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

Dairy farmers in the United States routinely replace their dairy cattle before they reach their maximum, productive potential. According to the dairy science literature, dairy cows reach maximum output at five years in the herd, yet they are typically removed at around three years. 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. Specifically, I analyze whether the costs of "unplanned mortality" cause dairy farms to replace at younger ages. 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 earlier than suggested by most dynamic simulations. Using the conditional choice probability method, the model results show that perceived mortality cost affects the replacement decisions of small dairies but not of large dairies. Instead, large dairies replace early because of differences in their production technology relative to small dairies.

Draft date: August 10, 2019

1 Introduction

Dairy farms in the United States routinely cull animals before the dairy science literature says they maximize their productive potential. A common finding in experimental herds is that a cow's annual milk output is quadratic in age, and reaches a maximum at about the fifth, year-long lactation (Mellado et al., 2011; Ray et al., 1992). Empirically, it is quite common for animals to be culled years before that, usually at lactation two or three (Hadley et al., 2006). Several dynamic programming simulations aimed at estimating the optimal culling rule suggest much lower rates of culling than actually practiced, yet 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 when they could be keeping animals longer and realizing higher production. What economic rationale could there be, if any, for such a pattern of asset replacement?

The recent trend in dairy cow health can potentially explain this replacement pattern. Over the past decades, milk production of dairy cattle worldwide has increased at about 3-4% annually, 50% of which is attributed to genetic selection alone (Pryce and Veerkamp, 2001; Thornton, 2010). Breeding for high milk production, however, has been met with declines in health and fertility. In the UK, rates of pregnancy on the first breeding attempt have fallen from 55.6% in 1975 to 39.7% in 1998, while 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 toll on the US dairy industry is evident in herd records, as death, infertility, injury, and disease together account for nearly 80% of cow exit (Pinedo et al., 2010; Smith et al., 2000). Compared to dairy cows born in 1960, a dairy cow's life span in the herd has decreased by 20% (decrease from 36 months to about 30 months) (De Vries, 2013).

The above trends suggest a rationale for the replacement policies seen in data. First, dairy farmers are incentivized to replace their old cows with younger ones because the subsequent generation produces more milk; newer dairy cows are "challengers" that increase the opportunity costs of keeping older cows (Perrin, 1972). At the same time, each generation is actually weaker than the previous generation, so they also develop health problems earlier. If unplanned mortality is costly to dairy farms, then they would also replace their animals at earlier ages as a precautionary measure to avoid the health problems that the animals develop at later ages. Both of these incentives create a cycle where dairy farmers replace earlier and earlier to both chase genetic gains while at the same time to buffer against the costs of their animals failing health.

This particular cycle is worrying for two reasons. First, the cycle of decreasing animal lifespan decreases both environmental quality and animal welfare, two issues of increasing importance to the general public. Two studies find that farms that replace more agressively cause more environmental damage, and that longer living, fertile cows are more beneficial to the environment (Bell et al., 2011; Garnsworthy, 2004). Decreasing cow health also negatively impacts animal welfare, which is of growing importance to consumers (Langford and Stott, 2012; Oltenacu and Broom, 2010).

The second reason, and the focus of this paper, is that replacing early to achieve genetic gains may not always benefit farmers in every market environment: De Vries (2017) argues that the gains in genetic progress accomplished by these breeding strategies do not outweigh the costs of declining health. Even with an agressive replacement strategy, higher rates of cows being forced to exit because of poor health still make dairy farms lose significant amounts of revenue. Since the gains from more milk production depend on the market price for milk, a drop in milk price will have a large impact on the dairy industry; the gains from high milk production in this scenario may not outweigh the costs incurred from declining health. Unfortunately, the costs borne by dairy farmers by having higher livestock mortality are harder to quantify than production traits. The economic return from increasing fat or protein production in dairy cows is reflected in the market prices for those milk components, but costs from unplanned mortality have no such market equivalent. Some recent attempts by

Heikkilä et al. (2012) and De Vries (2013) put forward estimates calculated from simulations of dynamic programming models, in the range of 500-1000 USD. These are good benchmarks, but these models simulate the costs rather than directly estimating the costs from data, which does not tell us as much about the actual costs dairy farmers face. A better alternative would be to use actual replacement decisions to estimate the costs that dairy farmers face.

I investigate whether the costs of unplanned mortality explain high replacement rates on Wisconsin dairy farms using a dynamic, discrete choice model. I define "unplanned mortality" as when a dairy cow dies or leaves the dairy herd when the dairy farmer did not plan to replace the animal, which includes cows dying on the farm as well as when a cow fails to conceive and is salvaged ahead of schedule. I contrast this with "planned mortality," which is when dairy cows are slaughtered as a planned decision. The theoretical model focuses on explaining replacement, a specific case of planned mortality, where the expected profit from the current cow is compared with the expected profit of a newer, replacement. I estimate this implied, ex-ante cost of unplanned mortality as well as the implied age at which production is maximized from the manager's perspective from data on dairy cow replacement decisions on over one-thousand farms, which represent about three-hundred thousand dairy cows.

I implement an estimator for dynamic discrete choice models pioneered by Hotz and Miller (1993) and Arcidiacono and Miller (2011), the conditional choice probability (CCP) estimator. The CCP estimator estimates the parameters of the structural model by approximating the continuation value of the dynamic programming problem; specifically, the continuation value is a function of the empirical probabilities of replacement from the data. Assuming type 1 extreme value errors, I estimate the structural parameters from a logit model predicting the decision to replace.

I estimate the structural parameters of the theoretical model using a novel, animal level dataset of over 300,000 dairy cows on over 1,000 Wisconsin dairy farms. I find that perceived costs of unplanned mortality explains replacement on small dairies but not on large dairies. The model estimates that small dairies perceive mortality costs to be as high as 9,000 USD per death, whereas large dairy replacement behavior is not affected by unplanned mortality. Instead, large dairies have differences in production technology that drive high replacement rates; specifically, large dairies believe that their dairy animals maximize production sooner than the dairy science literature predicts. The results of this paper show that unplanned mortality costs disproportionately affect the replacement behavior of small dairies; this implies that declining dairy cow health is more costly to small dairies than to large dairies.

