The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

The Role of Animal Breeding in Productivity Growth: Evidence from Wisconsin Dairy Farms

Jared Hutchins, Department of Agricultural and Consumer Economics at the University of Illinois at Urbana-Champaign, jhtchns2@illinois.edu
Yating Gong, Department of Agricultural and Applied Economics at the University of Wisconsin Madison, ygong42@wisc.edu
Xiaodong Du, Department of Agricultural and Applied Economics at the University of Wisconsin Madison, xdu23@wisc.edu

Selected Paper prepared for presentation at the 2021 Agricultural & Applied Economics Association Annual Meeting, Austin, TX, August 1 – August 3

Copyright 2021 by Jared Hutchins, Yating Gong and Xiaodong Du. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

The Role of Animal Breeding in Productivity Growth: Evidence from Wisconsin Dairy Farms

Jared Hutchins[*] Yating Gong[†] Xiaodong Du[‡]

June 9, 2021

Abstract

We examine the relationship between investments in genetics and productivity growth on Wisconsin dairy farms. Following the empirical industrial organization literature, we apply the control function approach to incorporate farm level annual expenditure on breeding and adopted genetics into the law of motion of productivity to understand the mechanism of genetics affecting farm productivity in term of milk production. In our sample of Wisconsin dairy farms, the results indicate that farmers who spent more on breeding three years prior had higher productivity as did those that chose genetics with high production yield. We also find that dairy farms with high productivity reap the smallest gains from increasing spending on breeding or high-yield genetics, which suggests that there are diminishing returns to investing in genetics. In additions, omitting the breeding or genetics contribution from the law of motion is found to affect estimated factor returns. Although the sample size is not large, we demonstrate the importance of accounting for inter-temporal investments in genetics when understanding factor shares and productivity growth in US dairy.

Keywords: control function, genetics, production function

JEL Classification: Q15, Q12, O13, C25, C26

[*]Assistant Professor, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign. Email: *jhtchns2@illinois.edu*.

[†]PhD Student, Agricultural and Applied Economics, University of Wisconsin, Madison. Email: *ygong42@wisc.edu*.

[‡]Associate Professor, Agricultural and Applied Economics, University of Wisconsin-Madison. Email: *xdu23@wisc.edu*.

1 Introduction

The dairy science literature attributes as much as 50% of the growth in milk yield over the past decades to genetic improvement in dairy cows (Pryce and Veerkamp 2001; Thornton 2010; VanRaden 2004). Figure 1 charts the growth path of average cow milk yield from 2003 to 2019, which grew on average 1.4% a year during this period. At the same time, the productivity of dairy bulls available on the market has grown at a higher rate: milk yield, butterfat yield, and protein yield of dairy bulls grew on average 3.3%, 5.1%. and 4% per year during this period. When combined into an index, its growth path matches that of cow milk yield. While its precise contribution is difficult to quantify, it is clear that genetic improvement is a powerful vector of productivity growth in dairy.

While choosing genetics is critical for productivity improvement of a dairy herd, producers making investments in their herd genetics has been largely ignored when estimating dairy farm production functions (Njuki, Bravo-Ureta, and Cabrera 2020; Jang and Du 2019; Mukherjee, Bravo-Ureta, and De Vries 2013). By omitting this decision, conventional production function estimation may misattribute this productivity growth to other input factors, which results in biased input coefficients (De Loecker 2013). Moreover, not understanding this vital vector of productivity growth in dairy farming makes it difficult to understand the effect of new genetic improvements on the future of the dairy industry.

This paper investigates the effect of genetics investment on dairy farm productivity in Wisconsin. Using two rich data sources on farm- and animal-level decisions, we integrate spending on genetic improvements into the first order Markov process of productivity in the style of Levinsohn and Petrin (2003) and De Loecker (2013). Since investments in genetics impact productivity three years from the date of investment, variation in breeding investment helps identify the parameters in the farm production function. Using this method, we analyze both the relationship of genetic investment with productivity growth as well as the extent to which omitting investment in genetics biases estimated factor shares in the production function.

We find that both spending on breeding and the production merit of the genetics they choose are significant factors in the law of motion for productivity. In our sample of Wisconsin dairy farms, farmers who spent more on breeding three years prior had higher productivity as did those that chose genetics with high production yield. We also find that dairy farms with high productivity reap the smallest gains from increasing spending on breeding or high-yield genetics, which suggests that there are diminishing returns to investing in genetics. By accounting for genetics investment in the law of motion, there are considerable (though not statistically

