Milked for All They Are Worth: Livestock Replacement in a Dynamic Discrete Choice Model

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

This is a preliminary draft.
Please do not cite or distribute without permission of the author.
Contact author for latest draft or click here.
Draft date: September 15, 2019

Abstract

This paper examines animal replacement behavior for over 1,000 Wisconsin dairy farms during the period 2011-2014 and analyzes the rationale for high replacement rates. Dairy farmers in the United States replace their cattle at around three years in the herd, earlier than what asset replacement models predict is profit maximizing. 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 paired with machine learning, I estimate the cost of mortality at 1,800 USD per death, 800 dollars more than estimates based on simulation studies. Utilizing farm size heterogeneity, I also find that mortality costs are three times higher on small dairies than on larger ones. In a counterfactual estimation, dairy farmers were willing to pay 1,300 USD to eliminate mortality risk completely for first year dairy cows. These results suggest that genetic selection in U.S. dairy favors relatively large farms and may be accelerating the exit of small farms.

Keywords: Dynamic Discrete Choice, Dairy, Asset Replacement, Investment Behavior
JEL Codes: C61, D25, Q12, Q16

1 Introduction

Dairy farms in the United States routinely cull animals before the economics and dairy science literatures claim is optimal. Asset replacement simulations which take into account prices and the management environment suggest keeping cows longer to maximize profit; despite this, dairy farmers have maintained consistently high rates of turnover in their herds by replacing earlier than considered optimal (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 cow mortality, equivalent to an asset unexpectedly becoming inoperable, account for high replacement rates on Wisconsin dairies. To test this hypothesis, I develop an asset replacement model in which the risk of unplanned mortality, an event where an animal dies or is removed from the herd unexpectedly, drives farmers to replace their animals earlier than without mortality. I use more than 1,000 Wisconsin dairy farms with over 260,000 cows to estimate a dynamic, discrete choice model which backs out the perceived cost of unplanned mortality. Using the conditional choice probability method paired with machine learning, I calculate the cost of unplanned mortality as 1,800 USD per cull, 80% higher than the largest estimates from dynamic programming simulations. Utilizing farm-size heterogeneity in the dataset, I also find that unplanned mortality costs are almost three times higher on small dairies than on larger dairies. Finally, using the structural model, I use compensating variation to calculate the average willingness-to-pay to eliminate mortality completely. On average, farmers are willing to pay about 90 USD ex-ante to eliminate unplanned mortality completely for first-year cows, implying an average, expected indemnity of 1,300 USD per death. My results suggest that discrepancies between the literature and reality are because models underestimate the cost of unplanned mortality, and that these costs are highest on small farms.

The default approach of the dairy science literature to understand economic behavior in animal replacement has been constructing dynamic programming simulations that prescribe rather than describe the optimal culling rule for dairy farms (Arendonk and Dijkhuizen, 1985; Stewart et al., 1977). This approach, however, has failed to explain discrepancies between the optimal culling rule calculated from a model and the replacement policies actually practiced. To understand this

discrepancy, I instead take the positive approach to the problem, following the principle of Rust (1987). By assuming dairy farmers are acting optimally, my behavioral model backs out parameters that actually describe the costs that farmers face. This exact approach is used by Miranda and Schnitkey (1995) to study dairy cow replacement, but their own model could not rationalize the high rates of culling found on dairy farms; their results found a large "premium" to culling dairy cattle that was not explained by the states in their model.

Their model, however, does not include an important component of the asset replacement decision on dairy farms: animal mortality. In fact, 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 in the past three decades, so has the incidence of infertility and disease (Pryce and Veerkamp, 2001). In the United Kingdom, 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). Decreasing dairy cow lifespan is also of concern because it has negative externalities affecting environmental quality and animal welfare, two issues of increasing importance to the general public (Garnsworthy, 2004; Oltenacu and Broom, 2010).

As a result, the dairy industry is subject to a technology that has high returns early in its life but is prone to failure later in its life. This technology potentially accelerates replacement rates in two important ways. First, dairy farms may replace earlier to acquire higher producing cows if milk production continues to increase. Holding an asset (a "defender") when its replacement (a "challenger") is more productive is a significant opportunity cost for the manager, causing earlier replacement as technological progress continues (De Vries, 2017; Perrin, 1972).

A second reason dairy farmers might replace early is to avoid the costs that are incurred when the asset fails in the middle of its production cycle. Asset failure, in this case unplanned mortality, is potentially very costly to dairies: cows either dying or having to be removed ahead of schedule will cause lost production, high treatment costs, and lost salvage value, resulting in a high penalty

for asset failure. If unplanned mortality costs are high, dairy managers may replace early 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). In fact, the reason dairy farms replace "too early" may be because we drastically *underestimate* the cost of unplanned mortality on dairy farms. These motives are mentioned but not explored in Miranda and Schnitkey (1995), meaning there is little understanding of how important these factors are in the replacement decision nor how costly this feature of the technology is for producers.

There has been no research using empirical modeling to estimate these costs from data. Since the costs of mortality are difficult to observe, Heikkilä et al. (2012) and De Vries (2013) instead 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 actually face. In addition to assuming knowledge all of the parameters of the decision, they cannot incorporate unobserved cow heterogeneity that may bias the decision. An empirical model, in contrast, does not assume knowledge of structural parameters: it instead calculates them directly from the data by assuming the observed behavior is an output of an economic model. An empirical model can also control for unobserved heterogeneity in animals that cannot be observed in data, but which may bias estimates of culling rules. Providing an empirical estimation of mortality cost to guide policy using a structural model is the main policy contribution of this paper.

In addition to informing policy, I also make a contribution to the literature on asset replacement by flexibly incorporating permanent, asset heterogeneity into a dynamic discrete choice model and using machine learning to facilitate estimation. Dynamic discrete choice models suffer from being computationally burdensome and difficult to estimate, especially using the nested fixed point algorithm of Rust (1987). To facilitate computation, 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. In my structural model, the Euler condition generates estimates of the cost of unplanned mortality, annual maintenance costs, the parameters of the production function, and the effect of a time trend on replacement, which proxies for technological progress. I also implement the random forest algorithm to facilitate the first stage probability estimation; this approach has the advantage of efficiently choosing bins for continuous variables and theoretically improving model efficiency by balancing the in-sample and out-of-sample properties of the probability prediction.

This paper calculates how costly unplanned mortality is on Wisconsin dairies using empirical replacement decisions while considering a range of motives, including technological progress. I find that the costs of unplanned mortality explain early replacement and are variable across herds. My estimate of 1,800 USD is 80% higher than the upper bound of De Vries (2013) and is the first empirical verification of these costs. In general, I do not find technological progress in the three year period in my data to be a significant factor explaining high rates of replacement; in fact, as time progressed, dairy farmers were more likely to keep cows rather than replace them. The model estimates that small dairies perceive unplanned mortality costs to be as high as 2,400 USD per death, which is nearly three times higher than the cost of unplanned mortality on larger dairies. Using welfare analysis, I calculate that dairy farmers would be willing to pay 90 USD to insure their first lactation cows against unplanned mortality, implying they perceive the unplanned mortality "indemnity" at about 1,300 USD.

The policy implications of this study are that unplanned mortality costs are a key determinant of replacement patterns and disproportionately affect small dairies. My results find that unplanned mortality costs are a much more significant predictor of replacement than time trends which proxy for technological progress in the market for dairy genetics. If U.S. dairy breeders continue to select for high milk production, this will likely decrease animal health, which disproportionately hurts the profitability of small farms. The results of this paper suggest an unintended consequence of dairy genetics which have high production but short lives: it may accelerate the exit of small farms, further contributing the consolidation of the dairy industry seen in recent years (Shepel, 2019).

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, that of 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 with previous methods.

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 model the environment. 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), whose paper in the *Journal of Dairy Science* explicitly modeled and solved 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 25 USD per lactation per year as a lower bound and about 100 USD as an upper bound. Heikkilä et al. (2012) calculates the cost of mastitis due to early exit as around 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 500-1,000 USD per exit in the United States 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, 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 is Miranda and Schnitkey (1995), who find that a large component of the gain from replacement was unexplained by their model. They hypothesized that annual costs that were linear in animal age are responsible for early replacement, but for all farms in their study this parameter is statistically insignificant. Instead, the alternative specific constant for the replacement decision, which is the location parameter of the distribution of the unobserved state ϵ, is large and significant compared to other factors in the model. They theorize that this constant represents 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 does not incorporate any mortality risk, the costs of mortality can instead manifest instead in the alternative specific constant.

Finally, this paper also contributes more generally to the literature analyzing dynamic decisions using dynamic discrete choice and the Conditional Choice Probability method. 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. It is more difficult with logit models to incorporate this heterogeneity and, in the case of conditional logit, impossible to estimate asset specific fixed effects (instead they must be conditioned out). Hotz and Miller (1993) show that continuation values in dynamic models can be approximated by empirical choice probabilities, which significantly opens up the number of ways to estimate dynamic models. Another recent advancement has been the Euler equation conditional choice probability (ECCP) method, which uses discrete analogues to Euler equation methods 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 than logit. By using the CCP and ECCP methods, I can more easily control for this time-invariant component of heterogeneity and estimate asset specific fixed effects. In addition to using this method, I implement the random forest algorithm in the first stage of the model to improve the efficiency of the estimator.

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 I recover cost and production function parameters to investigate the causes of replacement from data.

3 Theory Model

[Figure 1 about here.]

Consider the case where a dairy farm manager makes an annual decision at time t about a dairy cow at herd age a_t. Throughout the paper, the state a_t refers to the number of producing years the cow has spent in the herd, measured in year long "lactations," as opposed to number of years old (see Figure 1). The cow has an annual production function $y(a_t)$, which is the yield for that entire production year. 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. Given that dairy farms often have very large fixed costs and are

[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.

frequently in debt, it is not common to leave a stall empty; this would imply that profit margins are so low from milk production that it is more profitable to take a guaranteed loss. This is not often the case, and, in fact, dairy farms in many cases will respond to drops in price by expanding production (Atwood and Andersen, 1984). The power of this assumption for analysis is that I only need to focus on a binary decision. 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. A manager does 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]

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 100 days of their cycle, when they are weakest, meaning if an 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 the manager 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

[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 the 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.

replacement when the replacement is 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 be removed), or at least decreasing after some point. As age progresses, α will get larger and 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 2 using a quadratic, concave functional form for $y(a_t)$. The survival probability is modeled as a variant of the Weibull hazard rate function and is monotonically decreasing as age increases. The payoffs with and without asset

[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 (which we return to in Section 6).

failure are graphed in red and blue.

[Figure 2 about here.]

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 $\alpha = 0$. Notably, the payoff with mortality also has a smaller slope with respect to age; this is due to the effect of $S(a_t)$ on the other states, which decrease in importance as age increases. Because of this difference in curvature, even small increases in α will cause a large discrepancy between these culling rules.

