

BREEDING AND TECHNOLOGICAL INNOVATIONS IN DAIRY AND BEEF CATTLE
MARKETS

BY

VICTOR FUNES-LEAL

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Agricultural and Applied Economics
in the Graduate College of the
University of Illinois Urbana-Champaign, 2024

Urbana, Illinois

Doctoral Committee:

Assistant Professor Jared Hutchins, Chair
Associate Professor Mary Arends-Kuenning
Professor Peter Goldsmith
Professor Matthew Wheeler

ABSTRACT

Since humans domesticated cattle (*Bos Taurus*) around 10,000 years ago, we have taken advantage of all the byproducts of such animals (beef, milk, cheese, hides, etc.) However, until relatively recently, we did not possess adequate information to analyze decisions such as breeding, crossbreeding, or animal replacement. Until relatively recently, every dairy and beef cattle ranch operated in relative isolation, resorting to diverse methods to improve their gene pool, such as organizing in cooperatives (Hutchins and Hueth, 2023) to test the quality of their milk and acquire bulls to impregnate their cows.

Starting in the United States during the 1930s, The National Cooperative Dairy Herd Improvement Program started testing bulls for their genetic traits to provide better information to guide dairy farmer's breeding decisions with the help of breed associations and the United States Department of Agriculture (USDA). These tests called Large Scale Genetic Evaluations (Or simply "evaluations"), consist of measuring observed quantitative and qualitative traits from cattle to ascertain which portion of the realized ("expressed") value can be attributable to genetic factors, and particularly, the male line since none of the economic relevant traits of dairy cattle (milk, butterfat, and protein yields) can be directly observed on males. These evaluations are carried out continuously but published three times a year (April, August, and November), and their information is used for marketing dairy bulls.

Two of the chapters in this dissertation use Evaluation data from the National Association of Animal Breeders (NAAB), certified semen services, which have been routinely updated and published online since 2000 and consist of thousands of genetic trait data from every bull market in the United States. At the time of writing this dissertation, very few articles in Applied Economics have used such evaluation data; one notable exception is Whitt et al. (2019).

In this dissertation, I investigate three distinct aspects of the Economics of livestock genetics. In my first chapter, I analyze the consequences of a new technological development, Genomic Testing, and how it boosted the productivity of dairy cows in North America, but also the extent to which it led to negative consequences through increased inbreeding.

The second chapter deals with a different topic: the introduction of crossbreeds in Argentina as a form of climate adaptation driven by land use substitution from cattle grazing into soybeans. We show that such a displacement out of temperate areas has fueled a demand for specific traits, particularly more light-colored animals with lower birth weights but increased ability to gain weight.

Finally, the third chapter deals with the effects of winning the "Premier Sire" category in the World Dairy Expo on the price of bull genetics as a way to incorporate pedigree into a hedonic pricing framework. I show that the effect of winning the World Dairy Expo "trickles down" into the animal's family, but the sign of the effect is different from the expected, since winners are already well-known sires.

To my Wife and Mother.

ACKNOWLEDGMENTS

I want to acknowledge the support from my Advisor, Jared Hutchins, who introduced me to a whole new strand of literature that I had completely overlooked and had enough patience to guide me through numerous pathways during the last three and a half years of my PhD. I would also like to thank my committee members - Mary Arends-Kuenning, Matthew Wheeler, and Peter Goldsmith - for their insightful comments on my dissertation proposal and defense and their valuable feedback on my writing.

I would also like to acknowledge the support of my fellow graduate students, who provided valuable feedback throughout my years as a PhD student: Farzana Shirin, Christian Valencia, Noelia Romero, and my Job Market Cohort. I would also like to thank my two coauthors of the Economic Value of Brangus bull traits article, Nicolás Gatti and Ignacio Amaro from Argentina's Instituto Nacional de Tecnología Agropecuaria (INTA). Finally, all my friends who supported me when I first arrived at Champaign: Brian Feld, Natalia Canted, Anahid Bauer, and Felipe Díaz.

This dissertation received support from the Office of International Programs at the University of Illinois' College of Agricultural, Consumer, and Environmental Sciences through the Global Food Security Fellowship, which allowed me to fund part of my research in Argentina.