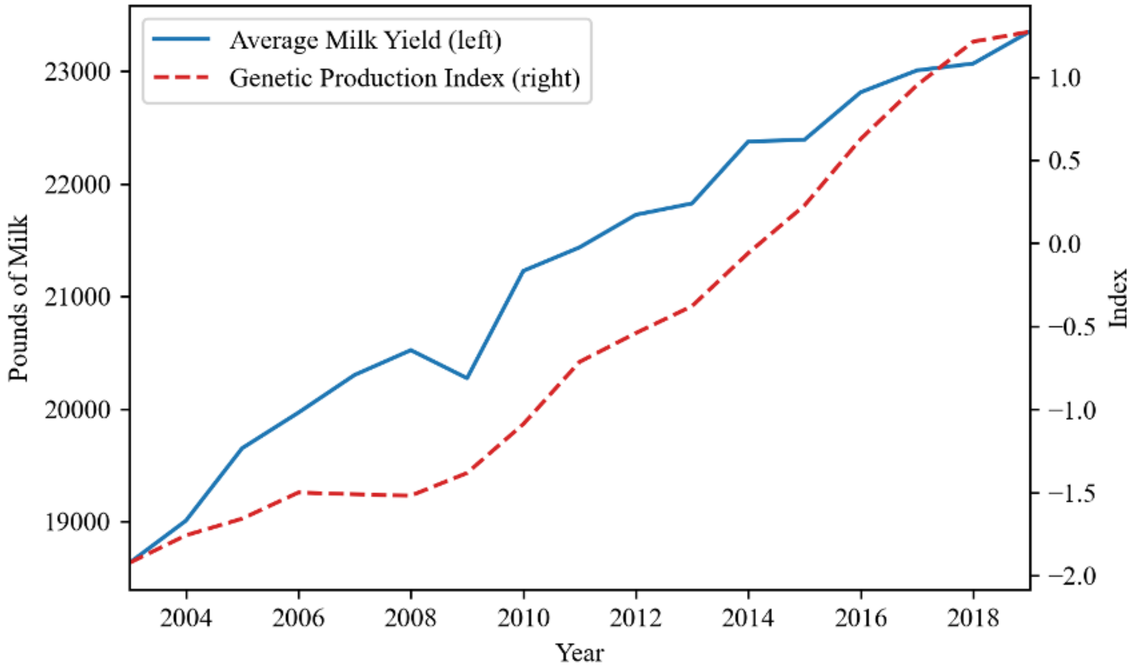

Figure 1: Average Milk Yield and Genetic Productivity, 2003-2019

significant) changes in the factor shares for feed, materials, and labor in the production function. While our study is limited by sample size, we demonstrate the importance of accounting for inter-temporal investments in genetics when understanding factor shares and productivity growth in US dairy. Despite its significance to the dairy industry, economic studies on the relationship between animal breeding and productivity growth are few and far in between. Studies on the crop sector have quantified the significant contributions plant breeding has made to the agricultural sector (Marasas, Smale, and Singh 2003; Babcock and Foster 1991; Nalley, Barkley, and Featherstone 2010). The few equivalent studies for the animal sector suggest this holds true in animal agriculture. For example, Townsend and Thirtle (2001) calculates that animal breeding research has a 35% return on investment in South Africa. Using a modified Malmquist productivity index, Atsbeha, Kristofersson, and Rickertsen (2012) find that 19% of the productivity growth rate in the Icelandic dairy sector is due to breeding. Apart from these few studies, the contribution of animal breeding to productivity growth is almost entirely unknown.

This is a glaring omission with respect to the estimation of factor shares in dairy farming. Of the existing literature, no studies have incorporated genetics as a dynamic investment into the dairy production function (Njuki, Bravo-Ureta, and Cabrera 2020; Jang and Du 2019; Mukherjee, Bravo-Ureta, and De Vries 2013). Much like the framework of Olley and Pakes (1996), dairy farms make investments every year in the genetics of their herd. Once a bull is chosen, it takes three years for its offspring to begin producing in the herd. This means farms every year make investments that will pay off in three years. The kinds of investments provide the kind of proxy variable we can use to identify production function coefficients (e.g. Levinsohn and Petrin 2003; Olley and Pakes 1996). Not only that, but these investments should also factor directly into the law of motion for productivity. De Loecker (2013) demonstrates in the case of exporting firms that neglecting critical variables in the law of motion leads to poorly estimated factor shares. In the case of dairy, by omitting inter-temporal breeding investments we may be misattributing output to various factors such as labor, feed, or even capital. Our work makes a significant contribution to our understanding of animal breeding and productivity by incorporating breeding investments directly into the law of motion. Ours is also the first study to examine the impact of genetic investments in health separate from production, which has been an important but so far unexplored aspect of productivity growth (see for example Townsend and Thirtle (2001)). There is an extensive literature on productivity and related empirical method for dealing with endogeneity in estimation. In the so-called control function approach, there are multiple choices for the proxy variable for the unobserved productivity, which is potential correlated with input choices, including, e.g., firm investment in Olley and Pakes (1996), intermediate input or material demand in Levinsohn and Petrin (2003), firm market value in Fan and Firestone (2010). In our case, we employ a unique proxy variable, genetic investment, observed for individual dairy farms. Different from capital investment, which can be lumpy and irregular, and intermediate input that may be less effective to capture productivity variation, genetic investment in the form of breeding cost, is an necessary and important cost that occurs every year. Although the amount of spending varies over time, it is a powerful tool for farmers to maintain and potentially improve composition, quality, and prosperity of the herd. Utilizing another unique feature of genetic investment, that is, a three-year gap between investment and effect being observed, we are able to quantify dynamic evolvement of herd productivity, which is a novel application of the method proposed in De Loecker (2013).

The paper proceeds as follows. First, we explain our empirical model which draws on De Loecker (2013) to incorporate genetic investments into the production function. We then describe our three data sources which provide farm- and cow-level information to estimate our model. We present the results of our analysis in the fourth section before concluding with a discussion of the significance of the results.

2 Empirical Model

In our model. farm i produces Q_{it} units of quality-adjusted milk in period t using these inputs: cows (C), labor (L), capital (K), feed (F) and intermediate inputs (M). The output is adjusted with the percentage composition of fat, the percentage composition of protein and the average somatic cell counts. Labor corresponds to hired workers. Capital refers to the stock of machinery and equipment, and intermediate inputs include the amount of fuel, electricity, veterinary services and custom services. Feed includes total amount of purchased feed. The production function is specified as:

$$Q_{it} = f(C_{it}, L_{it}, K_{it}, F_{it}, M_{it})exp(\omega_{it} + \epsilon_{it}) \tag{1}$$

where ω refers to Hicks-neutral farm-level productivity measure and ϵ is an i.i.d. measurement error term that accounts for random productivity shocks. The logarithmic form of Equation (1) is:

$$q_{it} = \beta_c c_{it} + \beta_l l_{it} + \beta_k k_{it} + \beta_f f_{it} + \beta_m m_{it} + \omega_{it} + \epsilon_{it} \tag{2}$$