3.2 Other Causes of Early Replacement

Above I demonstrate how the 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) claim 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 $S(a_t)$. The role of $S(a_t)$ 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 2, this is why the slope of the red line is different than the slope of the blue line. Two, $S(a_t)$ 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 replacement 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. Bethuyne (1998); Perrin (1972), but is not the focus of this paper. I do incorporate the fact that the payoff will change over time and that the option to replace in the year 2011 does not have the same value as that same option 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. The significance of this factor can provide evidence of whether there are technological progress expectations in this decision.

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 that represents the influence of states not observed in the data. The following are the two assumptions about the unobserved state ϵ to estimate the parameters in a regression model:

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.*

Assumption 2 is common to models of this type (e.g. (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 is 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, and only over the states x_t and z_t. Assumption 3 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 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 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 account, we can rewrite the payoff function:

[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 keep more than 200 cows, so market power is very dispersed.

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

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

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 is termed the "culling premium" by Miranda and Schnitkey (1995), and contains benefits to choosing to replace unexplained by the model. Now, we know the difference in current period payoffs but must next take into account the continuation value, which is the difference between

[5] To make the shock transmit only 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.

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 about here.]

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 $\Delta V(x, z) = E(V(x', z')|x, z, 1) - E(V(x', z')|x, z, 0)$. Rust's (1987) the solution to this problem is to solve the value function iteration problem to find $V^*(x, z)$ across all states, 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 CCP estimator derived by Hotz and Miller (1993) and expanded on by 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 "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" of Hotz and Miller

(1993)). Now, using the function ψ we can substitute $\bar{V}$ into 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) show 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, to exploit the principle of "limited dependence." Limited dependence is a special feature of models that involves 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 at $t+2$ that is unaffected by decisions in t (see the example of 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 $t+1$.

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

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

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

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

(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.

[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 $P_1 = 1$.

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 $\hat{P}$, both in-sample and out-of-sample. Unfortunately, we need to observe all combinations of ages, production shocks, and prices to have accurate estimates of P_1 across all states. 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 model 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 over-fitting).

I choose to predict P_1 using a random forest algorithm as a compromise between these two methods for the following two reasons. First, a random forest model 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. It is essentially a more sophisticated bin estimator that frees the econometrician from having to choose bins for continuous variables. When a random forest model 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 by Arcidiacono and Miller (2011), is a classic example of a model that will over-fit; it will produce accurate in-sample probabilities but will do poorly at predicting combinations of ages, 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 the production of the animal is 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 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; see, for example, 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 from the manager's perspective that this is actionable. The management actions that have an effect on cow death and infertility are broad structural changes that 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 lactation, which is most often an unplanned exit.

4.3 Second-Stage Estimation

With the above derivation, we could 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.

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

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, after we divide through by λ, are

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

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

So we can recover the structural parameters:

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

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 that of 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 parameters 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 the cow's 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 that manager to replace animals differently than

expected by the literature. 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 $y(1) - y(a+1) = 0$:

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

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 animals' production function is the same as what the literature says.

ECCP Method The issue with the above method is that it cannot flexibly incorporate permanent, unobserved asset heterogeneity which may bias our estimates. The typical way to control for such heterogeneity in a logit, for example, would be conditional maximum likelihood estimation. In such a method, however, fixed effects are not calculated, simply conditioned out, and so cannot be used to do estimates of marginal effects. Similarly, counterfactual estimation cannot be done considering these fixed effects if conditional logit is used.

Instead of conditional logit, I use the ECCP method which can flexibly incorporate fixed effects in the estimation equation. This 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:

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

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. A non-myopic decision maker also weights 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 (especially fixed effects estimators). I derive a regression equation using the moment condition the following way:

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

which leads to the regression equation:

$$
\begin{aligned}
\tilde{Y}_{jt} =& \mu_j + \tau t + \alpha(1 - S(a_{jt})) - \rho\eta_{jt}p_t - S(a_{jt})c_t + \\
& \gamma S(a_{jt})a_{jt} - (\beta_1 + 2\beta_2)S(a_{jt})a_{jt}p_t - \beta_2 S(a_{jt})a_{jt}^2 p_t + \tilde{\xi}_{jt}
\end{aligned}
$$

Note that with this method we do not have to condition out fixed effects, and so can get estimates of μ_j for every animal in the dataset. For doing counterfactual estimation of probabilities, for example, we can use Equation 7 using our estimate of θ without assuming the fixed effect is zero.

Unobserved benefits to replacement that are time variant 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 the data.

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. There are two main sources of endogeneity concern: herd environment and unobserved health information.

Thomsen and Houe (2006) show that herd environment affects the survival rate; more intensive dairies, for example, generally have higher rates of exit. Since management intensity (e.g., 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 controlled for using fixed effects. 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 is 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 endogeneity. 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 technical literature explaining cow health means dairy farmers would likely converge roughly on a survival rate curve 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

[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.

effects. This changes the interpretation of the model, however, because it excludes from the analysis cows with only one observed lactation. In particular, it omits all cows that do not survive 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 an untreatable health problem; the above theoretical model does not fit either of these exits, as neither is initiated with a replacement in mind. In the analysis that follows, I default to cow-level fixed effects.

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 ΔV. For now, I use SCC to robust check the model results by including it as a variable in the regression. This is making the assumption that SCC affects next period's revenue but does not impact the value function. While not perfect, this helps check whether spikes in SCC present an endogeneity problem for the model and robust check the results. More information on the robustness check is available in the Appendix C.3.

5 Data

My sample is for Dairy Herd Improvement (DHI) herds in Wisconsin served by one Dairy Records Processing Center, which represents 90% of the DHI herds in the state (DHI herds represent 50% of total dairy farms in the state). DHI Associations are producer cooperatives that were established in the early 20th century U.S. to collect information on dairy cow production for on-farm benchmarking of different cows. Herds that are a member of DHI have records collected every month on their cows' milk production, including fat and protein yield, as well as somatic cell

count (SCC) and calving and breeding decisions. Since every cow currently in the milking herd has records collected every month, cows that exit the data are cows that have exited the milking herd.

My data covers the period June 2011 to January 2015, about 1,500 herds, and about 260,000 cows which have 640,000 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 contains 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 could not have wild fluctuations in herd size, in this case meaning I dropped herds whose herd size has a coefficient of variation of more than one.[8] Herds which have wildly fluctuating herd size are generally considered unreliable. 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 would cause 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.

[Figure 3 about here.]

5.1 Exit Rates

A dairy cow "exit" is when a dairy cow leaves the dataset. In this data, if a dairy cow leaves the dataset less than 6 months before the end of the sample time frame, the cow is considered right censored rather than an exit. Figure 4 shows the rate of exit for cows that are uncensored in the data, on average, and also by herd size.

[Figure 4 about here.]

Exit rates are very high for dairy cows in this sample; around 50% of dairy cows leave the data in their first lactation. Fewer than 30% of dairy cattle make it to their output-maximizing age of five.

[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.

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 dataset, we also know whether a dairy cow was bred. Breeding is an important decision to observe because it is, in many cases, a signal of intention for the cow to be kept in the herd.[9] Specifically, while breeding does not indicate the cow was *not* replaced, purposefully *not* breeding is equivalent to taking the cow out of production. Whether or not this cow is sold or slaughtered, it must be replaced in those herds that are maintaining a roughly constant herd size. Figure 4 shows the rate at which cows are bred.

While exit rates are almost homogenous across dairy farm size, the rates of breeding are quite heterogeneous across herd size in terms of point estimates. The standard deviation of the replacement rate, however, is high enough such that none of the above curves are statistically independent. 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 first lactation cows unbred, consistent with the fact that their exit rate in Figure 4 is higher at the first lactation. At all levels they breed fewer cows. Larger farms may have an advantage in the fact that they can afford to inseminate in more cases at their scale without concern for cost; smaller farms may have to be more prudent in their decisions. Otherwise this may be evidence that smaller farms replace earlier than large farms.

Unfortunately, herd testing data often do 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 that 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, etc.). 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

[9] 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.

on the farm is similarly not considered "culling," as the manager did not plan for 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 make 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 clearly unplanned cull and culling for low production is a clearly planned cull, there is a continuum of circumstances between these that compromise 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 sudden health event.

As little information is available about why the cow exited (or the economic motive behind the exit), I first analyze replacement using the exit rate as the dependent variable (i_t). I then use an alternative, much stricter definition as a robustness check. Using just the exit rate 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, these action do not confound estimation if the increase in death rate is a linear shift in the exit rate. If this is the case, then fixed effects will control for 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 which exits are made with economic rationale in mind.

It also has the potential to bias our estimates of α upward, however. Since the empirical probability P_1^e of exit is one part the CCP generated from our model P_1 and one part unplanned exit, our model could mechanically pick up a correlation that biases α upwards since we currently rely on $1 - S_t$, a measure of unplanned exit, for our identification. Given, it is not explicitly clear how this will bias our estimator, since our outcome $\tilde{Y}$ contains non-linear combinations of P_1^e. Also, if the amount of truly unplanned exit is small, this will not affect our results very much.

Since the first strategy may produce upward biased estimates of mortality cost, I use a secondary definition that is not subject to this measurement error to robust check the results of the first method. The second strategy relies on breeding decisions as a proxy for replacement. Specifically, it restricts replacement to be only defined as when a cow is intentionally *not* bred. Similar to how fallowing a field is equivalent to taking it out of production, deciding not to breed a cow is

equivalent to taking the cow out of production, as cows must become pregnant and give birth in order to start the next lactation. The breeding decision also must be made typically in the first four months of the lactation, leaving a small window for unplanned mortality to enter the decision. While *not* breeding is equivalent to removing a cow, the converse is not true: breeding a cow does not imply the cow is being kept in the herd. Cows that have been already bred may be sold for a variety of reasons, and so we do not necessarily consider cows kept unless they appear in subsequent lactations. Also, records that are censored only four months into the lactation may be missing a breeding decision only because the cow died before it could be bred. For this reason, we do not consider these "early" exits a true replacement either. In fact, these early exits are most likely deaths, as most deaths happen in the first four months according to Hadley et al. (2006) (which is why later it becomes our estimate of $1 - S_t$). See Figure 5 for an explanation of how I label replacements for this definition.

[Figure 5 about here.]

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 6 shows what happens when we decompose exit rates by being at either less than or more than 120 days in milk (DIM), which is the number of days into the lactation).

[Figure 6 about here.]

The exit rate before 120 DIM has a bathtub shape, which is to be expected for 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, the same trend; however, rates from the literature do not show 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.

[Figure 7 about here.]

5.3 Market Prices and Shocks

[Figure 8 about here.]

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 income-over-feed-cost (IOFC), a measurement of the profit margin from producing one pound of milk "at test."[10] This measure of milk profitability is important to producers because it is the one used for the price support program Dairy Producer Margin Protection Program and formerly the Milk Income Loss Contract Program (Gould and Cabrera, 2011). 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 de-seasonalized, 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 which 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 adaptive expectation only for the term R, which means that managers value next period's expected revenue at the most recent prices. To calculate FV, I assume the probability of next period prices to be derived from 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. Figure 8 shows the price trends before and after seasonal correction from our data.