2 Literature Review

Attempting to estimate the "optimal" replacement policy for livestock, specifically 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. Prices were assumed completely constant and the penalty for "unplanned mortality," referred to as an "involuntary cull" in the dairy science literature, was to gain no milk production in the next period. In this formulation, the authors found that economic factors such as milk price and replacement price had very little influence on the optimal replacement policy. Since prices were parameters and not states in the model, this was done by estimating the model several times at different prices. This conclusion is generally reflected in how these models would end up being estimated in subsequent animal science literature: the most emphasis has been on biological factors 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 delivered much of the culling advice that has become standard in the dairy industry, though their advice tends to not be followed in practice. For example, it is commonplace when advising farmers to claim that 20-30% of the herd should be culled each year, though the culling rate is higher than 30% (De Vries, 2013; Hadley et al., 2006). In fact, most dairy cows are culled at the second or third cycle, even though they do not reach their highest production until about five (Mellado et al., 2011; Ray et al., 1992). These discrepancies have seldom been investigated since dynamic models of animal replacement estimate their parameters outside of the dynamic model and do not allow any unobserved states or dynamic prices to affect their behavior.

Allowing for such uncertainty, however, has been the de-facto approach of the economics literature when studying asset replacement ever since Rust (1987). Subsequent studies on replacement of "durable assets" built on this approach while steadily making the estimation method itself faster and more efficient, including Rothwell and Rust (1997), Cho (2011), and Schiraldi (2011). The power of this approach compared to computing "optimal policies" of replacement lies in the fact that it estimates parameters from data that will *rationalize* the observed decisions using the economic model. By doing this, the researcher can uncover patterns from the data that help explain the rationale for the replacement decisions not a priori known. Concerning dairy cow replacement, Miranda and Schnitkey (1995) takes this exact approach and finds that a large component of the gain from replacement is unobserved to their model; the alternative specific constant for the replacement decision, which represented the location parameter of the distribution of the unobserved state, 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.

I theorize that the large and significant gain to replacement can be partly explained by the ex-ante expected losses that dairy farms experience from unplanned mortality. The animals studied by Miranda and Schnitkey (1995) are not the same animals that the US dairy industry deals with today; though they produce more milk, they are even more prone to health problems that force their removal from the herd. Currently, about 80% of culling is reported as "involuntary," that is due to reasons unrelated to prodution (Pinedo et al., 2014). This is broadly reflective of the recent goals of breeders in the US industry, which have prioritized production over health (De Vries, 2017).

These trends represent real problems for dairy farms, and several papers have attempted to raise the alarm by pointing out the economic costs of declining longevity and cow health. 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 as 500-1000 USD per exit in the US, even when not considering lost production. Finally, a similar strand of literature has quantified the environmental costs of these policies, since high replacement rates require large amounts of replacements to be kept on the farm. Garnsworthy (2004) finds that less fertile dairy herds could lead to higher emissions. Bell et al. (2011) also suggests increasing herd longevity as a way to reduce greenhouse gas emissions from dairy farms since replacement herds would be smaller.

All the above cost estimates are not generated from actual data, rather only simulated. The costs actually faced by producers is much harder to determine. My paper takes a different approach to the previous literature to determine how costly declining health is to the dairy industry. Rather than relying on simulations, I construct a dynamic model where the cost of unplanned mortality is a parameter than can be estimated from data on replacement decisions. This paper joins others such as Pinto and Nelson (2009) and Scott (2013) in utilizing recent advances in dynamic discrete choice estimation via Hotz and Miller (1993) and Aguirregabiria and Magesan (2013) to estimate dynamic models of agriculture decision making on data. In the next section, I describe my methodological framework for analyzing dairy cow replacement to recover the costs of mortality from data.

3 Theory Model

Consider the case where a dairy farm manager makes an annual decision at time t about a dairy cow that produces annual milk output $y(a_t)$ which is a function of its age a_t (measured in year-long "lactations"). 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, so the herd size is fixed.*

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 were 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 individual stall, so we can focus on profit maximization at the cow level instead of the herd level.[1]

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. Specificall, $i_t \in \{0, 1\}$, 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. When the animal is replaced, a new animal with $a_t = 1$ produces next year and gives revenue $p_t y(1)$ and costs the manager c_t.[2] If the current animal is replaced, expected revenue for the next year will be $p_t y(1) - c_t$.

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 either die 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

[1] For a herd level model of animal replacement, see Chavas and Klemme (1986).

[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, there is no one price that 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 seasons every lactation. For now, I assume adaptive expectations, meaning the manager values next period's revenue at the most recent, observed prices. See Section 5 for more details on the prices.

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 the carcass, the costs of treating a sick animal that ultimately dies, lost production, or the costs of finding a replacement if one is not already lined up.

I model these costs from unplanned mortality as a "penalty term" α, which is added to the cost of replacement when the replacement was brought on by a death.[3] In the case of this forced replacement, 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:

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

where x_t and z_t age "endogenous states," states such as a_t that are affected by decisions, and "exogenous states," which are not (such as market states p_t and c_t).[4]

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 $R(x_t, i_t = 1) - R(x_t, i_t = 0)$, 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 dies, then α has no effect on the decision to replace. Managers would replace when $y(1) - y(a_t + 1) > c_t/p_t$, that is when the marginal return from resetting the age is more than the replacement cost; we know at some point the left hand side will be negative if y is a concave function.

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, cows that are older are more likely to have to be removed), or at least decreasing after some point. As age progresses, α will get larger and the production and replacement cost difference will get smaller. Thus the higher cost of mortality will cause higher rates of replacement, even if the animal is young.

