclusion

The objective of this study was to investigate the cause of high replacement rates on Wisconsin dairy farms using a structural dynamic model of animal replacement. Specifically, I tested the hypothesis that large costs of "unplanned mortality" on dairy farms cause high replacement rates; to this end, I derived a dynamic discrete choice model that explicitly incorporated a function of age to represent the probability of unplanned mortality in the dairy herd.

My model made significant progress in understanding more about the replacement rationale of dairy farmers by utilizing an expansive dataset and the lastest methods in dynamic discrete choice. This paper has several shortcomings, however, that should be addressed in future work. One issue with this analysis is that I can clearly delineate unplanned and planned mortality in theory but not in data. Most exits are not explicitly one or the other, but are a mix of economic behavior and adverse health events forcing replacement. This highlights a real issue with the "unplanned" and "planned" distinction that is frequently made in the dairy science literature. Most exits Even when exits are explicitly labeled with disposal codes, which many studies use, the reasons given by dairy farmers may be very adhoc and need not represent the actual economic motives behind the decision. My paper has made a vital contribution being one of the first to explore how to incorporate asset failure into an empirical replacement model and explicitly highlights the challenges of making this distinction. Future work should also do more to make several elements of this model realistic to the dairy cow replacement decision. For example, assumptions about the exogeneity of the survival rate, price expectations, and the way genetic progress enters the payoff were made to simplify the analysis. All of these could be relaxed in future work.

I found that, after conditioning out herd and cow fixed effects, that the cost of unplanned mortality was about 1,400 USD per death, roughly the average price of a replacement heifer. Farms under 250 dairy cows paid roughly four times more per death than dairies between 500-1,000 cows (about 2,000 versus 500). My model also estimated the production function parameters implied by the replacement decisions rather than using empirical data on milk production to test the hypothesis that farmers replaced early because they perceived milk production to peak faster than what the literature says. The age at which dairy cows maximize production was calculated as around 3, which is not in line with the dairy science literature, but is in line with the empirical production function calculated from this data. My results suggest that replacements before the "optimal" age at five can be explained by both the costs of unplanned mortality and a production technology fundamentally different than what is found in experiments.

Finally, I calculated compensating variation for how much a farmer would pay to have a payoff function with no probability of unplanned mortality. Aside from adding to the cost of replacement, the probability of survival in the model also decreased the importance of the production function at later ages, an effect I took into account by calculating a measure of compensating variation for removing unplanned mortality risk completely. On average, farmers were willing to pay about 2,100 USD to eliminate unplanned mortality completely, meaning the effect of unplanned mortality on the other factors of replacement is about 700 USD. Of the total cost of unplanned mortality to farmers, fully one-third comes from the limiting effect unplanned mortality has on the return from a replacement.

The policy implications of my results are that small dairy farms may disproportionately be paying the cost of the current trends in animal health. Volatility in milk price is therefore likely to change the distribution of farm size in the U.S. if swings in price disproportionately cause small dairies to exit. In this model, I estimated a measure or unplanned mortality cost that was independent of output price, and since small dairies pay the highest cost for declining cow health they are more vulnerable to price volatility. In times of low milk price, as in the past few years, small farms will exit at higher rates than large dairies because of these breeding strategies. Should it be the policy goal to keep small farms in business, steps have to be taken on breeding dairy cows for longer life, even if it at the expense of production.

References

Aguirregabiria, V. and A. Magesan (2013, December). Euler Equations for the Estimation of Dynamic Discrete Choice Structural Models. In E. Choo and M. Shum (Eds.), *Advances in Econometrics*, Volume 31, pp. 3–44. Emerald Group Publishing Limited.

Arcidiacono, P. and R. A. Miller (2011). Conditional Choice Probability Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity. *Econometrica 79*(6), 1823–1867.

Bar, D., L. W. Tauer, G. Bennett, R. N. González, J. A. Hertl, Y. H. Schukken, H. F. Schulte, F. L. Welcome, and Y. T. Gröhn (2008, June). The Cost of Generic Clinical Mastitis in Dairy Cows as Estimated by Using Dynamic Programming. *Journal of Dairy Science 91*(6), 2205–2214.

Bethuyne, G. (1998). Optimal replacement under variable intensity of utilization and technological progress. *The Engineering Economist 43*(2), 85–105.

Boström, H. (2008, December). Calibrating Random Forests. In *2008 Seventh International Conference on Machine Learning and Applications*, pp. 121–126.