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 mate, 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 lactation's latest Class III component prices, and sum them together. See Appendix B for more details about the calculation of $p\eta$.

[10] 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 the "average profitability" of producing one pound of milk. See Gould and Cabrera (2011) for details of calculation.

6 Results

Below I estimate the structural model using the ECCP method. Standard errors in all models are estimated as the standard deviation of 1,000 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 nonlinear 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. I present a robustness check of the results using my second definition of replacement and finish this section by estimating 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 over-fitting 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:

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

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

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

The prediction was done using monthly data that was de-seasonalized and CPI adjusted.[11]

[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

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 about here.]

Figure 9 shows the calculated revenue shock over age. As can be seen in the confidence intervals, they are highly variable; they could not reasonably be considered different than zero for any age. While across all ages the shock must mechanically be zero, given that it is a residual from a regression, it is not necessarily true across ages. Here, on average a given shock can reasonably be expected to be zero for any cow. The variation in this variable is high, reflected in the fact that the standard deviation is about 600 USD. The correlation between shocks is 0.3, implying only a mild autocorrelation between lactations.

[Figure 9 about here.]

Conditional Choice Probabilities

[Figure 10 about here.]

The random forest algorithm was trained using Brier score loss and 10-fold cross validation. Figure 10 shows the performance of the random forest estimator in-sample as compared with 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 (which comprise less than 1% of the data).

6.2 Second Stage Estimates

Table 3 presents the ECCP model estimates, which control for fixed, cow-level characteristics. Additionally, I estimate the model with a zero discount rate (a myopic decision maker). The model most similar to Miranda and Schnitkey (1995), which does not consider mortality, gives sensible results and parameters but has the lowest R-squared. My model estimates the penalty term α for

using the monthly pattern of prices. This implies that I am assuming dairy farmers look only 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 it be a direction of future research.

myopic and dynamic decision models, and I find the cost to be 840 USD for a myopic decision maker and 1,800 USD for a dynamic decision maker. As a benchmark, in the year 2011 the average market price for a dairy heifer replacement was on average 1,400 USD (USDA-NASS, 2011). It is also 800 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. This result is consistent with the idea that dairy farmers expect the cost of unplanned mortality to be a little above losing a dairy heifer replacement. Note also that our time trend is negative, which is not consistent with genetic progress being a motivating factor in animal replacement. Over time, managers are more likely to hold on to their current cows, in direct contradiction with the challenger versus defender model of asset replacement and technological progress. This result is robust to using an alternative time trend based on cow birth year, suggesting this is a robust feature of the decision and not a fluke of the data (see Appendix C.3 for more details). These results were also robust to including somatic cell count as a covariate, which had very little effect on the decision and no effect on the estimation of α (see Appendix C.3).

I noted before that this model calculated parameters of the production function implied by the behavioral model instead of from an empirical production function. Here I find a production function that is in line with the empirical production function, which I calculated in Appendix B; the age at which cows maximize production is about three, though it is actually lower in this model than what the empirical production function shows. This is one of the few, if any, studies that calculates production function parameters from both a behavioral model and an empirical production function. This allowed us to test an extra hypothesis about replacement behavior, namely that replacement is driven by manager's believing that dairy cows will maximize production sooner than five years. This is true in this behavioral model, meaning differing perceptions of the production function are another reason dairy cows are replaced early.

[Table 3 about here.]

Table 4 shows the same method but using instead the replacement definition based on breeding.

Recall that the probability of breeding a cow should not contain $1 - S_t$, and so should not upward bias the estimate of α. The estimates in this model, however, are larger than those just using the regular exit decision. The estimate of α is slightly higher here: about 2,300 USD. The age at which cows maximize production is also higher in this model, and is more in line with the empirical production function calculated in Appendix B. Note also that the production shock correlation ρ is closer to the first stage estimate of 0.3. The marginal cost here is also lower, around 120 USD versus 200 USD.

[Figure 11 about here.]

The ECCP model also estimates a distribution of cow-specific, fixed effects. These effects are individual "replacement premiums" that are unexplained by the model. They are presented in Figure 11 over age. In Miranda and Schnitkey (1995), they find a large "replacement premium," but in I instead find "keep premiums"; as cows get older and older, they are more likely to be kept. These unobserved traits, likely a factor of genetics that only the manager observes, appear to influence replacement behavior. This is some form of survival bias, but we can now see that the reasons for it are unobserved to this model. Note that these premiums are independent of observed performance, already captured in the state η. This shows how important it is for models analyzing asset replacement to control for this heterogeneity.

When comparing these two different model definitions on the same sample, I find that the effect on α is relatively small. The measurement error in the exit rate does not appear to be substantial as to bias our first definition too far upwards. The results of this comparison and the effects of different levels of fixed effects are show in Appendix C. Since the different definition does not change the estimate of α, the difference in α in Table 3 and Table 4 has to do with the observations we are dropping: these observations that are being dropped are specifically cows that exit already bred or within the first four months of their lactation. One reason the breeding sample has a higher mortality cost could be that these replacements are more prudent, forward looking decisions. These managers are making a conscious decision to not breed their dairy cows rather than selling them later in the lactation, and these exits appear to put more weight on mortality cost. Since we ideally want to use a range of dairy cattle, I proceed in all analysis with our first definition in order to have more heterogeneity in the data.

[Table 4 about here.]

6.3 Heterogeneity across Herds

The descriptive statistics show significant heterogeneity in exit and breeding rates across herd size. It is likely that different sized dairy farms 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.[12]

[Table 5 about here.]

In general, there is no significant heterogeneity in λ or ρ across farm size with the exception of dairy farms with more than 1,000 cows (which are the smallest category in the data). There appear to be slight differences in the age at which production is maximized, with bigger farms expecting cows to maximize slightly sooner than small farms. Smaller dairy farms appear to have cows that maximize production sooner, at 3.12 lactations for farms less than 100 cows versus 2.76 for farms with between 500 and 1000 cows.

The most heterogeneous estimates are the unplanned mortality cost α; farms with less than 250 dairy cows pay a mortality cost about three times higher than farms with between 500 and 1000 cows. Farms with between 250 and 500 dairy cows pay about 1,500 whereas the largest dairy farms, which make up a very small percentage of herds in this data, pay about 1,700 USD. Taking 1,400 USD as an average market rate for a dairy cow, only farms with 500 to 1000 cows pay less than market rate. 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 are cheaper than small farms costs simply 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. While the cause is unknown, these results suggest that small farms are disproportionately affected by reductions in dairy cow health.

[12] Appendix C presents these same estimates with the second definition of replacement.

7 Welfare Analysis

As a structural parameter, α is the added cost to replacement induced by an unplanned mortality event. As a counterfactual exercise, how much would farmers actually pay to eliminate mortality risk completely? Earlier, we saw how including $S(a_t)$ distinguishes this model from that of Miranda and Schnitkey (1995). The payoff with and without the probability of unplanned mortality is

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

Since $\theta^1 X^1_{jt}$ 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. This is a more comprehensive estimate of mortality cost: it takes into account the effects of S_t on the entire payoff function. This number would be equivalent to the premium of an insurance policy on animal mortality.

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 $\theta^0 X^0_{jt}$ to make them indifferent to transitioning to $\theta^1 X^1_{jt}$. Assuming that the value of λ is the same for every farmer, the average CV for transitioning to this payoff function is:

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

(see Small and Rosen (1981) for derivation).[13] To understand the relationship between mortality penalty α and $E(CV)$, the following partial derivative provides some intuition:

$$\frac{\partial E(CV)}{\partial \alpha} = (1 - S_t)P_0(x_t, z_t)$$

The intuition of this result is that the relationship between the penalty and the willingness to

[13] Recall also that I fixed the payoff to replacement to be zero having subtracted the payoff from replacement from both payoffs, so $-\theta^1 X^1_{jt}$ is the payoff from keeping the animal

eliminate mortality is the probability of the event weighted by the probability that cow would be kept. If the cow were definitely going to die, $S_t = 0$, and the manager were definitely going to keep the cow, $P_0 = 1$, then the relationship between α and $E(CV)$ is one-to-one. The less likely the manager is to keep the cow, the less they will pay to eliminate risk. This gives us a framework for thinking about how $E(CV)$ will be affected by the other states of the model; anything that increases the probability of keeping the animal will also increase $E(CV)$.

[Figure 12 about here.]

Since $E(CV)$ can be thought of as the premium for an insurance policy for animal mortality, dividing this by the probability of the event in some sense recovers an implied cost of mortality that is more comprehensive then the parameter α. It is, however, subject to the states in the model, unlike α which is a primitive parameter. For example, now the replacement premium μ_{jh} will also affect mortality cost. The bottom graph of Figure 12 shows how $E(CV)$ will be highest for cows with a low "replacement premium." This captures the ability of managers to see characteristics about cows that are unobserved in the model, giving us a sense for their importance in determining how managers value mortality.

[Figure 13 about here.]

In Figure 13 I graph both the CV over the whole sample as well as per age. On average, managers would insure their cows for 91.26 USD per lactation. The multi-modal distribution is a result of different distributions per age, as can be seen in the CV broken out by age. First lactation cows have the highest CV; this is intuitive, as managers would pay more to insure dairy cows that have more producing potential. Managers would pay around 100 USD to insure a new cow, implying a total mortality cot of about 1,300 USD. As cows age, the CV goes lower and lower, but becomes more variable. The increase in variability is due to the fact that older cows have the highest "keep premium" independent of the other states in the model (see Figure 11). In this sample, unobserved cow characteristics influence the willingness to pay to eliminate risk significantly. A subset of older cows in particular elicits a high $E(CV)$.

Broken out by herd size, the same pattern emerges as in Table 4. Smaller farms would pay the most to eliminate mortality, about three times more than larger farms.

[Table 6 about here.]

[Table 7 about here.]

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 incorporates the probability of unplanned mortality in the dairy herd. I also bolstered the model using the random forest algorithm, a technique that has not, to my knowledge, been used in estimating dynamic discrete choice. I also used the ECCP method to control for unobserved asset heterogeneity, finding that this heterogeneity was a significant factor in replacement. In contrary to Miranda and Schnitkey (1995), I found that most cows had a "keep" premium as opposed to a "cull" premium. Using these methods, we used the estimates of these fixed effects to calculate expected compensating variation that takes into account unobserved cow variation.