This dissertation received feedback and input from Seminars at the Department of Agricultural and Consumer Economics, particularly the Food, Agriculture, and Commodities (FACS) and the International Policy and Development (IPAD). I would also like to thank participants at the 2022 and 2023 annual American Agricultural and Applied Economics Association (AAEA) meetings, the 2023 annual International Consortium on Applied Bioeconomy Research, and the 2002 and 2024 NC1034 Research Group meetings. Finally, I would like to thank Roger Shanks and John B. Cole, who helped me improve my understanding of the Holstein Genetics Market.

Finally, I would like to acknowledge the support from my family, especially my wife, since none of this would have been possible without her support, particularly during the stressful days of the COVID-19 pandemic.

TABLE OF CONTENTS

Chapter 1	INTRODUCTION	1
Chapter 2	INFORMATION PROVISION AND NETWORK EXTERNALITIES: THE IMPACT OF GENOMIC TESTING ON THE DAIRY INDUSTRY	3
2.1	Introduction	3
2.2	Conceptual Framework	6
2.3	Data Description	8
2.4	Empirical Framework	11
2.5	Results	15
2.6	Robustness Checks	19
2.6.1	Total number of descendants	19
2.6.2	Quarter of birth effects	20
2.6.3	Date of birth of “superstars”	22
2.7	Discussion	22
Chapter 3	ECONOMIC VALUE OF BRANGUS CATTLE TRAITS IN ARGENTINA	24
3.1	Introduction	24
3.2	Beef Cattle Genetics Market in Argentina	25
3.2.1	Supply of Beef Cattle Genetics	25
3.2.2	Demand for Cattle Genetics	27
3.3	Theoretical Framework	28
3.3.1	Hedonic pricing model	28
3.3.2	Input Characteristics Model	29
3.3.3	Empirical Specification	30
3.4	Data	31
3.5	Results	32
3.6	Discussion	34
3.7	Conclusion	35
Chapter 4	THE PRICE OF FAME: THE EFFECT OF AWARD-WINNING PEDIGREES ON THE PRICE OF LIVESTOCK GENETICS	37
4.1	Introduction	37
4.2	Background and Conceptual Framework	38
4.2.1	The World Dairy Expo	38
4.2.2	Measurement of genetic traits	39
4.2.3	Input Characteristics Model	40
4.3	Data and Empirical Approach	42
4.3.1	Data description	42
4.3.2	Empirical framework	43
4.4	Results	44
4.4.1	Multiple-period differences-in-differences	46

4.4.2	Aggregate differences-in-differences	46
4.5	Conclusion	47
APPENDIX A: Appendix to Chapter 2		49
APPENDIX B: Appendix to Chapter 3		55
APPENDIX C: Appendix to Chapter 4		57
References		59

CHAPTER 1

INTRODUCTION

The goal of this dissertation is to use bull-level data to study different aspects of the cattle genetics market to explore topics such as: technology adoption, climate adaptation and the impacts of asymmetric information. This dissertation is a unique contribution to the applied economics literature by leveraging data from animal evaluations to answer different research questions using econometric methods.

In Chapter 1, I estimate the impact of a new testing technology, genomic selection, on the market for dairy cattle genetics in the United States. The genomic selection allows cattle breeders to test new cattle as soon as they are born for the presence of desirable, productive traits and more than double their pace of genetic improvement in milk production. At the same time, the rate of inbreeding, the extent to which dairy bulls are related to one another, greatly increased after the introduction of genomic testing.

Using a data set of all bulls marketed in the US from 2000 to 2020, we examine whether inbreeding rates increased due to increased competition for popular dairy cattle families in the face of greater market competition. I use a difference-in-difference and matching approach and find that inbreeding increased greatly in already popular families after the introduction of genomic testing relative to less popular families with similar production and health traits. The estimations show that the increased inbreeding rate of American bulls caused the entire industry to lose between 4 and 7.4 billion dollars from 2011 to 2019, accounting for between 4.4% to 8.1% of total revenue for the American Dairy Industry.

Chapter 2, deals with the consequences of a boom in demand for agricultural commodities that started in the early 2000s that displaced cattle ranching out of the most productive areas of the Pampas' prairie. The crossbreeds between Angus and Hereford with Brahman, i.e., Brangus and Braford, have been successfully adopted across Argentina. However, little is known about the specific bulls' traits that drive the demand for genetic selection of cattle outside the Pampas. Obtaining the economic value of traits would help to identify the demand for adapting livestock production to different ecosystems while preserving the meat quality of Angus and Hereford cattle. We estimated hedonic price models using Brangus bull sales data from two cattle breeding ranches in the north of Cordoba province.