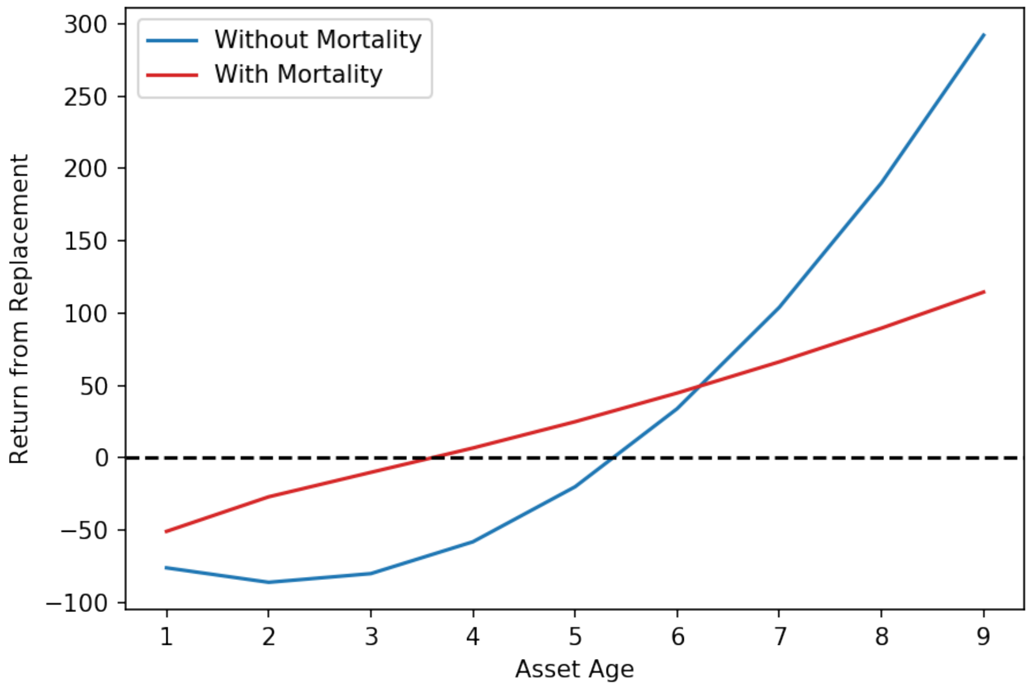
As an illustration, consider the parametric example in Figure 1 using the functional forms

[3] Note that this is unconditional of production; there are good arguments for making the penalty term actually 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 4).

[4] Saying that the prices p_t and c_t are exogenous is equivalent to saying that dairy farmers are price takers. This is generally true for dairy farms, especially dairy farms in Wisconsin that are usually less than 200 cows.

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. In essence, they are modeling something close to the costs of declining health. Linearity is quite restrictive, however, since going from age one to two must cost the same as going from age three to four. This may be why the maintenance cost they find, around 10-40 USD, was not statistically different than zero. In the above specification, age has another role, which is through S, which is typically non-linear. The main distinguishing feature between simple maintenance cost and mortality cost is not the functional form, however. It is the fact that as the important of α grows non-linearly, the effect of output price p_t and c_t decreases. 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. I include a linear maintenance cost function $M(a_t) = \gamma a_t$ to provide a way to compare these specifications.

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 could be a deviation from a group average, for example. 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. 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 a full, 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 time subscripts are replaced with prime notation to indicate a steady state solution. 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 in order 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) argue that in this case ϵ satisfies conditional indepence 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. 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.

As detailed in Table 1, 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 evolution of a_t will always go back to 1 if $i_t = 1$, but otherwise it will return to 1 with probability $1 - S(a_t)$ and go 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 $\bar{V}_0(x_t, z_t) = E(V(x_{t+1}, z_{t+1})|x_t, z_t, i_t = 0)$.

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

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

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

In this case, I include the effect of 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 reprecussions 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:

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

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 represents unobserved benefits to replacement not related to the rest of the state variables. Now we know the 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).

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.

[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.

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).

Note that this expression indefintely progresses by substitution of $v(x_{t+1}, z_{t+1}, k)$, but in the next section I discuss how choosing $k = 1$ makes the difference in payoffs finite. 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.

Application to Asset Replacement The most helpful choice of k is to assume that all cows are replaced next period and exploit the principle of "limited dependence," as in Aguirregabiria and Magesan (2013) and Arcidiacono and Miller (2011). 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, $f(x_{t+1}|x_t, 1) - f(x_{t+1}|x_t, 0)$. Remembering that $S(a_t) = f(a_t + 1|a_t, 0)$ and $1 - S(a_t) = f(1|a_t, 0)$, we can factor out the continuation value ΔV to get the following:

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

$$
\begin{aligned}
FV_1 &= \sum_{z_{t+1}=1}^{Z} \sum_{\eta_{t+1}=1}^{E} \Big(lnP_1(1, \eta_{t+1}, z_{t+1})\Big)\Big(f_\eta(\eta_{t+1}|\eta_t, 1) - f_\eta(\eta_{t+1}|\eta_t, 0)\Big) f_z(z_{t+1}|z_t) \\
FV_2 &= \sum_{z_{t+1}=1}^{Z} \sum_{\eta_{t+1}=1}^{E} \Big(lnP_1(1, \eta_{t+1}, z_{t+1}) - lnP_1(a_t + 1, \eta_{t+1}, z_{t+1})\Big)\Big(f(\eta_{t+1}|\eta_t, 0)\Big) f_z(z_{t+1}|z_t)
\end{aligned}
$$

[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 $P_1 = 1$.

(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 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 (it will overfit).

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. Rather than having the econometrician chooses the bins, the algorithm will choose how the bins are made using cross-validation. 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 simply increasing the number of covariates in a logit model.

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. The technique is used 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. In order to have a baseline curvature, I use the probability of "involuntary culling" from Van Arendonk (1985) while allowing for different farms to have different overall amounts of mortality risk.