My model made significant progress in understanding more about the replacement rationale of dairy farmers by utilizing an expansive dataset and constructing an empirical model estimated using standard methods in dynamic discrete choice. This paper has, however, several shortcomings 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. Even when exits are explicitly labeled with disposal codes, which many studies use, the reasons given by dairy farmers may be very ad hoc and need not represent the actual economic motives behind the decision. My paper has made a vital contribution by 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 cow fixed effects, the cost of unplanned mortality was about 1,800 USD per death, 800 dollars higher than previous simulations have documented. In addition to mortality cost, I found other determinants of the replacement decision: managers replace cows earlier than five years because on average they expect dairy cows to maximize production at three years. The also pay a compounding cost of 200 dollars per lactation that discourages keeping cows, which confirms the hypothesis of Miranda and Schnitkey (1995). An alternative robustness check suggested that that the estimate is even higher when using a definition of replacing based on breeding. In a welfare analysis, I estimated that farmers would be willing to insure against mortality by paying a premium of, on average, 90 USD per year, and would pay about 1,300 USD, the average cost of a replacement heifer, to completely eliminate mortality for their fresh cows. Contrary to the finding of Miranda and Schnitkey (1995), replacement decisions were explained by a "keep premium," a motive to keep cows that was unobserved in the model. This paper was the first to document the large amount of heterogeneity in "keep premiums" on Wisconsin dairies and explicitly calculate their distribution from data. This provides evidence that animal-level genetic factors, not herd-level factors, are the main driving force in replacement decisions. This suggests that future studies have to find a way to control for this unobserved, asset specific heterogeneity, and points to the important of animal-level management on Wisconsin dairies.

In addition to being the first empirical verification of mortality cost from a behavioral model, this paper also found large disparities in mortality cost across herd size that have important implications for policy. Farms under 250 dairy cows paid roughly three times more per death than dairies between 500 and 1,000 cows. Small farms were also willing to pay three times more to eliminate mortality risk completely. The policy implications of my results are that small dairy farms may be disproportionately 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 United States if swings in price disproportionately cause small dairies to exit. My results imply that the price environment is not the only thing to blame for small dairies going out of business; it is also the types of cows that dairy farms are using. When breeding for high milk production, it should also be taken into account

that declines in dairy cow health will actually impact the distribution of dairy farm size; small farms disproportionately pay the cost for increased milk production, and will disproportionately be impacted when milk price swings down. The dairy sector in the U.S. has already been following a trend of increased consolidation, especially in states like Wisconsin where the number of farms is going down but the number of cows is staying constant (Shepel, 2019).

These results suggest there is an ample margin for increasing dairy cow life at the cost of less milk production. There are many reasons to breed dairy cows for longer life, including reducing emissions and increasing animal welfare, but my results show that, since mortality is costly, there may be ways to breed for longer life without harming the dairy farm's bottom line. In fact, breeding for longevity may improve the dairy farm's bottom line, especially for small farms. The different cost environments are important to consider when improving dairy cow genetics to assure the survival of many dairy farms in the years to come.

References

Aguirregabiria, V. and A. Magesan (2013). 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.

Arendonk, J. A. M. v. and A. A. Dijkhuizen (1985). Studies on the Replacement Policies in Dairy Cattle. III. Influence of Variation in Reproduction and Production. *Livestock Production Science 13*(10.1016/0301-6226(85)90025-9), 333–349.

Atwood, J. D. and J. C. Andersen (1984). A Backward Bending Supply Response from Some Dairy Producers? *Economic Research Institute Study paper 84*(6), 1.

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). 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. *Engineering Economist 43*(2), 85.

Boström, H. (2008). 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 Empirical 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). 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). 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). 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.

Gould, B. W. and V. E. Cabrera (2011). USDA's Livestock Gross Margin Insurance for Dairy: What is it and How Can it be Used for Risk Management. *New York State Department of Agriculture and Markets*, 74.

Hadley, G. L., C. A. Wolf, and S. B. Harsh (2006). 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). 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). 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). 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 X Environment Interaction for Grazing versus Confinement. I. Production Traits. *Journal of Dairy Science 87*(2), 501–509.

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). Evaluation of Factors Associated with Increased Dairy Cow Mortality on United States Dairy Operations. *Journal of Dairy Science 91*(4), 1423–1432.

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

Newman, D. H. (2002). 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). Asset Replacement Principles. *American Journal of Agricultural Economics 54*(1), 60–67.

Pinedo, P. J., A. Daniels, J. Shumaker, and A. De Vries (2014). 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.

Rogers, G. W., J. A. M. Van Arendonk, and B. T. McDaniel (1988a). 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). 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). On the Optimal Lifetime of Nuclear Power 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). 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*.

Shepel, J. (2019). Dairy Consolidation Continues but Federal Program May Slow Rate. *Wisconsin State Farmer*.

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

Thomsen, P. and H. Houe (2006). Dairy Cow Mortality: A Review. *Veterinary Quarterly 28*(4), 122–129.

USDA-NASS (2011). Prices Received for Milk Cows - Dollars per Head. `https://www.nass.usda.gov/Statistics_by_State/Washington/Publications/Historic_Data/index.php`.

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). 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). 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 calculation 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_{jt}, 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_{jt} come from similar models estimated in animal science production models on DHI data (see Kearney et al. (2004) as an example).

The prediction model:

$$y_{jt} = \beta W_{jt} + h_j + \eta_{jt}$$

Contained in W_{jt}:

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

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

Table 8 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 8 about here.]

Appendix C Robustness Checks

C.1 Alternative Definition

[Table 9 about here.]

Table 10 shows the results of the alternative definition. Disparities between small farms and large farms are similar to those of the first definition. Small farms have three times higher mortality cost than larger farms (about 3,700 USD versus 1,300 USD). As in Table 4, there is a higher production maximizing age on smaller farms and also a higher consideration of production shocks (higher ρ).

[Table 10 about here.]

C.2 Different Levels of Fixed Effects

Below I estimate the ECCP model with different levels of fixed effects. Without any fixed effects, the results are relatively similar to what we estimate with fixed effects. The mortality cost appears to be attenuated to zero as a result of confounding factors at the cow level; herd fixed effects do not change the estimates, whereas cow fixed effects result in a higher mortality cost. It appears that whatever confounding factors there are for individual cows, they serve to attenuate the effect of the hazard rate on the replacement definition.

[Table 11 about here.]

C.3 Other Checks

In this section I implement two extra robustness checks: including SCC as a covariate and using an alternative time trend based on the year different cattle were born.

As mentioned previously, unobserved health shocks are a potential endogeneity issue for estimating the coefficient on the hazard rate. Specifically, health shocks may affect replacement and also update the probability of survival. The effects of health states on productivity are already captured in the state η_{jt}, which uses somatic cell count (SCC), a measure of milk bacteria count, in the production function. SCC is an important trait of milk because high counts of SCC are indications of mastitis, the most prevalent disease among lactating dairy cattle. However, any effect of health states on replacement independent of production is not captured in the model. To test the effect of SCC in the model, I explicitly include it as a covariate. Note that this is making a very specific assumption about how SCC affects replacement; by including it only as as covariate and not a state, we are assuming it affects next period's payoff but not the continuation value. In Table 12, the coefficient on SCC in the regression is positive, meaning cows with higher bacteria count are replaced more often, but the coefficient is quite small. A one standard deviation change in SCC, which is 150, implies a change in expected profit of just three dollars. In addition, it does not significantly change the estimate of α.

Another robustness check I implement is using an alternative time trend. In the main specification, I used a simple monthly time trend to capture technological improvement in replacements. My hypothesis was that the time trend would positively related to replacement since replacing now means taking advantage of new genetics. This is implied by the challenger versus defender model of asset replacement, and is theorized in Miranda and Schnitkey (1995) as the cause of the positive culling premium. In all specifications, I find that the time trend is negative, the opposite of what was hypothesized. Rather than being more willing to give up their cattle as time progresses, producers are less willing. The monthly time trend, however, may be picking up another economic condition unrelated to the technology. To test whether this negative time trend really has to do with the technology itself, I use the birth year of the cow as an alternative measure. Instead of a one unit increase being one month, now a one unit increase is one year, for example a cow born in 2009 versus a cow born in 2010. The relationship is still limited to be linear, meaning the difference between a 2011 cow and a 2012 cow must be the same between a 2009 and 2010 cow, but is variation strictly related to the cow itself and not to another trend.

The effect of the time trend remains negative, and about the same magnitude as the month trend. This confirms that the negative time trend has to do with the technology itself. The milk

production model in Table 8 confirms that there are higher returns in milk production associated with increases in this variable. The trend does not fit with the theory that genetic progress is an increase in opportunity cost incentivizing dairy farmers to replace earlier (De Vries, 2017). It instead suggests that newer dairy cows are being kept *more*. It could be that dairy farmers actually intend on holding newer cows longer because of their increased productivity. If they believe newer cows to have higher productivity, however, the negative trend implies that they do not expect the trend to continue. If they did, they would be incentivized to replace more and not less.

From this robustness check, I conclude that the negative time trend is not a fluke of the data, but actually a robust result of the behavioral model in this data. Dairy farmers' expectations of technological progress do not fit the typical challenger versus defender model theorized by Miranda and Schnitkey (1995) and De Vries (2017). Future work should investigate the robustness of this result in other datasets and possibly over farm characteristics.

[Table 12 about here.]

List of Figures

1 50
2 Payoffs 51
3 Record Counts and Herd Sizes 52
4 Percent Exiting/Bred at Each Lactation (Uncensored Cows Only) 53
5 Flow Chart for "Breeding" Replacement Definition 54
6 Exit Rate Decomposed and Compared to Literature Involuntary Rates 55
7 56
8 Price Trends Adjusted and Unadjusted for Seasonality 57
9 Revenue Shock by Age 58
10 CCP In-Sample Predictions for Exit Rate (top) and Breeding Definition (bottom) . 59
11 Fixed Effects Distribution over Age 60
12 Compensating Variation in Theory 61
13 Compensating Variation 62