Burt, O. R. (1965). Optimal replacement under risk. *Journal of Farm Economics 47*(2), 324–346.

Cho, S.-J. (2011). An Emprirical Model of Mainframe Computer Investment. *Journal of Applied Econometrics 26*(1), 122–150.

De Vries, A. (2013). Cow Longevity Economics: The Cost Benefit of Keeping the Cow in the Herd. In *DeLaval Cow longevity conference, Tumba, Sweden Aug*, pp. 28–29.

Dechow, C. D., G. W. Rogers, U. Sander-Nielsen, L. Klei, T. J. Lawlor, J. S. Clay, A. E. Freeman, G. Abdel-Azim, A. Kuck, and S. Schnell (2004, October). Correlations Among Body Condition Scores from Various Sources, Dairy Form, and Cow Health from the United States and Denmark. *Journal of Dairy Science 87*(10), 3526–3533.

Fetrow, J., K. V. Nordlund, and H. D. Norman (2006). Invited Review: Culling: Nomenclature, Definitions, and Recommendations. *Journal of dairy science 89*(6), 1896–1905.

Garnsworthy, P. C. (2004, February). The Environmental Impact of Fertility in Dairy Cows: a Modelling Approach to Predict Methane and Ammonia Emissions. *Animal Feed Science and Technology 112*(1), 211–223.

Hadley, G. L., C. A. Wolf, and S. B. Harsh (2006, June). Dairy Cattle Culling Patterns, Explanations, and Implications. *Journal of Dairy Science 89*(6), 2286–2296.

Heikkilä, A. M., J. I. Nousiainen, and S. Pyörälä (2012, January). Costs of Clinical Mastitis with Special Reference to Premature Culling. *Journal of Dairy Science 95*(1), 139–150.

Hotz, V. J. and R. A. Miller (1993, July). Conditional Choice Probabilities and the Estimation of Dynamic Models. *The Review of Economic Studies 60*(3), 497–529.

Kalantari, A. S., H. Mehrabani-Yeganeh, M. Moradi, A. H. Sanders, and A. De Vries (2010, May). Determining the Optimum Replacement Policy for Holstein Dairy Herds in Iran. *Journal of Dairy Science 93*(5), 2262–2270.

Kearney, J., M. Schutz, P. Boettcher, and K. Weigel (2004). Genotype× environment interaction for grazing versus confinement. i. production traits. *Journal of dairy science 87*(2), 501–509.

Knaus, W. (2009, May). Dairy cows trapped between performance demands and adaptability. *Journal of the Science of Food and Agriculture 89*(7), 1107–1114.

Magnac, T. and D. Thesmar (2002). Identifying dynamic discrete decision processes. *Econometrica 70*(2), 801–816.

McConnel, C. S., J. E. Lombard, B. A. Wagner, and F. B. Garry (2008, April). Evaluation of Factors Associated with Increased Dairy Cow Mortality on United States Dairy Operations. *Journal of Dairy Science 91*(4), 1423–1432.

Mellado, M., E. Antonio-Chirino, C. Meza-Herrera, F. Veliz, J. Arevalo, J. Mellado, and A. De Santiago (2011). Effect of lactation number, year, and season of initiation of lactation on milk yield of cows hormonally induced into lactation and treated with recombinant bovine somatotropin. *Journal of dairy science 94*(9), 4524–4530.

Miranda, M. J. and G. D. Schnitkey (1995, December). An Empirical Model of Asset Replacement in Dairy Production. *Journal of Applied Econometrics 10*(S1), S41–S55.

Newman, D. H. (2002, January). Forestry's golden rule and the development of the optimal forest rotation literature. *Journal of Forest Economics 8*(1), 5–27.

Oltenacu, P. A. and D. M. Broom (2010). The impact of genetic selection for increased milk yield on the welfare of dairy cows. *Animal welfare 19*(1), 39–49.

Perrin, R. K. (1972, February). Asset Replacement Principles. *American Journal of Agricultural Economics 54*(1), 60–67.

Pinedo, P. J., A. Daniels, J. Shumaker, and A. De Vries (2014, May). Dynamics of culling for Jersey, Holstein, and Jersey x Holstein crossbred cows in large multibreed dairy herds. *Journal of Dairy Science 97*(5), 2886–2895.

Pryce, J. E. and R. F. Veerkamp (2001). The Incorporation of Fertility Indices in Genetic Improvement Programmes. *BSAS occasional publication*, 237–250.