4.3 Second Stage Estimation

Logit Method With the above derivation, we can estimate the equation by 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_tc_t.

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

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

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

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 $y(1) - y(a+1) = 0$:

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

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.

Endogeneity Issues A final note on identification. A correct estimate of α depends on there being no correlation with some unobserved variable that affects the manager's perception of the survival rate. This can be a concern, because management actions can certainly have an effect on survival rate (for example, a more intensive milking system as Thomsen and Houe (2006) finds). The survival rate is also usually affected by large scale decisions, for instance decisions on housing and bedding, which cannot be so easily changed in the short run. I make the assumption that from the managers perspective, the survival rate is fixed in its curvature. I assume that the only heterogeneity in the perception of survival rate across farms is linear shifts in the curve at the farm level.

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; heterogeneity in perception of survival rates manifests through linear shifts in the survival rate rather than changes in curvature.*

The importance of the above assumption is that we can robust check the model results by including farm specific fixed effects to take care of any heterogeneity in the expected survival rate.[7]

A bigger issue, one that I cannot currently address, is the presence of certain hidden "health" states at the cow level that are signals of the probability that the animal survives but also affect exit. The precise story being told with this model is that managers spot health problems and which would affect the probability of survival and subsequently replace to solve this problem. The presence of an unobserved, cow-level characteristic would affect both the perception of the survival rate and the probability of replacement, making a clear endogeneity problem. There are, however, some promising "proxies" of health in the data that could be worked into the model in the future to proxy for this "health" state.

[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 more than 600,000 cows, bringing a total of 1.2 million lactation records. I look specifically at lactation level records, which record the total fat and protein at the end of the lactation rather than the production for that test day. The raw data contained many more herds, 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. Considering 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. It would appear that large farms try and keep cows at younger ages rather than have older ones in the herd.

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

[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	1,552
Cows	521,328
Cows with Full History	333,247
Lactation Records	866,010

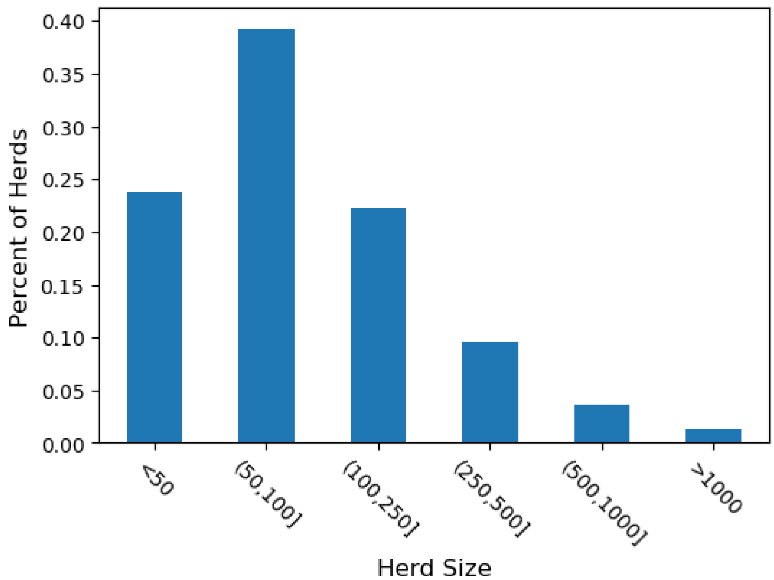


Figure 2: Record Counts and Herd Sizes

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 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* 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 rate at which cows are bred, but 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.

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

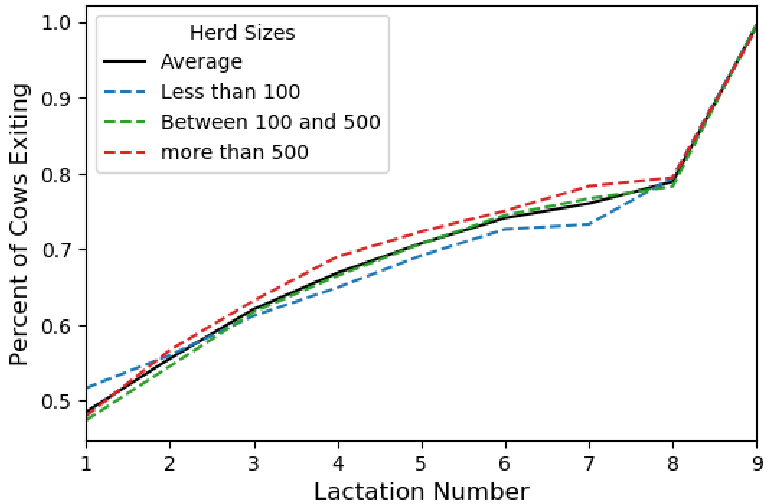


Figure 3: Percent Exiting at Each Lactation
(Uncensored Cows Only)

farms may have to be more prudent in their decisions. This is fairly consistent with the pregnancy rates at exit: larger farms have more pregnant dairy cows exiting. This is likely a mix of death and planned replacement, but unfortunately here it is indistinguishable.

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? Importantly for the model, how do 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 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, it is not exactly clear whether a cull was "forced" or is simply because