2.1 Input endogeneity

To control for input endogeneity, that is the correlation between farmers' input decisions and unobserved productivity, we adapt the OP/LP (Olley and Pakes 1992; Levinsohn and Petrin 2003) method to use investments in genetics, G_{it}, as a proxy for unobserved productivity. The demand for breeding and genetic investment is defined as $G_{it} = g_t(\omega_{it}, C_{it}, K_{it})$. If we assume that breeding and genetic investment is monotonic in ϵ_{it} for any combination of C_{it} and K_{it}, then the inverse, $\omega_{it} = g_t^{-1} = \omega_t(G_{it}, C_{it}, K_{it})$, is a valid proxy for productivity. We use two proxies for genetic investment: total spending on breeding and the productivity indices of the bulls chosen that year.

By using the investments in genetics as a proxy for unobserved productivity, our production function becomes:

$$q_{it} = \beta_l l_{it} + \beta_f f_{it} + \beta_m m_{it} + \phi_t(G_{it}, c_{it}, k_{it}) + \epsilon_{it}$$

where ϕ_t is a third-order polynomial approximation.

We depart from the standard OP/LP model by modeling genetic choices (breeding with bulls) in the law of motion for productivity. We assume that genetic investments at time

$t-3$ impact productivity at time t. This time lag is a result of the biological constraints of dairy farming. When a cow becomes pregnant after breeding, it gives birth ten months later. That offspring is then bred at about one year old so that it can begin producing at two years old. Altogether, the breeding decision and the productivity resulting from that decision are about three years apart. If ξ_{it} denotes the productivity shock, the corresponding law of motion becomes:

$$\omega_{it} = \varphi_t(\omega_{it-1}, G_{it-3}) + \xi_{it} \tag{3}$$

For Equation (3), we consider two cases. In the first case, productivity follows an AR(1) process and is a simply linear function of genetic choice three years ago and a productivity shock. It is represented as:

$$\omega_{it} = \rho\omega_{it-1} + \gamma G_{it-3} + \xi_{it} \tag{4}$$

In the second case, productivity follows an AR(1) process and we model an interaction between the level of productivity last period and the level of investment that will be realized next period:

$$\omega_{it} = \rho\omega_{it-1} + \theta_1 G_{it-3} + \theta_2 G_{it-3}\omega_{it-1} + \xi_{it} \tag{5}$$

This formula is similar to the law of motion used by De Loecker (2013), where export status is interacted with productivity lags. In that case, the interactions allow firms with different levels of productivity to have different returns to exporting. In our case, it allows farms with different levels of productivity to have different returns to genetic investments. In essence, ω_{it-1} reflects the productivity of last year's cow cohort and ρ reflects how much those effects persist. The genetic investment G_{it-3} reflects the productivity of the incoming cohort, with θ_1 measuring the incoming cohort's impact on productivity. The effect of adding the new cohort to the existing cohort is measured by θ_2. Since the new cohort is usually genetically related to the current cohort, the interaction between current and incoming cohorts is a significant part of how productivity evolves on dairy farms. The sign of θ_2 also sheds light on whether there are decreasing or increasing returns to genetic investments in dairy. A positive sign would indicate that the most productive dairy farms get the most out of genetics investment (increasing returns), whereas a negative sign would indicate that the returns to genetics are lowest for the most productive farms (decreasing returns).

We assume $E[G_{it-3}\xi_{it}] = 0$ to obtain identification, which assumes that breeding decisions made three years prior are independent of the evolution of current productivity. This assumption gives a moment condition we can use to identify the parameters of interest:

$$E\{\xi_{it}(l_{it-1}, k_{it-1}, c_{it-1}, f_{it-1}, m_{it-1})\} = 0$$

Note that incorporating genetic investments is key to identifying the parameters in the production function. If we use the endogenous productivity changes in the model, that is $\omega_{it} = \phi_t(\omega_{it-1}) + \xi_{it}$, then the productivity shock _it contains the impact of genetic investment on productivity. If genetic investment correlates with any of these inputs, omitting genetic investment from the law of motion would bias the factor shares.

With the method, we address two questions concerning the role of genetic improvement in dairy productivity growth. First, we examine whether farmers' choice in genetics affects the law of motion in productivity. By incorporating farmers' choices in genetics, we elucidate to what extent genetic improvement in dairy cows has increased on-farm total factor productivity. Second, we examine the extent to which ignoring genetic improvement biases estimated factor shares in the production function.

2.2 Quality adjustment of output

As quality and quantity jointly determine milk price, dairy farmers take both into consideration when making genetic choices. Therefore we adjust milk production by quality attributes to measure productivity and its growth accurately. Following Atsbeha et al. (2012), we generate a milk quality index with three key attributes affecting milk price: nutrient component percentage, including the butterfat percentage (C_{it}^{fat}) and protein percentage ($C_{it}^{protein}$), and hygienic quality attribute represented by somatic cell count in unit milk (SCC_{it}). The higher the butterfat and protein percentage, the better the milk quality; The lower the SSC, the better the milk quality.