Figure 1

Dairy Cow Life Cycle

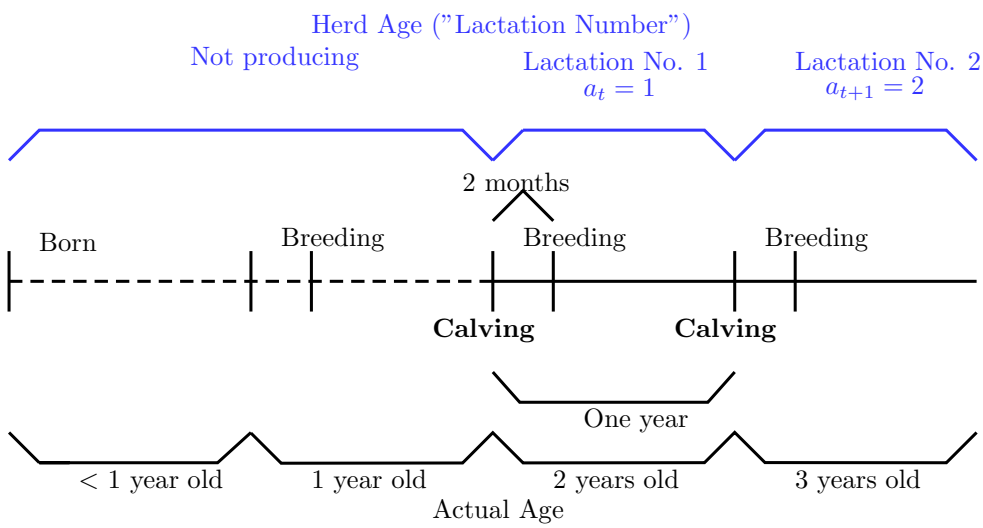
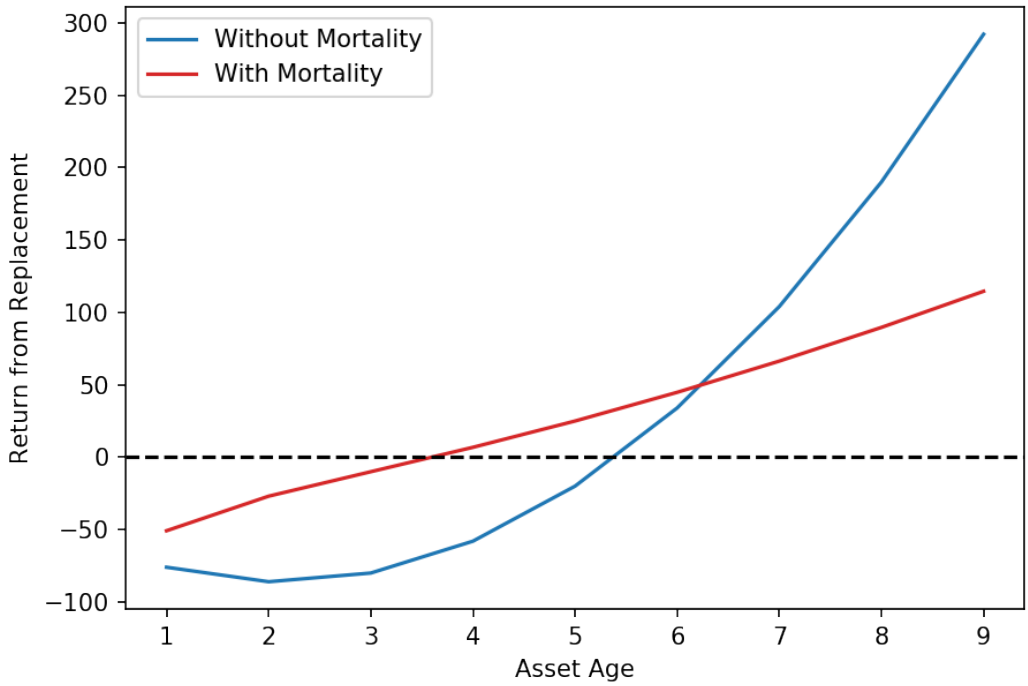


Figure 2: Payoffs



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

	Record Counts
Herds	1,511
Cows	267,341
Cows with Full History	169,484
Lactation Records	640,352

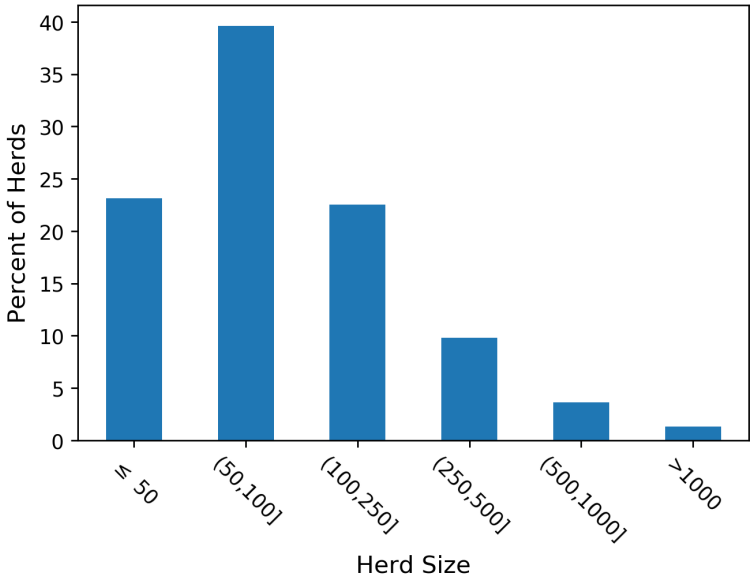


Figure 3: Record Counts and Herd Sizes

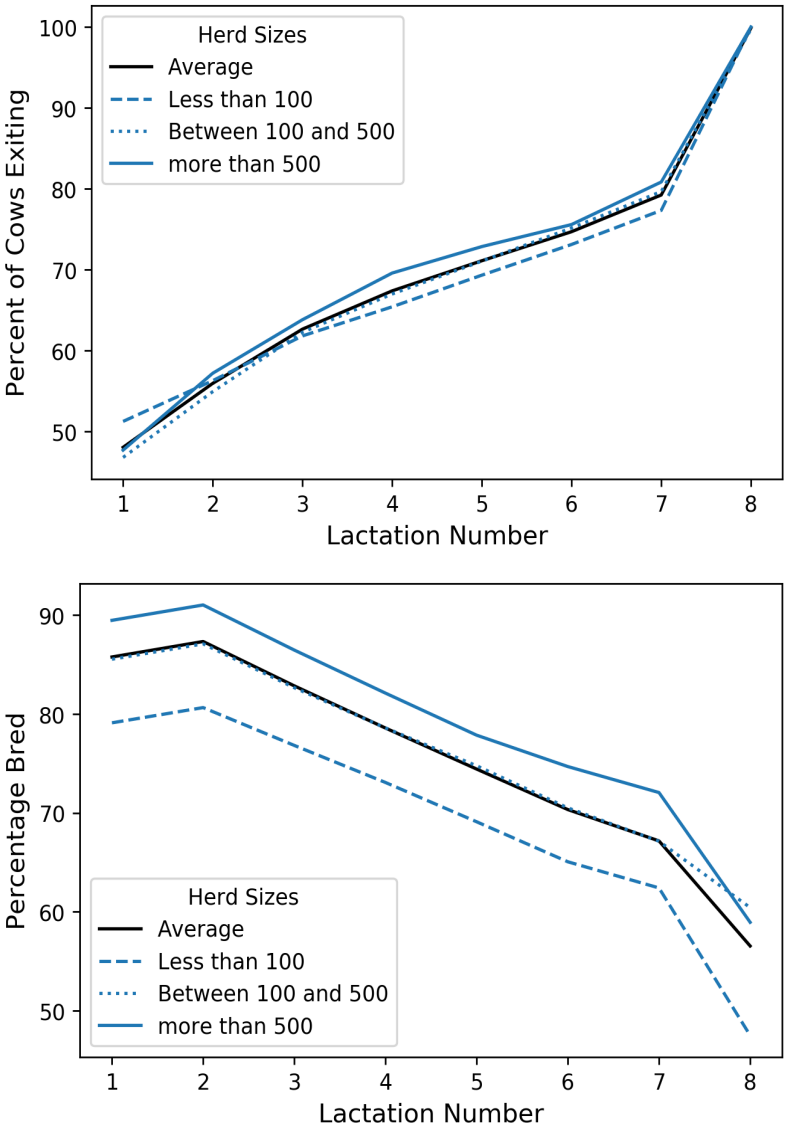


Figure 4: Percent Exiting/Bred at Each Lactation
(Uncensored Cows Only)