I find that cattle ranchers prefer observed traits such as weight, coat color, and age, while genetic indicators such as Expected Progeny Differences (EPDs) have secondary importance. We argue that stronger preferences for red-coated bulls, as opposed to black-coated bulls, could be associated with the demand for reducing heat stress; the weak association between EPDs and prices may be related to unobservable variables such as ranchers' characteristics and that the value of genetics is implicit in the studs' reputation.

In chapter 3, I estimate the impact of winning an award at a cattle show on the price of cattle genetics and that of their relatives. Dairy farmers choose from various dairy bulls that possess genetic traits that reflect their productivity, resiliency to disease, and physical characteristics. Another relevant attribute of dairy bulls is their pedigree prestige, which dairy farmers can use as a proxy for quality. We test the importance of

pedigree prestige in determining dairy bull prices by examining the winners and runners-up of “Premier Sire” at the annual World Dairy Expo. Using an event study framework, we find that bulls that win Premiere Sire experience no significant price premium relative to other runner-ups. Further research is needed to determine whether pedigree prestige is a significant determinant of price.

CHAPTER 2

INFORMATION PROVISION AND NETWORK EXTERNALITIES: THE IMPACT OF GENOMIC TESTING ON THE DAIRY INDUSTRY

2.1 Introduction

The United States is the world's largest supplier of cattle genetics. In 2022, the United States alone exported \$295 million of bovine genetics, representing 47% of the world's exports in value terms.

This means that the continued health and quality of the US cattle genetics gene pool are important to the US livestock industry and livestock industries worldwide.

For about six decades, milk production per cow has consistently increased in the US herd, and about half of this growth has been attributed to improvements in cattle genetics. Dairy breeders improve cow milk yield by identifying which bulls have the most productive female offspring and then continuing to breed off of these lines.

In previous years, the pace of this genetic improvement was limited by biological factors. Determining whether a bull had good milk production ability took at least three years since the bull had to have offspring born from their use on dairy farms and those offspring would need to take two years to grow to production age.

In the past ten years, an innovation in testing technology has revolutionized the industry: genomic testing. Introduced in 2008, genomic testing allows breeders to link each animal's genes and productive outcomes such as milk yield and fat yield. The practical upshot is that determining a dairy bull's quality can now be done as soon as it is born instead of in three years.

There are two important impacts of genomic testing in the market for dairy bulls. First, the pace of genetic improvement has more than doubled in dairy cattle since breeders can find productive bulls at much younger ages and breed new generations even quicker. Second, dairy companies have more than doubled the number of bulls for sale on the market.

However, the introduction of genomic testing has coincided with a less desirable outcome: an increase in inbreeding. The inbreeding rate measures the relatedness of a dairy bull to the rest of the dairy cattle population. Inbreeding is an example of a negative externality in the cattle genetics industry since higher inbreeding rates expose the industry to more disease risk and depressed yields in future generations, costs not borne directly by the breeder in the short run.

Since dairy breeders continually breed from each other's genetic stock, the breeding decisions of one breeder will affect all other breeders through the shared public resource that is the genetic pool available in the US. In one case, breeding excessively from a famous dairy bull "Pawnee Farm Arlinda Chief" led to more

than 500,000 spontaneous abortions and nearly 420 million dollars in losses because of a deleterious mutation in a gene that was passed to his offspring (Adams et al., 2016). Given the US’s role in the global cattle genetics market, damaging the diversity of the dairy cattle gene pool could also damage livestock markets worldwide.

In this paper, we test whether the introduction of genomic testing caused an increase in competition for already popular dairy bull genetic lines and, thus, an increase in the inbreeding rate. Our analysis is complicated because popular dairy bull lines are usually superior to other lines in terms of production and health traits. To get a clean comparison, we construct a counterfactual line that is comparable to popular lines in terms of productivity using propensity score matching.

Using a difference-in-difference approach, we then examine inbreeding rates in popular families relative to our counterfactual line before and after the introduction of genomic testing in 2010. If genomic testing increased competition for quality lines independent of previous popularity, we expect that inbreeding rates for these two groups will be the same both before and after its introduction.

We find that, after the introduction of genomic testing, popular dairy bull lines became more inbred than our counterfactual. Relative to our control group, the inbreeding rate of the most popular lines increased 1-2.8 percentage points after 2010. Before genomic testing, there was no statistically significant difference in inbreeding rates across our groups. Using an estimate of the costs of inbreeding from the Holstein breed association, we estimate that this increase in inbreeding led to \$3.6-6.7 billion in losses to the industry from 2011-2019.