Pryce, J. E., R. Woolaston, D. P. Berry, E. Wall, M. Winters, R. Butler, and M. Shaffer (2014). World Trends in Dairy Cow Fertility. In *Proceedings, 10th World Congress of Genetics Applied to Livestock Production*, Volume 680.

Ray, D., T. Halbach, and D. Armstrong (1992). Season and lactation number effects on milk production and reproduction of dairy cattle in arizona. *Journal of dairy science 75*(11), 2976–2983.

Rogers, G. W., J. A. M. Van Arendonk, and B. T. McDaniel (1988a, December). Influence of Involuntary Culling on Optimum Culling Rates and Annualized Net Revenue. *Journal of Dairy Science 71*(12), 3463–3469.

Rogers, G. W., J. A. M. Van Arendonk, and B. T. McDaniel (1988b, December). Influence of Production and Prices on Optimum Culling Rates and Annualized Net Revenue. *Journal of Dairy Science 71*(12), 3453–3462.

Rothwell, G. and J. Rust (1997, April). On the Optimal Lifetime of Nuclear Powe Plants. *Journal of Business & Economic Statistics 15*(2), 195–208.

Rust, J. (1987). Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. *Econometrica 55*(5), 999–1033.

Schiraldi, P. (2011, June). Automobile Replacement: a Dynamic Structural Approach. *The RAND Journal of Economics 42*(2), 266–291.

Scott, P. T. (2013). Dynamic discrete choice estimation of agricultural land use. *Toulouse School of Economics Working Paper 526.*

Small, K. A. and H. S. Rosen (1981). Applied welfare economics with discrete choice models. *Econometrica (pre-1986) 49*(1), 105.

Stewart, H. M., E. B. Burnside, J. W. Wilton, and W. C. Pfeiffer (1977). A Dynamic Programming Approach to Culling Decisions in Commercial Dairy Herds. *Journal of Dairy Science 60*(4), 602–617.

Stott, A. W. (1994, January). The Economic Advantage of Longevity in the Dairy Cow. *Journal of Agricultural Economics 45*(1), 113–122.

Thomsen, P. T. and H. Houe (2006). Dairy cow mortality. A review. *Veterinary Quarterly 28*(4), 122–129.

USDA, N. Prices Received for Milk Cows - Dollars per Head.

Van Arendonk, J. (1985). A model to estimate the performance, revenues and costs of dairy cows under different production and price situations. *Agricultural systems 16*(3), 157–189.

Van Arendonk, J. A. M. (1988, April). Management Guides for Insemination and Replacement Decisions. *Journal of Dairy Science 71*(4), 1050–1057.

Weigel, K. A., R. W. Palmer, and D. Z. Caraviello (2003, April). Investigation of Factors Affecting Voluntary and Involuntary Culling in Expanding Dairy Herds in Wisconsin using Survival Analysis. *Journal of Dairy Science 86*(4), 1482–1486.

Appendix A Future Value Calculation

$$v(x_t, 1) - v(x_t, 0) = R(x_t, 1) - R(x_t, 0) + \delta \sum_{x_{t+1}=1}^{X} \Big(R(x_{t+1}, 1) + ln(P_1(x_{t+1}), 1) \Big) \tag{10}$$

$$\Big(f(x_{t+1}|x_t, 1) - f(x_{t+1}|x_t, 0) \Big) \tag{11}$$

The last term multiplied by δ is what I call ΔV, and I derive it below. Remembering that all the states evolve independently of one another and only a_t and η_t depend on i_t, we can factor the probabilities out this way:

$$f(x_{t+1}|x_t, 1) - f(x_{t+1}|x_t, 0) = \Big(f(a_{t+1}|a_t, 1) f(\eta_{t+1}|\eta_t, 1) - f(a_{t+1}|a_t, 0) f(\eta_{t+1}|\eta_t, 0) \Big) f(z_{t+1}|z_t)$$

When considering a_t, recall that a_t is a discrete state that can only transition to $a_{t+1} = 1$ or $a_{t+1} = a_t + 1$; the age must either go up by one or go back to 1, so it sufficient to only consider the cases where $a_{t+1} = a_t + 1$ or $a_{t+1} = 1$ when calculating the expected value. Because of unplanned exit, the probability of transitioning from age a_t back to age 1 is:

$$f(a_{t+1} = 1 | a_t, i_t) = \begin{cases} 1 & i_t = 1 \\ 1 - S(a_t) & i_t = 0 \end{cases}$$

And that the probability of a_t going to age $a_t + 1$ is:

$$f(a_{t+1} = a_t + 1 | a_t, i_t) = \begin{cases} 0 & i_t = 1 \\ S(a_t) & i_t = 0 \end{cases}$$

The shock state η_t is also dependent on the decision to replace. Recall that shocks are autocorrelated with coefficient ρ but only in the case that the cow is not replaced; should the cow be replaced, the performance of the previous cycle does not affect the new occupant.