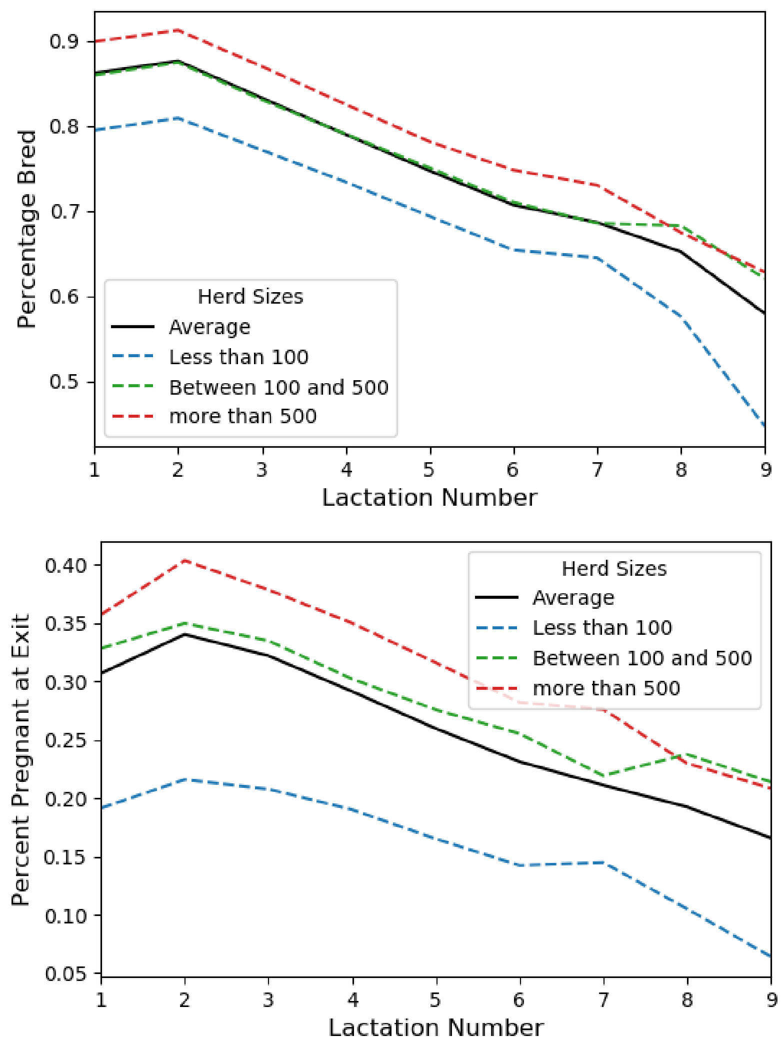


Figure 4: Breeding Attempted and Pregnancy Rate (Uncensored Cows Only)

the disease affects the productivity of the cow.

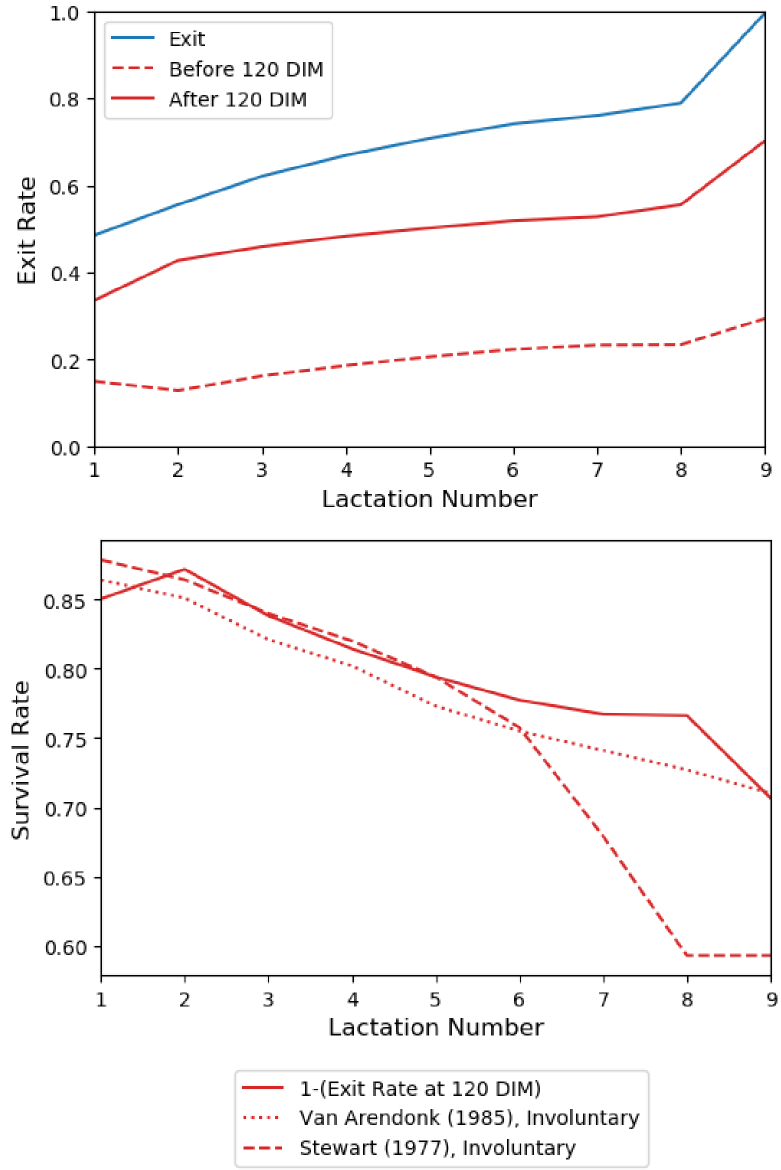
As little information is available about why the cow exited (or the economic motive behind it), I analyze replacement using the exit rate as the dependent variable (i_t).The argument behind using just the exit rate hinges on the extent to which unplanned mortality is "unplanned"; if truly unancticipated, there is no reason that the state variables predict the decision. However, there may not be a reason they bias the parameters either. The only damage thus done is to underestimate the costs we are interested in.[9] This strategy particularly relies on anything pushing the exit rate up to be both farm 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 independent of time. If this is the case, then herd 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.

5.2 Survival Rate

Finally, we also need an estimate of the survival rate, a difficult task since we do not know which exits are deaths. 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). After this decomposition, the exit rate is a bit lower. 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 face risks that older cows do not. 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. Despite this, in the first six lactations they roughly how the same rate of decrease after lactation two. In later lactations, the literature appears more optimistic about survival. Since these rates were calculated some time ago, it may be that technology has allowed dairy cows that do survive to that age to continue surviving; note for example how the surival trends upward slightly at lactation seven.