Let I_{it} denotes the unit milk value of farm i at time t. As the milk unit value is linear in protein percentage and fat percentage and close to linear in SSC (Atsbeha et al., 2012), we estimate the following hedonic price equation:

$$I_{it} = \alpha_0 + \alpha_1 C_{it}^{fat} + \alpha_2 C_{it}^{protein} + \alpha_3 SCC_{it} + \eta_{it}, \tag{6}$$

where the term η_{it} is a normally distributed random error. With the estimated parameter $\hat{\alpha}_0$,

$\hat{\alpha}_1$, $\hat{\alpha}_2$, and $\hat{\alpha}_3$, we calculate the average unit milk value $\bar{\hat{I}}$ with average milk quality attributes $\bar{C}^{fat}$, $\bar{C}^{protein}$ and $\overline{SCC}$ as follows:

$$\bar{\hat{I}} = \hat{\alpha}_0 + \hat{\alpha}_1 \bar{C}^{fat} + \hat{\alpha}_2 \bar{C}^{protein} + \hat{\alpha}_3 \overline{SCC} \tag{7}$$

The term $\hat{I}_{it}/\bar{\hat{I}}$ is specified as the milk quality index, which equals one when a farm's milk quality is equivalent to the sample average, greater (or lower) than one when a farm's milk quality is better (or lower) than average. The quality-adjusted milk output, $\tilde{Q}_{it}$, is then calculated as:

$$\tilde{Q}_{it} = \frac{\hat{I}_{it}}{\bar{\hat{I}}} Q_{it}$$

We next turn to our data sources for estimating the production function and understanding the role of genetic investment in dairy farm productivity.

3 Data

The main data used in this study consists of farm-level observations of output and input for 60 farms in Wisconsin over 2008-2018. The data is collected by the University of Wisconsin Center for Dairy Profitability (CDP). We also combine the CDP data with records of breeding decisions from Dairy Herd Improvement Associations (DHIAs) over 2012-2018 to measure the genetic quality of bulls each farm selects.

For the dairy farms in the CDP data set, we observe extensive output and input information in each year. On the output side, we observe revenue from selling milk, the quantity of sold milk, butterfat, and protein, and the somatic cell counts (SCC) of unit milk. On the input side, we observe the expenditure on hired labor, breeding, feed, fuel, and utility, the number of cows, and the value of building and machinery and equipment owned by the farms. To compute the real value of output and input, the revenue and the expenditure are deflated by price indexes from the USDA Quick Statistics. The milk revenue is deflated by the dairy product price index. The value of machine and equipment is deflated by a machinery price index. The value of buildings is deflated by a building material price index. The expenditure on hired labor and feed is deflated by wage rate and forage feed price index respectively. The expenditure on utilities is deflated by Wisconsin electricity average retail price (cents/kWh). The breeding fee is deflated by CPI index.

Table 1 presents the descriptive statistics of the inputs and output. During 2008-2018, the average farm size measured by average milk revenue and herd size shows a growing trend. The average milk revenue increases from 73 million in 2008 to 132 million in 2018 and grows by 80 percent. The average herd size grew 64%, from 177 in 2008 to 290 in 2018. During the same period, the average expenditure on breeding grows by 42 percent. Table 2 shows the descriptive statistics of milk quality indexes containing butterfat percentage, protein percentage, and SCC. It shows that there is minor change in butterfat percentage and protein percentage over 2008-2018. However, SCC decreased by 40 percent during the period. SCC is typically used as a proxy for the incidence of mastitis, so its decrease implies increasing milk quality and cow health.

Our data on genetic indices comes from The Council on Dairy Cattle Breeding (CDCB). The CDCB estimates genetic indices for dairy sires called "predicted transmitting ability" (PTAs) three times a year which are publicly available. The indices measure how much a sire will "transmit" performance relative to a base bull. For example, if a bull has a butterfat PTA score of 50, then the farmer can expect that offspring from that bull will produce 50 more pounds of milk than the base bull (whose PTA is zero). Genetics companies use these indices to price dairy bulls as well as by dairy farmers to inform their choices.

We use over 140,000 dairy bull evaluations from the period 2012-2019 and focus on five traits: milk yield, fat yield, protein yield, somatic cell score, and daughter pregnancy rate. The first three are production related, the fourth is related to health, and the last trait is a measure of fertility. To construct production and health indices, we use principal component analysis. In our calculations, the first two components captured 83% of the variance in these five traits. The first component was highly correlated with the production traits, while the second component was highly correlated with somatic cell score and daughter pregnancy rate. We use the first component as our "production index" and the second as our "health index." The details of our PCA calculation can be found in the Appendix.

After estimating these indices, we match this data to DHIA data of farm breeding choices. DHIAs are farmer cooperatives that collect monthly data for genetic evaluation and farm benchmarking. In addition to collecting production data for individual cows, they collect data on which bull was used to breed which cow. By matching the bull index in DHIA to its evaluation from the CDCB, we know the bull's production and health index scores from the time it was chosen. Unfortunately, a limited number of our CDP farms are also in DHIA so we can only use the genetic index scores for a subset of our data. We take the average index scores for each farm in each year, which produces some noise in the measure. Also, much like breeding