Now we calculate the difference in transition probabilities for $a_{t+1} = a_t + 1$ and $a_{t+1} = 1$.

$$\begin{aligned} f(1, \eta_{t+1} | a_t, \eta_t, 1) - f(1, \eta_{t+1} | a_t, \eta_t, 0) &= (1) f(\eta_{t+1} | \eta_t, 1) - (1 - S(a_t)) f(\eta_{t+1} | \eta_t, 0) \\ &= f(\eta_{t+1} | \eta_t, 1) - (1 - S(a_t)) f(\eta_{t+1} | \eta_t, 0) \end{aligned}$$

$$\begin{aligned} f(a_t + 1, \eta_{t+1} | a_t, \eta_t, 1) - f(a_t + 1, \eta_{t+1} | a_t, \eta_t, 0) &= (0) f(\eta_{t+1} | \eta_t, 1) - S(a_t) f(\eta_{t+1} | \eta_t, 0)) \\ &= -S(a_t) f(\eta_{t+1} | \eta_t, 0)) \end{aligned}$$

Now back to caluclation of ΔV, first ignoring the states p_t and c_t since they are not influenced by the decision:

$$
\begin{aligned}
\Delta V =& \sum_{x_{t+1}=1}^{X} \Big(R(x_{t+1},1) + ln(P_1(x_{t+1})\Big)\Big(f(x_{t+1}|x_t,1) - f(x_{t+1}|x_t,0)\Big) \\
=& \sum_{\eta_{t+1}=1}^{E} \Big(R(a_{t+1}=1,\eta_{t+1},z_{t+1},1) + ln(P_1(a_{t+1}=1,\eta_{t+1},z_{t+1})\Big) \\
& \qquad \Big(f(\eta_{t+1}|\eta_t,1) - (1-S(a_t))f(\eta_{t+1}|\eta_t,0)\Big) - \\
& \sum_{\eta_{t+1}=1}^{E} \Big(R(a_{t+1}=a_t+1,\eta_{t+1},z_{t+1},1) + ln(P_1(a_{t+1}=a_t+1,\eta_{t+1},z_{t+1})\Big)\Big(S(a_t)f(\eta_{t+1}|\eta_t,0))\Big) \\
=& -S(a_t) \sum_{\eta_{t+1}=1}^{E} \Big(R(a_{t+1}=a_t+1,\eta_{t+1},z_{t+1},1) + ln(P_1(a_{t+1}=a_t+1,\eta_{t+1},z_{t+1})\Big)\Big(f(\eta_{t+1}|\eta_t,0)\Big) \\
& + \sum_{\eta_{t+1}=1}^{E} \Big(R(a_{t+1}=1,\eta_{t+1},z_{t+1},1) + ln(P_1(a_{t+1}=1,\eta_{t+1},z_{t+1})\Big)\Big(f(\eta_{t+1}|\eta_t,1)\Big) \\
& - \Big(1-S(a_t)\Big) \sum_{\eta_{t+1}=1}^{E} \Big(R(a_{t+1}=1,\eta_{t+1},z_{t+1},1) + ln(P_1(a_{t+1}=1,\eta_{t+1},z_{t+1})\Big)\Big(f(\eta_{t+1}|\eta_t,0)\Big) \\
=& S(a_t) \sum_{\eta_{t+1}=1}^{E} \Big(ln(P_1(a_{t+1}=1,\eta_{t+1},z_{t+1}) - ln(P_1(a_{t+1}=a_t+1,\eta_{t+1},z_{t+1})\Big)\Big(f(\eta_{t+1}|\eta_t,1)\Big) \\
& + \sum_{\eta_{t+1}=1}^{E} \Big(R(a_{t+1}=1,\eta_{t+1},z_{t+1},1) + ln(P_1(a_{t+1}=1,\eta_{t+1},z_{t+1})\Big)\Big(f(\eta_{t+1}|\eta_t,1) - f(\eta_{t+1}|\eta_t,0)\Big)
\end{aligned}
$$