[9] This is not necessarily the case for unobserved health states; see the discussion in the conclusion.

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



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, and the replacement cost is calculated as the salvage value of a 1400 pound dairy cow minus the market price for a new heifer.[10] 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 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.

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 and fixing characteristics, including 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 $p_t\eta_t$, I calculate the residual for both fat and protein, multiply them by that months Class III component prices, and sum them together.

Figure 6 shows the calculated rewvenue shock over age. As can be seen in the confidence intervals, they are highly variable; they could not reasonably be considered different than zero. See Appendix B for more details about the calculation of $p_t\eta_t$.

6 Results

Below I estimate the structural model using logit and conditional logit (to attempt to control for fixed herd characteristics). Standard errors in the logit estimates 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

[10] Income-over-feed-cost is a common measure of dairy profit margin due to certain beliefs about the production process. There is a general belief that there is no intensive margin, that is within lactation production decisions, to make that affect production y. This views the lactation curve as completely deterministic, an assumption that also allowed us to write production function y only as a function of age.

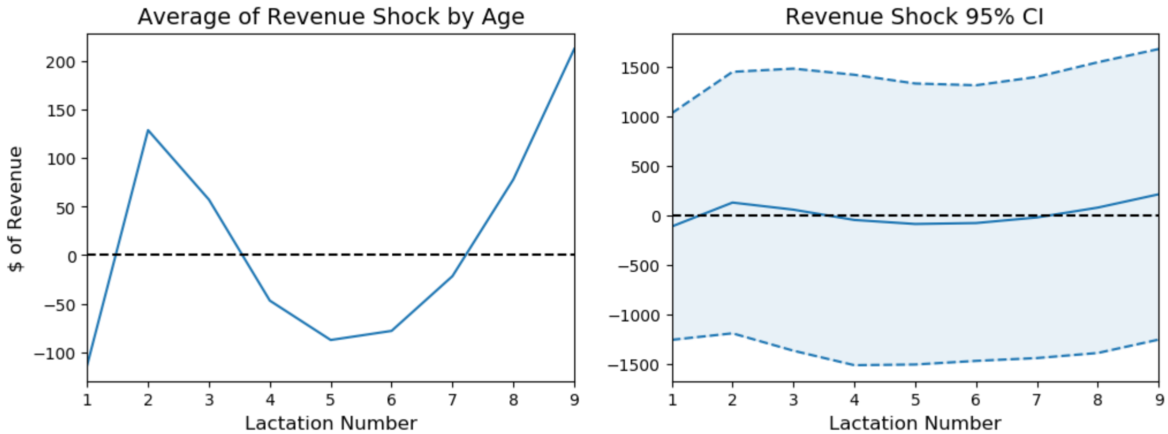


Figure 6: Revenue Shock by Age

particular combination of states is not often seen in the data. In all models the discount rate is fixed at .99 unless otherwise specified.

6.1 First Stage Estimation

The first stage estimation of FV proceeds in two steps. First, the transition probabilities f_x and f_z are used in calculating the expection over the CCP P_1. Second, the CCP P_1 is calculated on every combination of states after being trained on the sample. A random forest algorithm is used to prevent overfitting and assure good out-of-sample properties of the estimator.

State Transitions An AR(1) normal regression was estimated for each of the three states in order to take expectations over the out of sample probabilities. I used the following equations, assuming that the error term is normally distributed:

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

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

The prediction was done using monthly data that was deseasonalized and CPI adjusted.[11] 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.

[11]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.

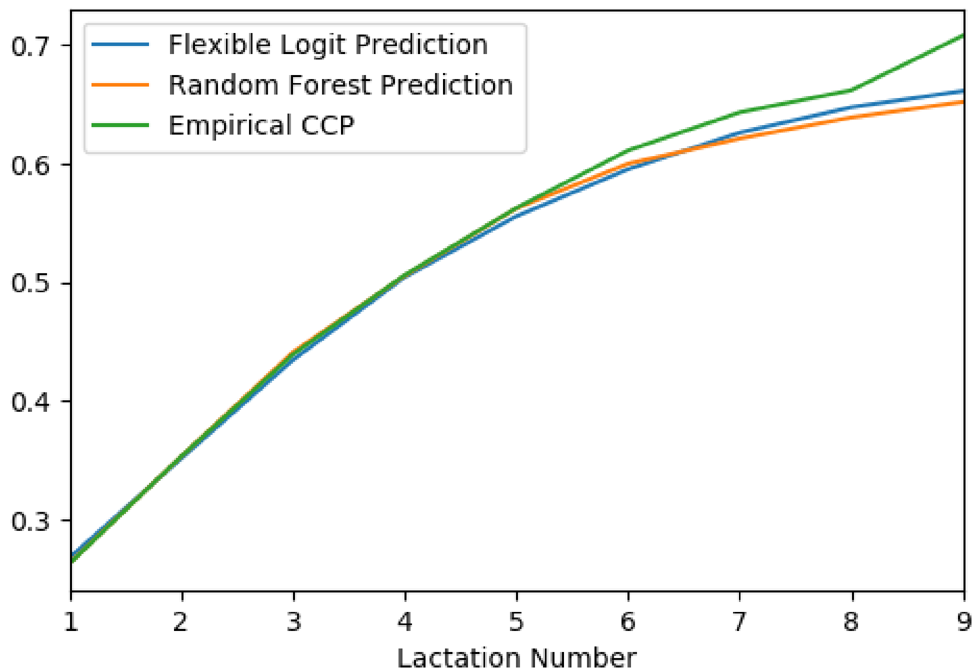


Figure 7: CCP In-Sample Predictions

Conditional Choice Probabilities The random forest algorithm was trained using brier score loss and 10-fold cross validation. Figure 7 shows the performance of the random forest estimator in-sample as compared to flexible logit, a popular first stage method for CCP estimators. Both estimators perform very similarly, with only slight differences in performance at higher ages (which represent less than 1% of the data). While in-sample performance was similar to flexible logit, I ultimately chose random forest to protect at overfitting.[12]

6.2 Structural Logit Model Estimates

In the following two tables, I estimate the structural model with conventional logit and conditional logit (which condition on the herd) on the whole sample and also on different herd size categories. In the preliminary results that follow, I estimate the model on exit decisions and analyze heterogeneity across herd type.