Figure 5: Flow Chart for "Breeding" Replacement Definition

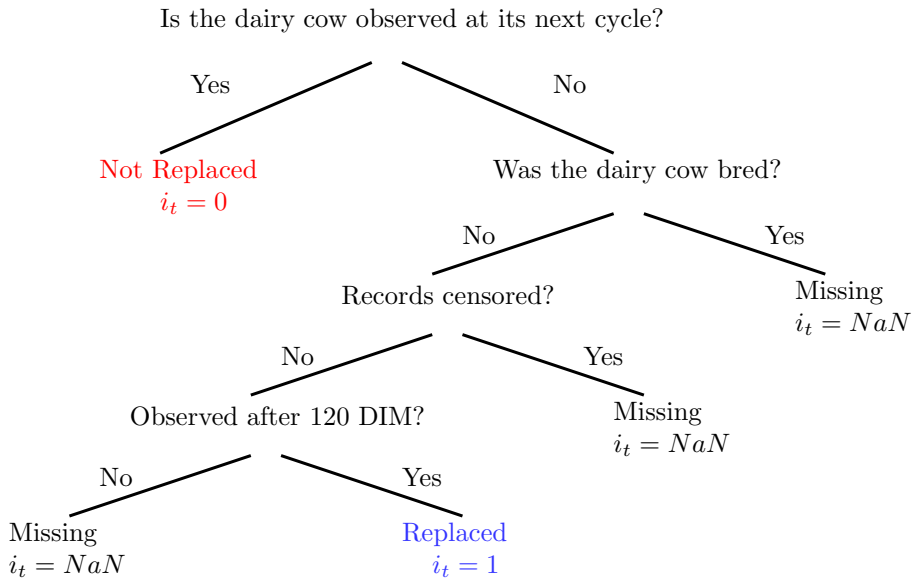


Figure 6: Exit Rate Decomposed and Compared to Literature Involuntary Rates

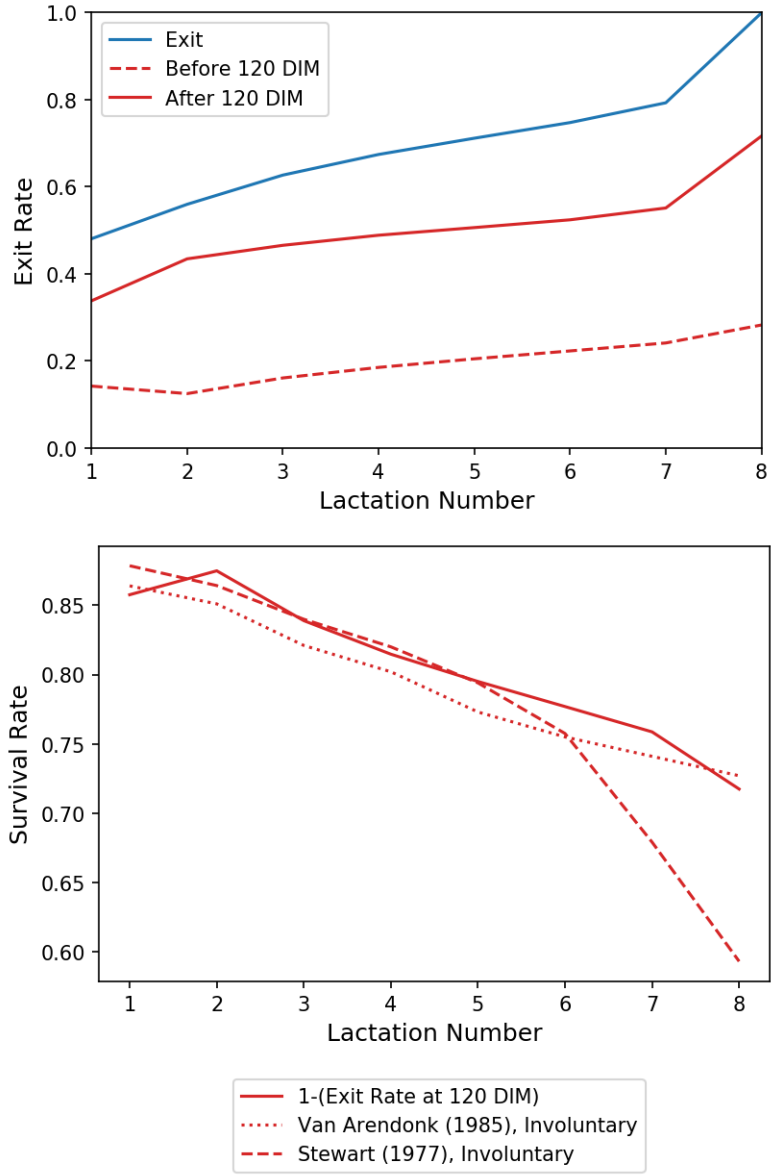


Figure 7

Exit Definitions in the Cow Production Cycle

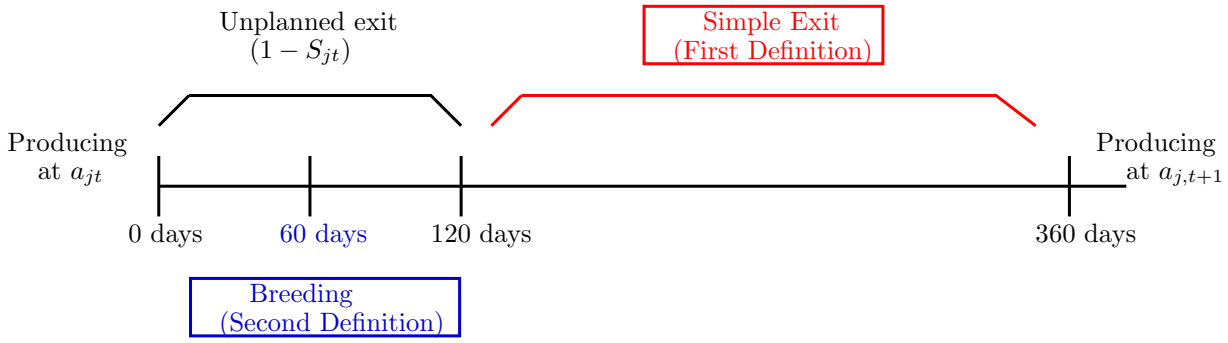
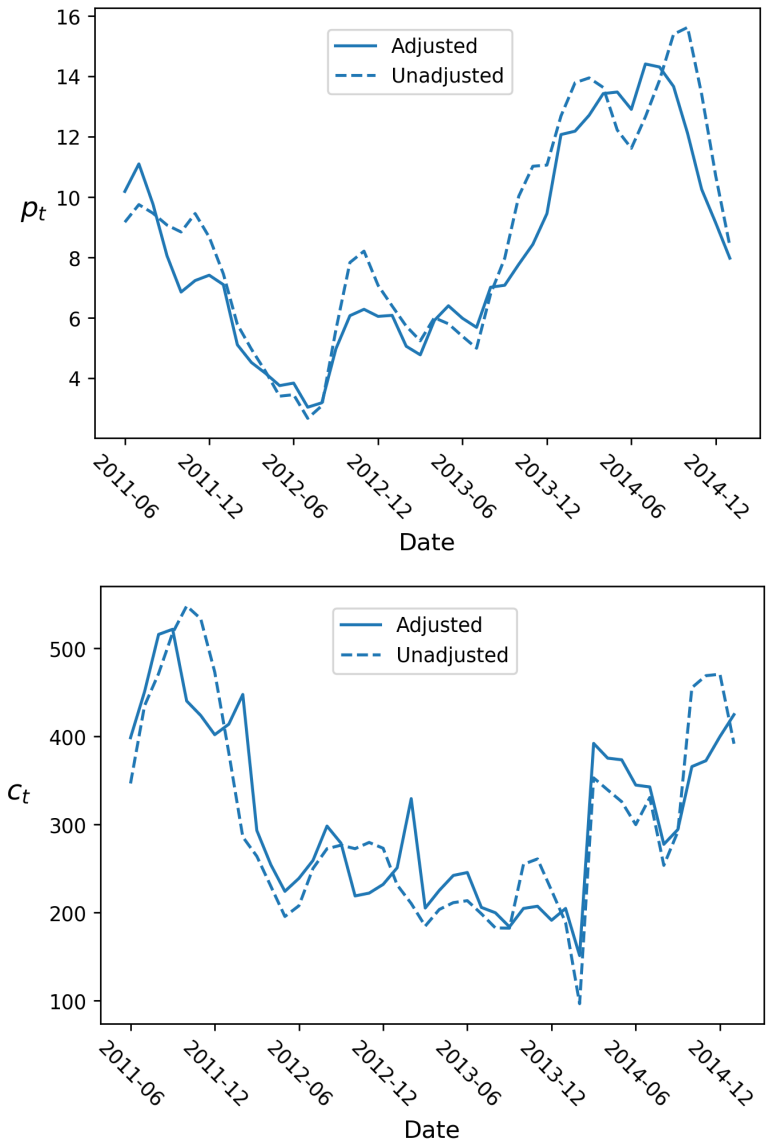


Figure 8: Price Trends Adjusted and Unadjusted for Seasonality



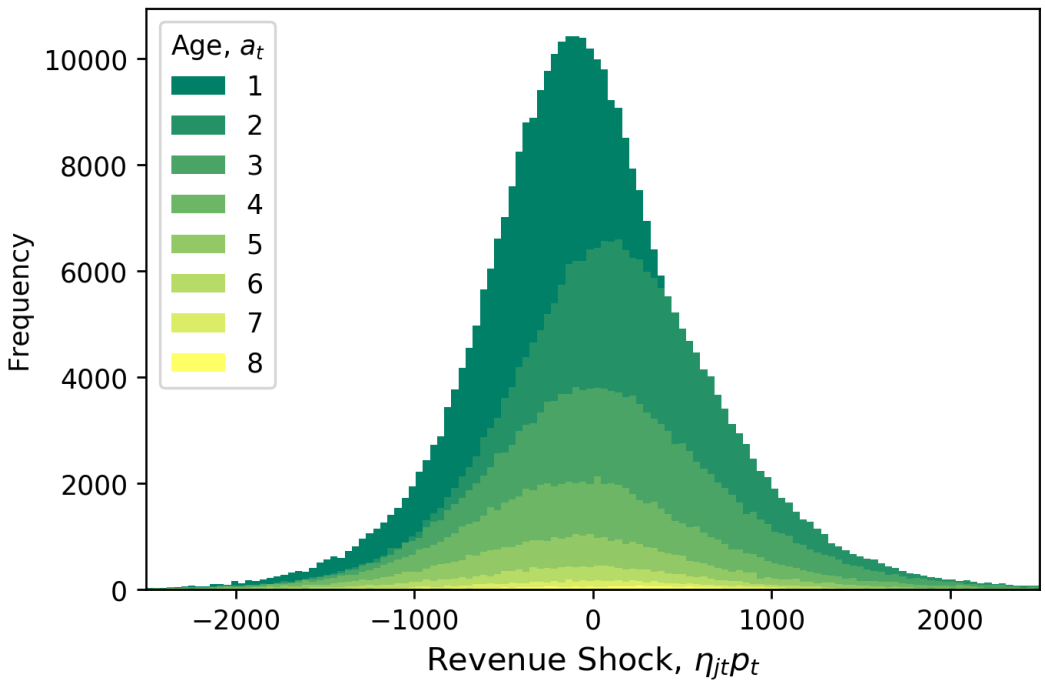


Figure 9: Revenue Shock by Age

				Percentiles		
	Mean	Std Dev	Count	2.5%	50%	97.5%
Age						
1	-110.75	584.87	322,842	-1297.59	-105.55	1040.42
2	136.05	676.23	241,250	-1171.22	120.21	1538.50
3	64.81	730.42	153,147	-1307.91	40.48	1589.48
4	-40.32	746.20	83,689	-1445.45	-60.69	1522.15
5	-80.61	724.43	39,950	-1477.22	-86.83	1394.08
6	-74.21	713.33	16,718	-1498.12	-61.58	1309.33
7	-19.51	725.10	6,217	-1547.24	4.17	1352.78
8	85.86	741.36	2,198	-1490.17	132.89	1442.50