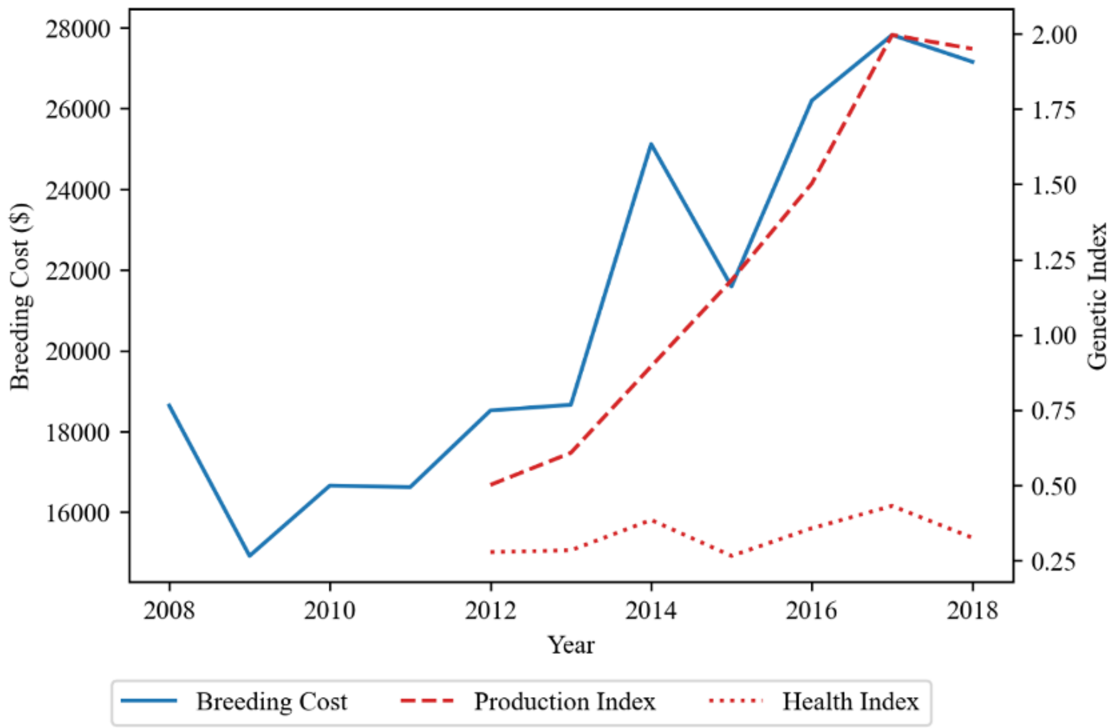

Figure 2: Average Breeding Expenditure and Genetic Production Index, 2008-2018

spending, not all choices the farms make result in offspring. Still, this noisy measure of genetic investment is completely unique since it can capture both the production and health aspects of their investments. By examining the subset of data belonging to both the CDP data and the DHIA, we have a never-before-seen look into how dairy farmer breeding choices translate into productivity.

Combining both CDP and DHIA data, we obtain two proxy measures for genetic investments: spending on breeding and the average production and health indices of their bull choices. Figure 2 shows the growing trend of average breeding expenditure and the production ability of the chosen sires in our sample. These Wisconsin dairy farms made breeding decisions similar to the rest of the country: the genetic production index is growing rapidly over time just as in Figure 1. In contrast, the health index scarcely grew at all during this period. There seems to be little variation in the health index during this period, reflecting a tendency for dairy farmers to invest in milk production more than in health traits.

In the next section, we present our estimation results from this unique dataset to explore the role of breeding in accelerating or decelerating productivity growth in dairy. Comparing across methods, we also determine whether omitting genetic investment leads to biased estimates of factor shares.

Table 1: Summary Statistics of CDP Sample

	All	2008	2013	2018	Change % over 2008-2018
Milk sold income (1,000\$)[1]	91,215 (104,043)	73,788 (86,610)	88,498 (97,403)	132,895 (142,901)	80
Milk (1,000 Pounds)	5,469 (6,152)	4,347 (5,080)	5,463 (5,992)	7,786 (8,195)	79
Butter fat (1,000 Pounds)	202 (224)	161 (186)	202 (214)	297 (312)	84
Protein (1,000 Pounds)	165 (185)	131 (154)	166 (181)	234 (246)	79
Somatic cell count index	186 (88)	224 (100)	180 (69)	120 (33)	-46
Herd size	213 (213)	177 (183)	208 (204)	290 (282)	64
Capital, buildings, and equipment (1,000\$)[2]	376 (462)	326 (523)	357 (398)	516 (527)	58
Feed (1,000\$)[3]	306 (356)	225 (313)	282 (305)	391 (391)	74
Hired labor (1,000\$)[4]	117 (154)	105 (146)	122 (156)	136 (187)	30
Utility (100 kWh)	19 (16)	18 (18)	18 (13)	24 (20)	33
Breeding fee (1,000\$)[5]	20 (20)	19 (21)	19 (19)	27 (23)	42
# Observations	372	45	35	19	-

[1] Milk sold income is deflated by dairy product price index.
[2] The market value of building is deflated by building material price index. The market value of machinery and equipment is deflated by machinery price index.
[3] The expenditure on feed is deflated by forage feed price index.
[4] The expenditure on hired labor is deflated by wage rate price index.
[5] The expenditure on breeding is deflated by CPI.
* Standard deviations are in parentheses.

Table 2: Summary Statistics of Milk Quality Indexes

Symbol	Description	Unit	All	2008	2013	2018	Change % over 2008-2018
I_{it}^{fat}	Butterfat	%	3.75 (0.24)	3.76 (0.26)	3.80 (0.19)	3.84 (0.14)	2
$I_{it}^{protein}$	Protein	%	3.04 (0.19)	3.09 (0.37)	3.06 (0.11)	3.04 (0.08)	-2
I_{it}^{SCC}	Somatic cell count (SCC)	1,000 cells/ml	200 (90)	218 (93)	180 (70)	134 (40)	-39
$\frac{\hat{V}_{it}}{\hat{V}}$	Linear Adjuster	#	0.91 (0.03)	0.92 (0.03)	0.88 (0.02)	0.90 (0.01)	-2
Q_{it}	Milk production	1,000 Pounds	4514 (5448)	4,087 (4,751)	4,707 (5,504)	7,334 (7,937)	79
$\tilde{Q}_{it}$	Adjusted milk production	-	4080 (4916)	3,791 (4,436)	4,120 (4,786)	6,565 (7,093)	73
# Observation			738	60	46	28	

Note: The estimated parameters for Equation (6) are butterfat: $1.06^{***}(0.19)$, protein: $0.57^{**}(0.24)$, and SCC: $-0.003^{***}(0.001)$. Standard errors are in the parentheses.
*** p<0.01, ** p<0.05, * p<0.1.