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. Unlike in their model, the constant term μ, which they call the "premium," is negative. The marginal cost paramater γ is also

[12]Data appendix detailing the in and out of sample performance of these methods in this particular problem is forthcoming.

ten times the size of their estimate. In this data, it appears that high rates of exit are being well explained by age, as the logit did not have to compensate with a large, positive constant term. Shock correlation ρ is in the direction as expected, but is quite small. The implied production function parameters suggest that the age at which production is maximized is about four lactations, which is only one year off from the dairy science literature estimate of five. While we would expect the time-trend to be positive, indicating the increasing benefit from adopting new cows, in this specification it is quite small; it implies that every month the option to replace loses 7 USD. 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.

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 signficant. Once it is divided by the marginal utility of income, the scale λ, it is quite large; these results imply the cost of an unplanned exit is about 3,700 USD, which is more than twice the typical annual profit of a dairy cow (about 1,600 USD). While marginal cost stays about the same, the age of maximum production goes down slightly. In this specification, the replacement rates are mostly being driven by a high perceived cost of mortality and a lower production maximum than normally expected in dairy cows.

The last column estimates the same model using conditional maximum likelihood to parse out herd-specific factors to exit that are time-invariant. The coefficient α becomes quite small and is statistically not distinguishable from zero. The marginal cost estimate shrinks, and the age at which production is maximized goes to 4.5, much closer to 5 than before. This is an interesting result, as Miranda and Schnitkey (1995) did not have a large enough dataset look at more than about five herds, and could not look at how the model performs on a variety of farms. Unconditioned, the model suggest a role for the function of age that we have used here. After conditioning on the herd, this specification does not fit as well. It would appear that everything attributed to the hazard rate is actually attributed to time-invariant, herd characteristics.

In the descriptive statistics, there was significant heterogeneity in exit and breeding rates across herd size. In Table 4, I estimate the same conditional logit model as Table 3 but on different categories of herd size. The first column shows that on farms with less than one-hundred cows the model predicts a mortality cost. As before, it is quite large, on the order of 9,000 USD. In contrast, the linear age term is actually negative, predicting an increase in 100 dollars in profit from each year that it increases. The shock correlation is also the opposite direction than expected, implying that cows with higher than expected production are more likely to be culled (though the estimate is quite small, as to be almost insignificant to the decision). The age of max production on such farms is as the literature says, around five.

For larger herds, the hazard rate does not have the expected sign, but rather the opposite. Instead, the linear age term has the most explanatory power, being much larger than in Table 3. The age at which production is maximized is also lower than on smaller dairies.

These results show that the model derived above is in general not suited to handle a large level of herd heterogeneity. While the specification estimated a positive α on small farms, the linear age specification does the best job of representing health costs on large farms. Conditioning out all herd effects, the "marginal cost" term γ does the best job of representing health costs across the whole sample, as the hazard rate had no explanatory power in the conditional logit. The model did uncover, however, heterogeneous expectations of the production function; larger dairies expect their dairy cows to maximize production at younger ages. Small farms, in contrast, have expectations in line with the dairy science literature. This could explain why exit rates on large dairies are higher for older cows, whereas it is the opposite on smaller dairies.

7 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. I found that, after conditioning out herd fixed effects, this hypothesis was not supported on the whole sample. Instead, small dairies, less than one-hundred cows, were the only farms whose replacement decisions were motivated by the survival rate of dairy cows. For larger dairies, more intensive production appeared to be the reason for early replacement: larger farms expected their dairy cows to maximize production sooner than predicted by the dairy science literature. From this study, there are a few points of discussion important for the economics literature as well as the dairy industry: the role of heterogeneity in culling models and the disproportionate effect of declining health on small dairies.

First, heterogeneity across herds is substantial, and one structural model of animal replacement is not likely to fit all production systems. For example, this simple model had very different interpretations depending on what sort of herd it was applied to. For small herds, it appeared that unplanned mortality was quite important to replacement, whereas this hardly seemed to factor into large farms' replacement decisions. Perhaps part of the failing of dynamic simulations to understand culling is that it attempts to impose the same structure to multiple farms who actually have quite distinctly different production systems, and maybe even objectives. Under these circumstances, it may then seem quite natural for many dairy farms to ignore culling advice from such simulations. One this study has done is highlight the significant heterogeneity in the replacement decision across farms, which necessitates different modeling approaches for different sizes of dairy farm.

A final point of discussion, and one important for policy, is that if declining health matters to the bottom line of the dairy farm, it matters most to small farms. In the analysis, smaller farms were much more motivated to replace dairy cattle because of survival rates. As an implied cost, the number from this study, about 9,000, is much too large to be a realistic estimate of mortality cost, but pinpoints that declining health will affect the replacement decisions of smaller farms the most. This may be because larger farms have sufficient scale to deal with animal death cost-effectively, and thus it does not factor into the replacement decision as much. For these farms, they appear to factor in an annual, marginal cost to progressing to each lactation rather than replacing to avoid a mortality cost. Since health appears to matter to small farms the most, it is likely that declines in cow health will hit these farms the hardest. It should be factored into policy, then, that declining animal health will have distributional effects in the industry.

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.

Bell, M. J., E. Wall, G. Russell, G. Simm, and A. W. Stott (2011). The effect of improving cow productivity, fertility, and longevity on the global warming potential of dairy systems. *Journal of Dairy Science 94*(7), 3662–3678.

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

Chavas, J.-P. and R. M. Klemme (1986). Aggregate Milk Supply Response and Investment Behavior on U.S. Dairy Farms. *American Journal of Agricultural Economics 68*(1), 55–66.