So now apply the normalization that $R(x_t, 1) = 0$, and using the shorthand $P_1(a_t, \eta_t, p_t, c_t) = P_1(a_t, \tilde{x}_t)$ we can write:

$$
\begin{aligned}
\Delta V =& S(a_t) \sum_{z_{t+1}=1}^{Z} \sum_{\eta_{t+1}=1}^{E} \Big(lnP_1(1,\eta_{t+1},z_{t+1}) - lnP_1(a_t+1,\eta_{t+1},z_{t+1})\Big)\Big(f_\eta(\eta_{t+1}|\eta_t,1)\Big)f_z(z_{t+1}|z_t) \\
& + \sum_{z_{t+1}=1}^{Z} \sum_{\eta_{t+1}=1}^{E} \Big(ln(P_1(1,\tilde{x}_{t+1})\Big)\Big(f(\eta_{t+1}|\eta_t,1) - f_\eta(\eta_{t+1}|\eta_t,0)\Big)f_z(z_{t+1}|z_t)
\end{aligned}
$$

So now we have factored out the survival function $S(a_t)$ so we only can estimate its parameters inside the main model. The other state transitions, however, still have to be estimated separately. Note that the value FV_1 has to do with the fact that shocks are correlated, since when $\rho = 0$ then $f(\eta_{t+1}|\eta_t, 1) = f(\eta_{t+1}|\eta_t, 0)$ and $FV_1 = 0$, whereas FV_2 is an adjustment term for the change in the probability of replacing next period if replacement is done today.

Appendix B Milk Production Model

One of the covariates in our model is $\eta_{jt}p_t$, which is the shock in revenue from the current cycle. To get an estimate of η, which is the deviation from the production function, I do a linear prediction of fat and protein yield for each cow using their covariates. The covariates W_{jkt} come from similar models estimated in animal science production models on DHI data (see Kearney et al. (2004) as an example).

The prediction model:

$$y_{jkt} = \beta W_{jkt} + h_k + \eta_{jkt}$$

Contained in W_{it}:

- Lactation number
- Lactation number squared
- Proportion Days Milked 3x
- Lactation Length (DIM)
- Calving Month
- Birth Year
- Age at first calving

and h_k is a herd intercept. I then predict the residual $\hat{\eta}_{jkt} = y_{jkt} - \hat{\beta} W_{jkt} - \hat{h}_k$ for fat and protein and multiply them their Class III component prices prevailing in the month the record was taken.

Table 6 shows the results of the milk production model. Calculated from this production function, the optimal lactation number at which production is maximized is around three to four, which is in line with Miranda and Schnitkey (1995). This indicates that another reason dairy cows are replaced earlier than typically calculated by simulations is that production is maximized much sooner than five lactations. The birth year effects show something akin to genetic progress in milk production; independent of all factors, cows that were born in more recent years have higher milk production.

Table 6: Milk Production Model

	Fat Yield	Protein Yield	Energy Corrected Milk (ECM)
Lactation Number	66.852***	69.778***	2110.466***
	(0.631)	(0.444)	(15.373)
Lactation Number Sqrd	-10.973***	-9.686***	-321.745***
	(0.057)	(0.040)	(1.401)
Proportion Milked 3x	99.577***	76.034***	2860.178***
	(1.456)	(1.024)	(35.489)
Lactation Length	3.051***	2.573***	86.508***
	(0.001)	(0.001)	(0.033)
Age in Years	53.974***	33.396***	1358.848***
	(0.461)	(0.324)	(11.239)
Somatic Cell Score	-0.121***	-0.060***	-3.009***
	(0.001)	(0.001)	(0.026)
Birth Year			
2006	34.405***	30.421***	945.791***
	(0.984)	(0.692)	(23.993)
2007	65.036***	57.995***	1754.510***
	(0.954)	(0.671)	(23.255)
2008	75.858***	67.684***	2037.064***
	(0.969)	(0.681)	(23.612)
2009	89.866***	78.058***	2363.118***
	(0.999)	(0.703)	(24.356)
2010	109.260***	91.072***	2807.903***
	(1.037)	(0.730)	(25.284)
2011	110.927***	94.334***	2862.303***
	(1.101)	(0.775)	(26.851)
2012	118.899***	101.299***	3101.160***
	(1.238)	(0.871)	(30.183)
2013	101.899***	90.072***	2510.211***
	(3.035)	(2.135)	(73.983)
Observations	1,172,293	1,172,293	1,172,293
Adjusted R^2	0.86	0.90	0.89