4 Results

Our two proxy measures capture different aspects of genetics investments. Total expenditure on breeding contains the cost of the bull, labor fees, and management fees. On average, the most productive genetics tend to be the most expensive, so we can assume breeding expenditure is weakly increasing in genetic quality. Our genetic indices measure only the average traits chosen in that year, which helps us measure the quality of the investments being made. Taken together, we have a unique picture of what sorts of investments dairy farmers are making.

Table 3 presents the partial correlation between genetic measures and the farm's inputs. The results show that dairy farms with larger herd size, more employees, and more capital stock tend to spend more on breeding. However, the genetic production index is not correlated with herd size, capital stock, employment or feed when all these inputs, the year fixed effect and county fixed effect are controlled. The labor and management fees in the breeding process are strongly correlated with herd size, capital stock, and employment, but the quality of genetics is not.

Table 4 presents the parameters estimated with four models: an OLS model, a standard LP model in which breeding expenditure is the proxy variable, an LP model with breeding

Table 3: Genetic Investment and Inputs

	Breeding Cost			Genetic Production Index		
	(1)	(2)	(3)	(1)	(2)	(3)
Herd Size	0.874***	0.616***	0.676***	0.465***	0.109	0.020
	(0.042)	(0.110)	(0.113)	(0.088)	(0.225)	(0.157)
Capital	0.105***	0.110***	0.102***	-0.025	-0.064	-0.060
	(0.028)	(0.028)	(0.031)	(0.062)	(0.063)	(0.039)
Employment	-	0.116***	0.172***	-	-0.118*	0.090
		(0.039)	(0.040)		(0.061)	(0.064)
Feed	-	-0.032	0.002	-	0.386***	0.082
		(0.073)	(0.081)		(0.146)	(0.134)
Material	-	0.144	-0.097		0.153	0.193
		(0.094)	(0.095)		(0.191)	(0.122)
Genetic health index	-	-	-	Yes	Yes	Yes
Year FE	No	No	Yes	No	No	Yes
County FE	No	No	Yes	No	No	Yes

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors are in parentheses.

Table 4: Estimated Parameters of Productivity Model

	OLS	Standard LP Model Proxy Var = Breeding Cost	LP model with Breeding expenditure in Law of Motion of Productivity		Diff between coefficients in column (2) and column (4) (%)
	(1)	(2)	(3)	(4)	
Labor	0.074*** (0.012)	0.067*** (0.017)	0.057*** (0.015)	0.057*** (0.017)	-15
Capital	0.024*** (0.008)	0.016** (0.008)	0.014* (0.007)	0.014 (0.008)	-1
Herd Size	0.804*** (0.027)	0.802*** (0.050)	0.792*** (0.053)	0.791*** (0.063)	-1
Feed	0.120*** (0.021)	0.094* (0.056)	0.122** (0.057)	0.125* (0.064)	33
Material	0.023 (0.020)	0.089 (0.055)	0.038 (0.032)	0.039 (0.052)	-56
Constant	8.179*** (0.145)	-	-	-	-
Interact term of productivity and breeding cost	-	-	No	Yes	-
# Observations	372	372	372	372	-
Productivity	-	8.04	8.36	8.35	-

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are obtained by block bootstrapping.

Table 5: Estimated Parameters of the Law of Motion

	LP model with Breeding expenditure in Law of Motion of Productivity (3)	(4)
Productivity, t-1	0.612***	1.04***
	(0.050)	(0.255)
Breeding Cost, t-3	0.015***	0.423*
	(0.004)	(0.240)
Interact term of Productivity (t-1) and Breeding expenditure (t-3)	-	-0.049*
		(0.029)
Year FE	Yes	Yes
Obs.	296	296
R-squared	0.732	0.735

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors are in parentheses.

expenditure in the law of motion (Equation 4), and an LP model where breeding expenditure is interacted with productivity in the law of motion (Equation 5). In these last two models breeding expenditure is used as the proxy variable.

The results show that including breeding expenditure in the law of motion of productivity has no effect on the capital and herd size coefficients. However, when we add breeding expenditure to the law of motion, the coefficients for labor and materials shrink and the coefficient for feed expenses grows. We see a 15% decrease in the feed coefficient, a 56% decrease in the materials coefficient, and a 33% increase in the feed coefficient. Because of the small sample size, these differences are likely not statistically significant. Regardless, there is evidence that inclusion of breeding investments in the law of motion will alter the estimation of the factor returns. Including a non-linear interaction term does not appear to affect estimation given that the Column 3 and Column 4 estimates are essentially the same.

Table 5 presents the coefficients for the law of motion of productivity corresponding to columns 3 and 4. The results indicate that breeding expenditure that occurred three years ago has a positive effect on current productivity. The interaction effect is negative, which means farms with higher productivity have lower returns to spending on breeding. In other words, the productivity boost of a better cohort diminishes the more productive the current cohort is. One way to interpret this pattern of coefficients is that there are diminishing returns to investments in genetics.