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.

De Vries, A. (2017, May). Economic Trade-Offs Between Genetic Improvement and Longevity in Dairy Cattle. *Journal of Dairy Science 100*(5), 4184–4192.

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.

Langford, F. M. and A. W. Stott (2012). Culled early or culled late: economic decisions and risks to welfare in dairy cows. *Animal welfare 21*(1), 41–55.

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.

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.

Pinedo, P. J., A. De Vries, and D. W. Webb (2010, May). Dynamics of Culling Risk with Disposal Codes Reported by Dairy Herd Improvement Dairy Herds. *Journal of Dairy Science 93*(5), 2250–2261.

Pinto, A. D. and G. C. Nelson (2009, June). Land Use Change with Spatially Explicit Data: A Dynamic Approach. *Environmental and Resource Economics 43*(2), 209–229.

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*.

Smith, J. W., L. O. Ely, and A. M. Chapa (2000, December). Effect of Region, Herd Size, and Milk Production on Reasons Cows Leave the Herd. *Journal of Dairy Science 83*(12), 2980–2987.

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.

Thornton, P. K. (2010, September). Livestock Production: Recent Trends, Future Prospects. *Philosophical Transactions of the Royal Society B: Biological Sciences 365*(1554), 2853–2867.

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

$$v(x_t,1) - v(x_t,0) = R(x_t,1) - R(x_t,0) + \delta \sum_{x_{t+1}=1}^{X} \Big(R(x_{t+1},1) + ln(P_1(x_{t+1}),1)\Big) \tag{8}$$

$$\Big(f(x_{t+1}|x_t,1) - f(x_{t+1}|x_t,0)\Big) \tag{9}$$

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 previous 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)
- Breed
- 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. Thus we get an estimate of the shock in revenue, ηp.

Table 1: Model Summary

Endogenous	a_t	Age
States (x_t)	η_t	Production shock
Exogenous	p_t	Output price
States (z_t)	c_t	Replacement cost
Controls	$i_t \in \{0, 1\}$	Replacement decision
Technology	$y(a_t) = \beta_0 + \beta_1 a_t + \beta_2 a_t^2 + \eta_t$	Total milk output
	$\eta_t \sim N(0, \sigma_y^2)$	of lactation a_t.
	$M(a_t) = \gamma a_t$	Maintenance cost
		function.
	$S(a_t)$	Survival rate
	$p_{t+1} = \tau_{0p} + \tau_{1p} p_t + \xi_t^p$	AR(1) process of prices
	$\xi_t^p \sim N(0, \sigma_p^2)$	
	$c_{t+1} = \tau_{0c} + \tau_{1c} c_t + \xi_t^c$	
	$\xi_t^c \sim N(0, \sigma_c^2)$	
	$P(a_{t+1} = 1 \vert i_t) = \begin{cases} 1 & i_t = 1 \\ 1 - S(a_t) & i_t = 0 \end{cases}$	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, $R(x_t, z_t)$	$p_t y(1) - c_t$	If $i_t = 1$
	$S(a_t)\Big(p_t y(a_t + 1)\Big) +$	If $i_t = 0$
	$(1 - S(a_t))\Big(p_t y(1) - c_t - \alpha\Big)$	
Parameters	β_0, β_1, β_2, γ	Production and maintenance cost parameters
	$\tau_{0p}, \tau_{1p}, \sigma_p^2, \tau_{0c}, \tau_{1c}, \sigma_c^2$	Price parameters
	$\delta \in [0, 1)$	Discount factor
	α	Cost of mortality
	μ,λ	Location and scale of error term

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

Table 3: Structural Model Estimates

		No Mortality	Mortality Risk	Conditional Logit
Premium	μ	-1852.2	-1754.93	
		(124.691)	(110.753)	
Time Trend	τ	-7.721	-7.757	-4.647
		(0.335)	(0.306)	(0.194)
Penalty	α		3709.71	154.081
			(402.736)	(225.948)
MC	γ	257.099	230.747	144.453
		(17.459)	(17.484)	(9.009)
Shock Correlation	ρ	0.035	0.029	0.025
		(0.002)	(0.002)	(0.001)
Age of Max	$-\frac{\beta_1}{2\beta_2}$	3.92	3.741	4.548
		(0.045)	(0.04)	(0.048)
Age of Free	$-\frac{\beta_1+\beta_2}{\beta_1}$	6.84	6.481	8.096
		(0.091)	(0.08)	(0.096)
Observations		866,010	866,010	866,010

Bootstrapped “standard errors” in parentheses

Discount rate set to .99

Table 4: Structural Model on Different Herd Sizes

		Less than 100	100 to 250	250 to 500	500 to 1000	More than 1000
Time Trend	τ	-3.532	-5.487	-5.034	-1.331	-4.622
		(0.698)	(0.374)	(0.483)	(0.446)	(0.555)
Penalty	α	9048.46	143.785	-14.865	-2026.36	-5175.62
		(1230.041)	(415.221)	(558.975)	(427.911)	(768.354)
MC	γ	-91.975	133.422	195.352	163.843	358.464
		(21.529)	(16.383)	(24.62)	(16.468)	(38.504)
Shock Correlation	ρ	-0.025	0.012	0.044	0.032	0.051
		(0.003)	(0.001)	(0.004)	(0.002)	(0.004)
Age of Max	$-\frac{\beta_1}{2\beta_2}$	5.286	4.717	4.253	4.219	3.955
		(0.196)	(0.098)	(0.085)	(0.102)	(0.081)
Age of Free	$-\frac{\beta_1+\beta_2}{\beta_1}$	9.571	8.433	7.506	7.438	6.909
		(0.391)	(0.196)	(0.17)	(0.205)	(0.163)
Observations		193,921	204,682	201,011	142,449	123,947

Bootstrapped "standard errors" in parentheses

Discount rate set to .99