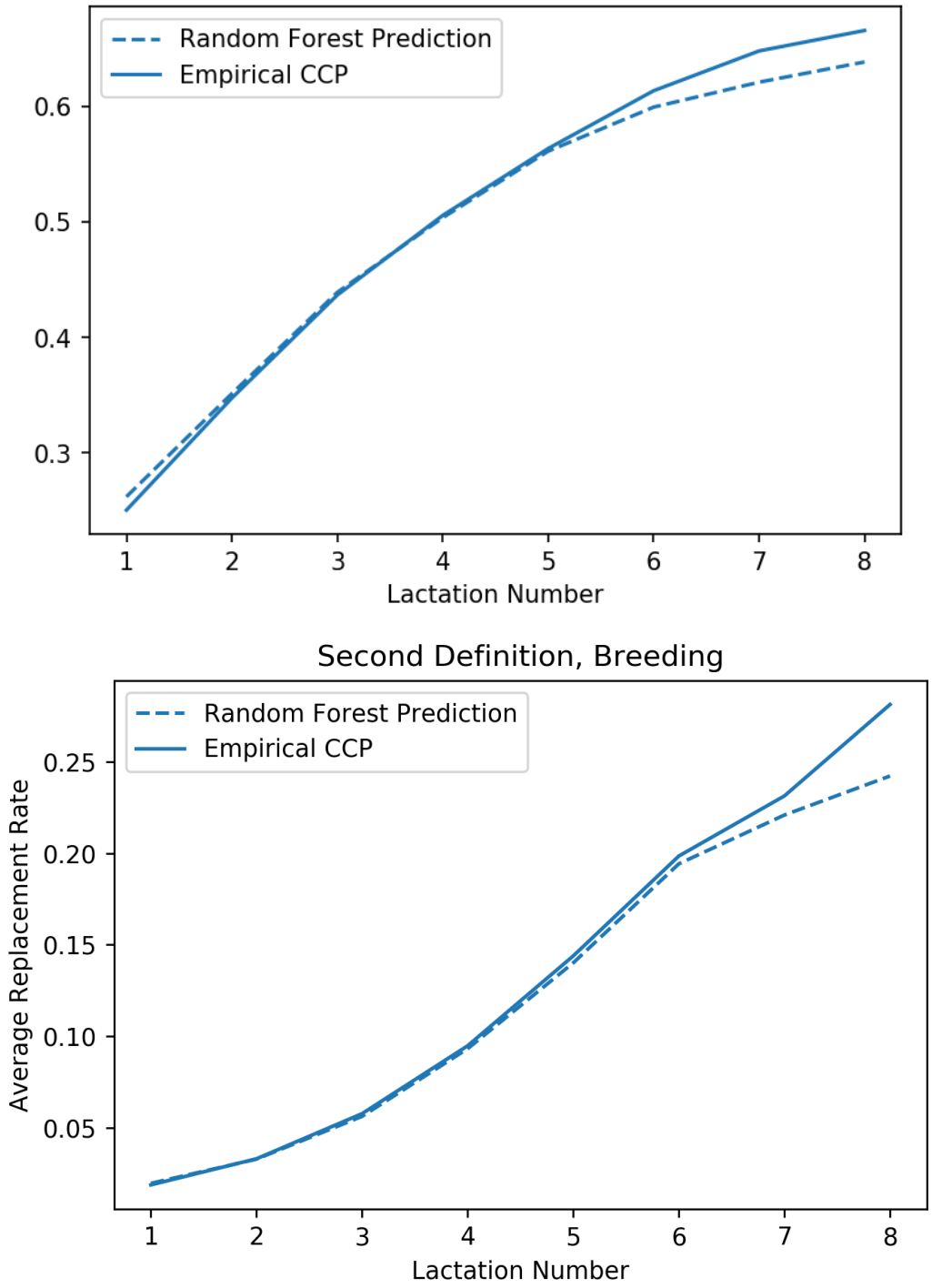
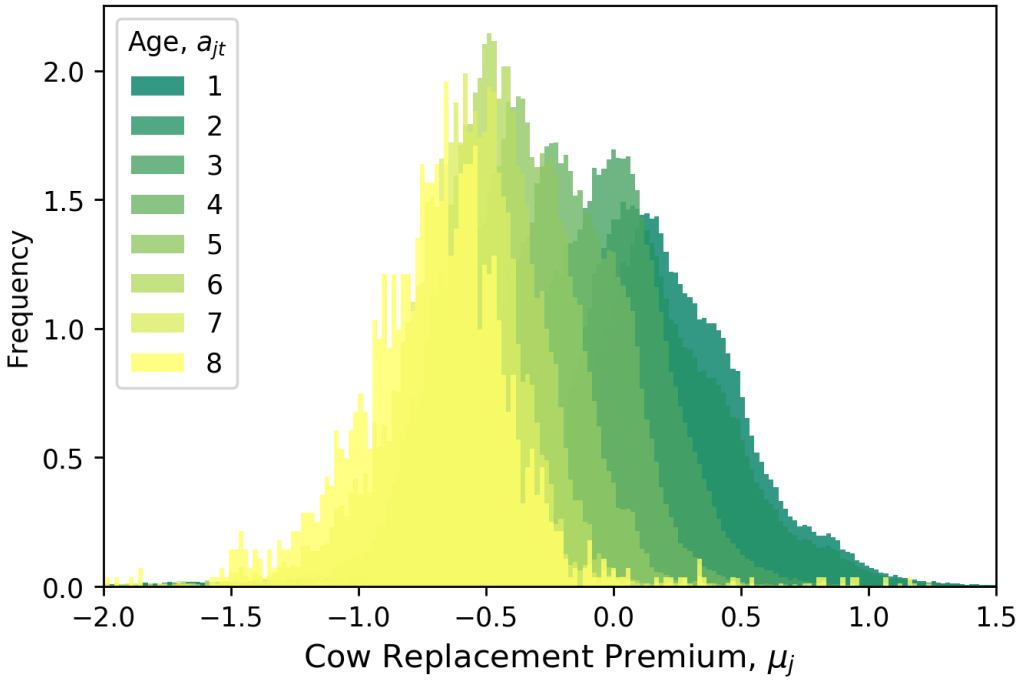


Figure 10: CCP In-Sample Predictions for Exit Rate (top) and Breeding Definition (bottom)

Figure 11: Fixed Effects Distribution over Age



Age	Mean	Std Dev	Count	2.5%	50%	97.5%
1	0.15	0.48	163,915	-1.09	0.18	0.89
2	0.10	0.44	215,697	-0.92	0.11	0.85
3	-0.05	0.30	134,365	-0.56	-0.03	0.45
4	-0.21	0.29	72,019	-0.68	-0.20	0.24
5	-0.38	0.27	33,800	-0.81	-0.37	0.06
6	-0.51	0.26	13,934	-0.98	-0.50	-0.12
7	-0.62	0.29	5,122	-1.21	-0.60	-0.21
8	-0.71	0.34	1,500	-1.39	-0.69	-0.13

Figure 12: Compensating Variation in Theory

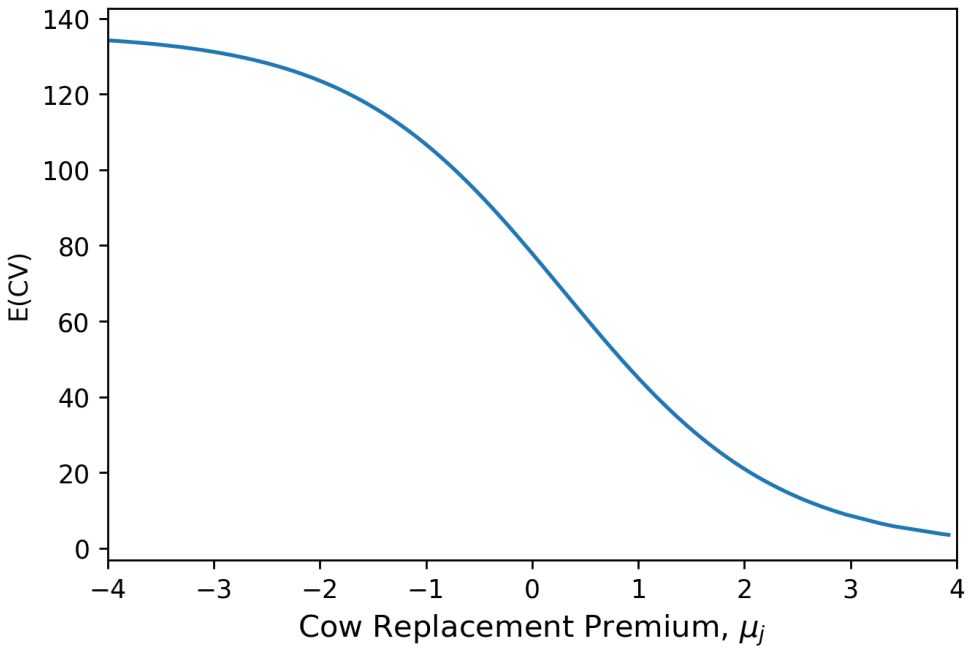
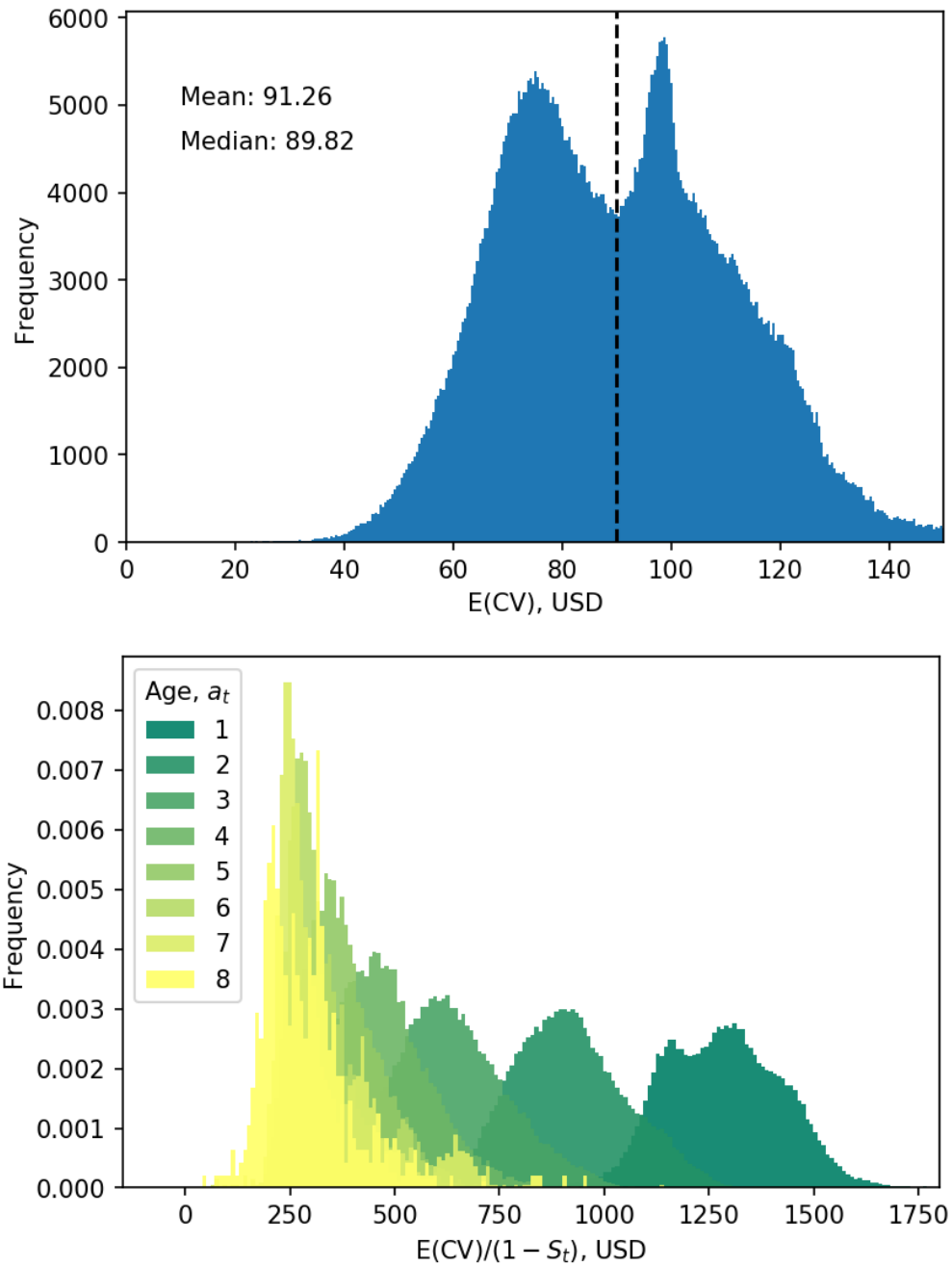


Figure 13: Compensating Variation



List of Tables

1 Model Summary 64
2 AR(1) Regressions for State Transitions 65
3 ECCP Model Estimates 66
4 Breeding Definition, ECCP Method 67
5 Model Results across Herd Size 68
6 Average CV by Age (Adjusted by Hazard Rate) 69
7 Average Indemnity for New Cows, by Herd Size 70
8 Milk Production Model 71
9 Comparison of Methods on Restricted Dataset 72
10 Second Definition across Herd Size 73
11 Different Fixed Effects Specifications 74
12 Other Robustness Checks 75