To explore the impact of genetic progress on productivity in-depth, we merge CDP data with our DHIA data to introduce genetic production and health indices into the model. The summary statistics for this sample are in Table 6. The merged CDP-DHI sample has 140

Table 6: Summary Statistics for CDP-DHI Sample

	Mean	Std. Err.	Minimum	Maximum
Milk sold income (1,000\$)[1]	127,665	137,428	10,140	632,094
Milk (1,000 Pounds)	7,469	7,867	640	35,528
Butter fat (1,000 Pounds)	279	288	24	1,305
Protein (1,000 Pounds)	225	236	20	1,062
Somatic cell count index	148	58	61	374
Herd size	275	259	38	1,177
Capital, building & machinery & equipment (1,000\$)[2]	517	536	4	2,401
Hired labor (1,000\$)[3]	147	191	0	855
Feed (1,000\$)[4]	424	446	13	2,150
Utility (100 kWh)	23	19	4	84
Breeding fee (1,000\$) [5]	26	24	1	92
Genetic production index	0.80	0.56	-1.30	1.91
Genetic health index	0.30	0.44	-1.32	1.40
# Observation		88		

[1] Milk sold income is deflated by dairy product price index.
[2] The market value of building is deflated by building material price index. The market value of machinery and equipment is deflated by machinery price index.
[3] The expenditure on hired labor is deflated by wage rate price index.
[4] The expenditure on feed is deflated by forage feed price index.
[5] The expenditure on breeding is deflated by CPI.

observations over 2012-2018. The average herd size in the merged sample is 262 cows per farm, larger than the average herd size in the CDP sample, 213 cows per farm. The average output and input are also larger for dairy farms in the merged CDP-DHI sample.

Table 7 compares the coefficients of the model using breeding expenditures, genetic indices, or both in the law of motion. The coefficients change very little between these models, except the materials coefficients that are not statistically different from zero in any model. Table 8 shows how the genetic indices factor in the law of motion compare to that of breeding expenditure. Like breeding expenditure, production and health indices from bulls chosen three years ago correlate positively to current productivity (though only the production index is statistically different than zero at the 90% level). The interaction effects are negative, so similar to breeding expenditure there may be diminishing returns to increasing the genetic indices of dairy cattle. Unfortunately, the sample size is so small that few of these coefficients are estimated with any precision. Regardless, the positive point estimates for both the production and health indices suggest that there are more interesting relationships to explore in this area.

After incorporating both spending and the indices into the law of motion, the effect of the production index on productivity diminishes to .460 while the health index still has a coefficient of 2.839. This would suggest that breeding expenditure is the most correlated with production

Table 7: Estimated Parameters of Productivity Model with CDP-DHI Sample

	LP model with Breeding expenditure in the Law of Motion of Productivity	LP model with Genetic Indexes in the Law of Motion of Productivity	LP model with Both Breeding expenditure and Genetic Indexes in the Law of Motion of Productivity
Labor	0.057***	0.059	0.054***
	(0.013)	(0.077)	(0.018)
Capital	0.014**	0.012	0.014**
	(0.006)	(0.021)	(0.007)
Herd size	0.814***	0.817***	0.823***
	(0.045)	(0.216)	(0.059)
Feed	0.128***	0.146**	0.128*
	(0.045)	(0.071)	(0.067)
Material	0.005	0.026	0.001
	(0.045)	(0.063)	(0.059)
# Observation		88	
Productivity	8.56	8.12	8.59

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors in parentheses are obtained by block bootstrapping.

traits. Interestingly, for the same level of spending health traits would appear to have the most important effects on productivity (though only in point estimate). This provides some descriptive evidence that investments in animal health may have some important effects on productivity growth.

5 Conclusion

Our work examines the relationship between investments in genetics through breeding and productivity growth in the US dairy industry. Using a detailed dataset of Wisconsin dairy farms, we incorporate expenditures on breeding services and genetic indices into the productivity law of motion. We test whether or not omitting genetics investments impacts the estimation of factor shares in the dairy production function. We find that including genetics investments in the law of motion changes the point estimates for feed, labor and materials in the production function. While breeding investments made three years ago positively impact productivity in the current period, the returns to investment are lowest for farms with high productivity. When a dairy cow cohort is already productive, investing in a new cohort does not have as high returns. Put differently, we find that there are diminishing returns to investments in genetics on Wisconsin dairy farms.

Table 8: Estimations for the Law of Motion Model with CDP-DHI Sample

	LP model with Breeding expenditure in the Law of Motion of Productivity	LP model with Genetic Production Index in the Law of Motion of Productivity	LP model with Both Breeding expenditure and Genetic Indexes in the Law of Motion of Productivity
Productivity,t-1	0.434 (1.602)	1.013*** (0.317)	0.906 (1.692)
Breeding Cost, t-3	0.082 (1.274)	-	0.378 (1.545)
Genetic Production Index, t-3	-	2.890* (1.518)	0.460 (2.147)
Genetic Health Index, t-3	-	4.598 (2.978)	2.839 (3.372)
Productivity,t-1 * Breeding expenditure, t-3	-0.007 (0.149)	-	-0.041 (0.179)
Productivity, t-1 * Genetic Production Index, t-3	-	-0.358* (0.187)	-0.056 (0.250)
Productivity,t-1 * Genetic Health Index, t-3	-	-0.563 (0.366)	-0.328 (0.392)
Year FE	Yes	Yes	Yes
Obs.		55	
R-squared	0.580	0.408	0.600

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors are in parentheses.

Our work has several limitations that may be able to be addressed in future work. First, the sample size in this study was very small. The sorts of data needed to understand investments in genetics in the context of other inputs are hard to come by. While datasets like the Agricultural Resource Management Survey (ARMS) have detailed information about capital and labor inputs, they do not have information about breeding investments. This a critical data gap that still exists when studying dairy farm operations and leads us to our second limitation. In this study, we used expenditures on breeding as a proxy for genetic investments. This may be a crude proxy for what we are actually trying to measure: investments in genetic quality. Data on individual breeding decisions from DHI can help fill this gap, but there was so little cross-over in our sample between DHI and CDP farms that the estimation was very imprecise. Further data collection efforts should particularly focus on trying to understand how much dairy farmers pay for different kinds of genetics. This kind of data is vital to understanding the link between investments in genetics and productivity growth in a more precise way.