Since dairy breeders are incentivized to inbreed for short-run gains even at the expense of their genetic pool, this indicates a role for regulation to penalize the industry for inbreeding in the short run to prevent damages to the industry in the long run.

This article fills a critical gap in the literature on livestock markets and the behavior of animal breeders. Few economics papers analyze the strategic actions of animal breeders and even fewer analyze how these actions interact with inbreeding. Early modeling of the animal breeder’s decision dates back to studies such as Ladd and Gibson (1978), Kerr (1984), and Melton et al. (1994) which use hedonic regression and linear programming to determine the implicit market value of different livestock traits. Richards and Jeffrey (1996) and Schroeder et al. (1992) apply this method to dairy bulls to determine the economic value of each dairy bull trait implied by the results of a hedonic regression. Thus far, the economics literature has not given sufficient attention to the network nature of livestock genetics markets nor appreciated how the decisions of animal breeders may lead to worse outcomes for the whole sector.

Specifically, the few economics papers on the dairy genetics market have so far missed an important development in the US dairy herd: the rise of inbreeding. Inbreeding in this context is an example of a negative “network externality.” Network externalities were defined by (Katz and Shapiro, 1985) as usually positive externalities that result from increasing a product’s user base. For example, a technology may become more valuable to a user as the user base grows and the user network has more connections. Thus, the more dense the network is, the greater the potentially positive externality.

Similarly, inbreeding is a network measure of the extent to which a dairy bull is related to all other dairy bulls in a population. As the bulls become more related, the network becomes denser and the average inbreeding rate of the population goes up. Breeders can be incentivized in the short run to increase inbreeding by breeding cattle from lines that are already popular which may boost genetic improvement and capture more market share. However, instead of causing a positive externality, inbreeding causes a negative externality by exposing the network to a greater chance of harmful recessive traits and weakening the genetic pool’s

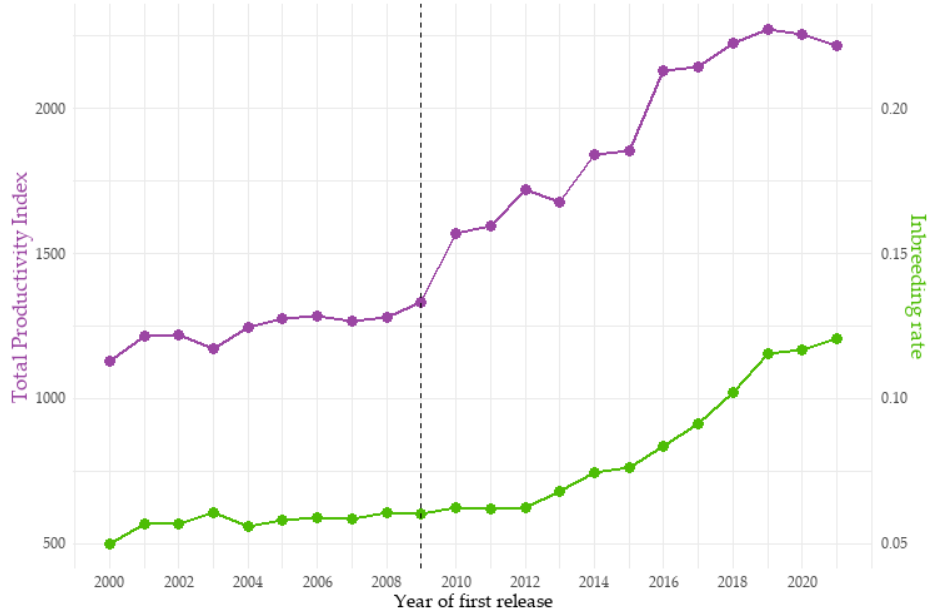


Figure 2.1. Dairy Bull Productivity versus Inbreeding Rate

resiliency. Animal breeders do not directly pay the cost of inbreeding in the short run, which can lead to a rate of inbreeding, which is ultimately inefficient for the market.