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
	$P(a_{t+1} = 1 \mid i_t) = \begin{cases} 1 & i_t = 1 \\ 1 - S(a_t) & i_t = 0 \end{cases}$	Evolution of age a
	$P(a_{t+1} = a_t + 1 \mid i_t) = \begin{cases} 0 & i_t = 1 \\ S(a_t) & i_t = 0 \end{cases}$	
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

Table 2: AR(1) Regressions for State Transitions

	η	p	c
μ	43.09	8.03	291.34
σ	647.43	0.94	54.73
ρ	0.30	0.94	0.79

Table 3: ECCP Model Estimates

		No Mortality $\delta = .99$	Mortality Risk $\delta = 0$	Mortality Risk $\delta = .95$	Mortality Risk $\delta = .99$
Time Trend	τ	-17.60	-16.60	-15.67	-15.63
		(0.32)	(0.27)	(0.30)	(0.29)
Penalty	α		840.44	1806.31	1851.58
			(62.99)	(65.39)	(70.08)
MC	γ	200.72	306.97	197.35	192.25
		(4.68)	(5.97)	(5.22)	(5.17)
Shock Correlation	ρ	0.202	0.185	0.180	0.179
		(0.004)	(0.003)	(0.003)	(0.003)
Age of Max	$-\frac{\beta_1}{2\beta_2}$	3.20	1.65	2.89	2.93
		(0.03)	(0.04)	(0.02)	(0.02)
Age of Free	$-\frac{\beta_1+\beta_2}{\beta_1}$	5.40	2.30	4.78	4.87
		(0.06)	(0.09)	(0.05)	(0.05)
Scale	λ	0.0020	0.0025	0.0023	0.0023
		(0.00003)	(0.00004)	(0.00004)	(0.00004)
Observations		640,352	640,352	640,352	640,352
Adjusted R^2		0.084	0.226	0.103	0.099

Bootstrapped "standard errors" in parentheses

Table 4: Breeding Definition, ECCP Method

		No Mortality	Mortality Risk $\delta = 0$	Mortality Risk $\delta = .95$	Mortality Risk $\delta = .99$
Time Trend	τ	-21.98	-12.82	-16.38	-16.64
		(0.86)	(0.40)	(0.69)	(0.71)
Penalty	α		1911.50	2330.74	2371.84
			(70.80)	(116.10)	(115.30)
MC	γ	132.33	299.47	127.35	114.84
		(10.90)	(7.93)	(10.25)	(10.42)
Shock Correlation	ρ	0.405	0.226	0.333	0.341
		(0.0113)	(0.0040)	(0.0087)	(0.0092)
Age of Max	$-\frac{\beta_1}{2\beta_2}$	3.83	2.35	3.55	3.60
		(0.026)	(0.021)	(0.022)	(0.023)
Age of Free	$-\frac{\beta_1+\beta_2}{\beta_1}$	6.65	3.69	6.10	6.20
		(0.052)	(0.041)	(0.045)	(0.046)
Scale	λ	0.0014	0.0029	0.0017	0.0017
		(0.00004)	(0.00004)	(0.00004)	(0.00004)
Observations		355,734	355,734	355,734	355,734
Adjusted R^2		0.215	0.462	0.242	0.234

Bootstrapped "standard errors" in parentheses

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	-14.97	-16.82	-16.24	-13.98	-14.52
		(0.58)	(0.61)	(0.68)	(0.68)	(0.64)
Penalty	α	2444.91	2548.06	1569.63	854.382	1707.36
		(150.99)	(163.92)	(146.54)	(182.93)	(163.72)
MC	γ	189.26	206.32	206.98	189.70	142.47
		(11.014)	(11.68)	(11.62)	(12.89)	(10.09)
Shock Correlation	ρ	0.168	0.185	0.186	0.185	0.150
		(0.008)	(0.009)	(0.007)	(0.009)	(0.006)
Age of Max	$-\frac{\beta_1}{2\beta_2}$	3.12	2.96	2.80	2.77	3.06
		(0.05)	(0.06)	(0.05)	(0.08)	(0.05)
Age of Free	$-\frac{\beta_1+\beta_2}{\beta_1}$	5.24	4.92	4.61	4.54	5.12
		(0.10)	(0.12)	(0.11)	(0.16)	(0.10)
Scale	λ	0.0019	0.0020	0.0023	0.0023	0.0030
		(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Observations		143,585	152,898	148,934	104,391	90,544
Adjusted R^2		0.004	0.063	0.161	0.148	0.127

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

Table 6: Average CV by Age
(Adjusted by Hazard Rate)

Age	Mean	Std Dev	Count	2.5%	50%	97.5%
1	1,319.93	143.51	163,915.0	1,077.61	1,304.37	1,610.97
2	1,028.42	190.31	215,697.0	736.20	993.53	1,487.16
3	767.50	193.27	134,365.0	481.05	729.33	1,228.66
4	604.27	187.53	72,019.0	353.46	558.80	1,085.97
5	520.35	204.34	33,800.0	282.86	455.25	1,069.50
6	494.92	242.87	13,934.0	228.49	403.51	1,110.74
7	523.91	307.30	5,122.0	185.67	414.10	1,258.98
8	601.02	425.03	1,500.0	121.25	429.15	1,551.97

Table 7: Average Indemnity for New Cows, by Herd Size

Herd Size	Mean	Std Dev	2.5%	50%	97.5%
Less than 100	1,666.14	184.20	1,360.79	1,653.39	2,069.68
100 to 250	1,772.00	184.49	1,461.41	1,753.35	2,188.50
250 to 500	1,113.94	126.80	902.68	1,098.01	1,364.25
500 to 1000	596.39	102.05	448.63	573.75	803.11
more than 1000	1,341.54	129.28	1,120.18	1,326.16	1,593.92

Table 8: 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

Table 9: Comparison of Methods on Restricted Dataset

		Myopic		Dynamic	
		First Definition	Second Definition	First Definition	Second Definition
Time Trend	τ	-11.56	-12.82	-10.76	-16.64
		(0.46)	(0.40)	(0.53)	(0.71)
Penalty	α	1870.64	1911.50	2657.39	2371.84
		(75.67)	(70.80)	(82.72)	(115.30)
MC	γ	187.30	299.47	105.59	114.84
		(7.66)	(7.93)	(8.11)	(10.42)
Shock Correlation	ρ	0.141	0.226	0.137	0.341
		(0.003)	(0.004)	(0.003)	(0.009)
Age of Max	$-\frac{\beta_1}{2\beta_2}$	1.88	2.35	2.47	3.60
		(0.03)	(0.02)	(0.02)	(0.02)
Age of Free	$-\frac{\beta_1+\beta_2}{\beta_1}$	2.75	3.69	3.94	6.20
		(0.06)	(0.04)	(0.04)	(0.04575)
Scale	λ	0.0036	0.0029	0.0032	0.0017
		(0.00006)	(0.00004)	(0.00006)	(0.00004)
Observations		355,734	355,734	355,734	355,734
Adjusted R^2		0.173	0.462	0.059	0.234

Bootstrapped standard deviations in parentheses
Discount rate set to 0.99

Table 10: Second Definition across Herd Size

		Less than 100	100 to 250	250 to 500	500 to 1000	More than 1000
Time Trend	τ	-21.8664	-19.9254	-18.3192	-13.4993	-13.9214
		(2.1021)	(1.9091)	(1.6495)	(1.6045)	(1.5072)
Penalty	α	3795.79	3616.42	2062.43	1367.64	2152.54
		(394.8643)	(359.5386)	(258.3953)	(286.2498)	(220.6126)
MC	γ	166.384	134.18	153.396	101.388	62.0186
		(30.3989)	(27.4336)	(24.4909)	(24.0195)	(20.735)
Shock Correlation	ρ	0.4404	0.411	0.364	0.3191	0.2056
		(0.0347)	(0.0284)	(0.0211)	(0.0216)	(0.0103)
Age of Max	$-\frac{\beta_1}{2\beta_2}$	3.9938	3.8205	3.453	3.3769	3.4549
		(0.0631)	(0.0571)	(0.0482)	(0.0555)	(0.044)
Age of Free	$-\frac{\beta_1+\beta_2}{\beta_1}$	6.9876	6.6409	5.906	5.7537	5.9098
		(0.1262)	(0.1142)	(0.0965)	(0.111)	(0.088)
Scale	λ	0.0011	0.0013	0.0016	0.0018	0.0028
		(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Observations		108,523	127,626	132,592	92,458	80,053
Adjusted R^2		0.430	0.422	0.427	0.408	0.349

Bootstrapped standard deviations in parentheses
Discount rate set to .99

Table 11: Different Fixed Effects Specifications

		No Fixed Effects	Herd Fixed Effects	Cow Fixed Effects
Premium	μ	-830.85		
		(20.37)		
Time Trend	τ	-14.08	-14.10	-15.63
		(0.11)	(0.11)	(0.29)
Penalty	α	1062.96	1088.78	1851.58
		(76.53)	(76.79)	(70.08)
MC	γ	254.68	254.54	192.25
		(5.75)	(5.68)	(5.17)
Shock Correlation	ρ	0.279	0.280	0.179
		(0.005)	(0.005)	(0.003)
Age of Max	$-\frac{\beta_1}{2\beta_2}$	2.53	2.53	2.93
		(0.03)	(0.03)	(0.02)
Age of Free	$-\frac{\beta_1+\beta_2}{\beta_1}$	4.06	4.06	4.87
		(0.06)	(0.06)	(0.05)
Scale	λ	0.0013	0.0013	0.0023
		(0.00002)	(0.00002)	(0.00004)
Observations		640,352	640,352	640,352
Adjusted R^2		0.182	0.184	0.099

Bootstrapped standard deviations in parentheses

Discount rate set to .99

Table 12: Other Robustness Checks

		Main Specification	SCC Shock	Birth Year Time Trend
SCC Shock			0.0239	
			(0.0067)	
Time Trend	τ	-15.63	-15.63	-160.35
		(0.29)	(0.29)	(11.02)
Penalty	α	1851.58	1843.16	1899.24
		(70.08)	(73.89)	(103.35)
MC	γ	192.25	191.49	80.71
		(5.17)	(5.21)	(5.33)
Shock Correlation	ρ	0.179	0.179	0.279
		(0.0034)	(0.0034)	(0.0060)
Age of Max	$-\frac{\beta_1}{2\beta_2}$	2.93	2.93	2.28
		(0.02)	(0.02)	(0.03)
Age of Free	$-\frac{\beta_1+\beta_2}{\beta_1}$	4.87	4.87	3.45
		(0.05)	(0.05)	(0.07)
Scale	λ	0.0023	0.0023	0.001
		(0.00004)	(0.00004)	(0.00003)
Observations		640,352	640,352	640,352
Adjusted R^2		0.099	0.087	0.099