Strategic breeding has been one of the most important innovations in animal agriculture in the past century. Yet, it has been practically ignored in the economics field when studying productivity growth in industries like the dairy sector. Our work is a crucial first step in determining just how important genetic improvement has been in the context of other technological innovations in dairy. The next century will no doubt pose new challenges for animal industries such as dairy, and new directions in animal breeding will be a part of addressing them. Understanding how farmers make breeding investments currently is essential to understanding how the industry will evolve tomorrow.

References

Atsbeha, Daniel Muluwork, Dadi Kristofersson, and Kyrre Rickertsen. 2012. "Animal Breeding and Productivity Growth of Dairy Farms." American Journal of Agricultural Economics 94 (4): 996–1012. https://doi.org/10.1093/ajae/aas033.

Babcock, Bruce A., and William E. Foster. 1991. "Measuring the Potential Contribution of Plant Breeding to Crop Yields: Flue-Cured Tobacco, 1954–87." American Journal of Agricultural Economics 73 (3): 850–59. https://doi.org/10.2307/1242837.

De Loecker, Jan. 2013. "Detecting Learning by Exporting." American Economic Journal: Microeconomics 5 (3): 1–21. Fan M, Firestone S. 2010. International journal of industrial organization. Internatinal Journal of Industrial Organization 28(5):434–440.

Jang, Heesun, and Xiaodong Du. 2019. “Evolving Techniques in Production Function Identification Illustrated in the Case of the US Dairy.” Applied Economics 51 (14): 1463–77. https://doi.org/10.1080/00036846.2018.1527457.

Levinsohn, James, and Amil Petrin. 2003. “Estimating Production Functions Using Inputs to Control for Unobservables.” The Review of Economic Studies 70 (2): 317–41.

Marasas, Carissa N., Melinda Smale, and R. P. Singh. 2003. “The Economic Impact of Productivity Maintenance Research: Breeding for Leaf Rust Resistance in Modern Wheat⋆.” Agricultural Economics 29 (3): 253–63.

Mukherjee, Deep, Boris E. Bravo-Ureta, and Albert De Vries. 2013. “Dairy Productivity and Climatic Conditions: Econometric Evidence from South-Eastern United States.” Australian Journal of Agricultural and Resource Economics 57 (1): 123–40.

Nalley, Lawton L., Andrew P. Barkley, and Allen M. Featherstone. 2010. “The Genetic and Economic Impact of the CIMMYT Wheat Breeding Program on Local Producers in the Yaqui Valley, Sonora Mexico.” Agricultural Economics 41 (5): 453–62.

Njuki, Eric, Boris E Bravo-Ureta, and Víctor E Cabrera. 2020. “Climatic Effects and Total Factor Productivity: Econometric Evidence for Wisconsin Dairy Farms.” European Review of Agricultural Economics 47 (3): 1276–1301. https://doi.org/10.1093/erae/jbz046.

Olley, Steven, and Ariel Pakes. 1996. “The Dynamics of Productivity in the Telecomunications Equipment Industry.” Econometrica 64: 1263–97.

Pryce, J. E., and R. F. Veerkamp. 2001. “The Incorporation of Fertility Indices in Genetic Improvement Programmes.” BSAS Occasional Publication, 237–50.

Thornton, Philip K. 2010. “Livestock Production: Recent Trends, Future Prospects.” Philosophical Transactions of the Royal Society B: Biological Sciences 365 (1554): 2853–67. https://doi.org/10.1098/rstb.2010.0134.

Townsend, Robert, and Colin Thirtle. 2001. “Is Livestock Research Unproductive? Separating Health Maintenance from Improvement Research.” Agricultural Economics 25 (2–3): 177–89.

VanRaden, P. M. 2004. “Invited Review: Selection on Net Merit to Improve Lifetime Profit.” Journal of Dairy Science 87 (10): 3125–31. https://doi.org/10.3168/jds.S0022-0302(04)73447-5.

Appendix

Principal Component Analysis

In order to construct a production and health index, we apply Principal Component Analysis (PCA) to a large dataset of dairy bull evaluations. Dairy bulls are evaluated three times a year, and in each period their scores are updated. The scores are calculated by the CDCB using a statistical formula that calculates the contribution of a bull to the performance of its offspring. The offspring data is from DHIA, meaning as bulls are used more widely there is more and more data available to calculate the scores. These scores are used by genetics companies to price bulls and by farmers to decide which bulls to purchase.

For our purposes, we use five traits: milk yield, fat yield, protein yield, somatic cell score, and daughter pregnancy rate. The first three are production traits whereas the last two are related to health. The first two principal components explained about 85% of variation in this data and capture the fact the production traits tend to be correlated with production traits and health traits with health traits. Table A1 shows the loading scores for each component, coefficients which determine the hyperplane through the trait space. The larger these coefficients are the stronger the association between this component and the trait is.

Table A1: Principal Component Analysis Loading Scores

Loading Scores	Component 1	Component 2
Milk Yield	-0.567	-0.029
Fat Yield	-0.552	0.055
Protein Yield	-0.580	0.009
Somatic Cell Score[1]	-0.118	0.714
Daughter Pregnancy Rate	0.149	0.697
Percentage of Variation Explained	58.06	27.25
Observations	1,055,869	

[1] We use the negative of the somatic cell score in order to maintain the interpretation that higher values mean higher health.

The first component has large and negative loading scores for the three production traits whereas the second component has large and positive loading scores for the health components. This suggests that the first component can be a proxy for production traits and the second component can be a proxy for health traits. Since the production traits have negative load scores for the first component, we use the negative of the first component as our production index so that higher values are interpreted as more production (like the health index).