We use this framework to understand how genomic testing may have increased inbreeding rates in the US dairy cattle population. Before its introduction, it was theorized that genomic testing would actually decrease inbreeding by allowing breeders to find productive dairy bulls outside known genetic lines more easily. Instead, the average inbreeding coefficient of dairy bulls increased along with dairy bull productivity. Figure 2.1 shows that, along with an increase in the total productivity index (an index measure of dairy cow productivity), there has been an increase in the average inbreeding rate of the US dairy bull population. Animal scientists have theorized that this increase in inbreeding is due to breeders using already popular genetics lines to gain a competitive edge (Cole, 2019). We provide evidence of this theory by comparing the inbreeding rates of groups of bulls that are similar in genetic traits but different in popularity before and after genomic testing. If genomic testing led to an increase in the discovery of less popular lines, we should expect no difference in the inbreeding rates of these groups since animal breeders would have selected them only on the basis of their genetic traits. Instead, we find that genetic lines already popular before genomic testing experienced an increase in the average inbreeding rate, leading to significant economic damage to the dairy sector.

The paper is structured as follows: Section 2.2 presents the relevant background details and a detailed causal mechanism to explain why introducing genomic selection led to higher inbreeding rates in subsequent generations of dairy bulls. Section 2.3 explains the data sources and shows how some descriptive statistics can help us delineate the relationship between our observed variables. Section 2.4 outlines the identification strategy and methodological framework. Finally, in section 2.7, we explain the results and discuss their implications for the future of the US dairy industry.

2.2 Conceptual Framework

Selective breeding has been practiced since prehistoric times; however, the discovery of inheritance mechanisms and genetics in the 19th and 20th centuries allowed breeders to identify the link between desired traits in animals and their inheritance mechanisms (Lush, 1937). From an economic standpoint, an animal breeder supplies a product (animal genetics) with an inseparable bundle of traits embodied in a single animal. To produce such an animal, the breeder combines a set of inputs that can be split into two groups: a series of inputs exogenously supplied by the producer (labor, feed, capital, veterinary services) and a series of inputs embodied in the animal (traits.) Those inputs are combined on a production function subject to the prices of exogenous inputs (observed) and genetic inputs, whose prices are implicit but can be estimated based on the price paid for a bull’s genetics using the hedonic approach (Schroeder et al., 1992; Sy et al., 1997).

Melton et al. (1979, 1994) add an additional element: breeders are paid based on the (positive) deviation of the trait levels of their bulls relative to the population average. Breeders compete to maximize profits; to do so, they will attempt to release bulls into the market with the highest possible trait values. Dairy farmers demand genetics to improve their productivity in the next period by choosing the bulls with the best characteristics. However, the traits that constitute an animal are marketed as a “bundle” of traits embedded in each animal. Dairy farmers are only interested in the productivity of their cows because their profits depend on it; they are not directly interested in how inbred their cows are; therefore, inbreeding rates cannot be priced directly as an additional cost to dairy farmers.

Inbreeding is the mating of individuals more closely related to each other than the average for the population (Bourdon, 2000, p. 337). Inbreeding has two consequences: first, it increases the likelihood that an animal’s offspring will perform similarly to its ancestors; for example, if a particularly tall bull is mated with a related cow, its offspring will likely be taller than average because the genes that carry such trait are likely to be present in both parents. Second, it favors the expression of deleterious recessive alleles; since an animal has two copies of a recessive gene from each parent, it is more likely to express any such traits than an outbred animal.

The increment in the rate of change in hereditary traits caused by genomic selection can be decomposed into a set of components: heritability, selection intensity, and generation interval (Kerr, 1984). Equation 2.1 is called the “breeder’s equation”:

$$\text{genetic improvement} = \frac{\text{heritability} \times \text{selection intensity} \times \text{trait variance}}{\text{generation interval}} \quad (2.1)$$

Heritability is a population parameter specific to each trait that measures the strength of the relationship between the genetic traits of parents and their offspring, and so is the standard deviation of the observed trait values. **Selection intensity** is the differential in average trait values between selected animals and the general population at a given time. Selection intensity measures how “choosy” breeders are in deciding which individuals are selected; more intensity of selection means that only the best animals (according to some criterion) are selected to have offspring. If animals were randomly selected, those magnitudes would not differ. The generation interval is the time required to replace one generation with the next: the shorter the interval, the faster the genetic change. Genetic improvement is driven by two forces: selection intensity and generation interval. An increase in selectivity or a decrease in the generation interval increases genetic gains per generation.

Before 2009, it took one to five years for a bull to father a large enough number of daughters to test them for the presence of genetic traits such as milk, protein, or fat yields. With genomic testing, an animal can be

tested as soon as it is born for the presence of specific traits in its genome. Consequently, the generation interval, the average age of parents when offspring are born (G. R. Wiggans and Carrillo, 2022), has decreased from around five and a half years to less than two years, close to the biological minimum of one year.

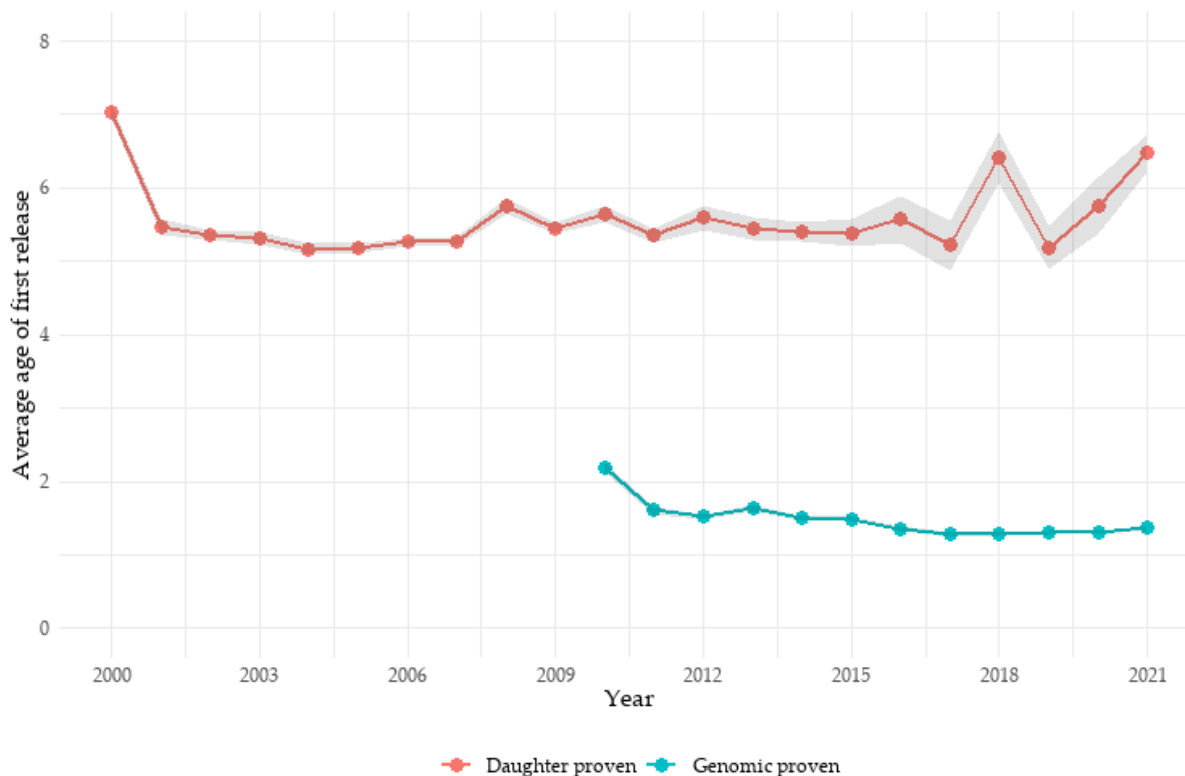


Figure 2.2. Average age at first proof for Holstein bulls
Source: NAAB

The most important consequence of a shorter generation interval is an increase in genetic progress. Figure 2.2 also shows that the confidence interval of our estimation of the mean age has widened; we interpret that as evidence of a decrease in the number of daughter-proven bulls.

A key element of our analysis is the existence of what we call “supersires”, high-prestige bulls who became the founders of lines (families). The more genetics companies breed the same lines, the more newborn bulls will be related to a specific ancestor. We define a bull as a supersire in terms of the size of its offspring; since we have records of all male descendants from each sire that have been released on the dairy genetic market, we can then identify which bulls have had the largest offspring; more precisely, a bull born between 1997 and 2007 is a supersire if the size of his offspring is on the 95th percentile of the offspring distribution.

We test the hypothesis that genomic selection caused higher inbreeding rates by constructing a counterfactual line of bulls that possessed similar traits but were less popular at the introduction of genomics. We identify their descendants and respective inbreeding rates; they increased disproportionately more for the descendants of supersires. A likely explanation for this process is that genetics companies did not invest in the most productive bulls but instead in the most famous lines in an attempt to gain market share. The more companies breed into those lines, the higher the inbreeding rate of the next generation of bulls. This leads to a deterioration of the expected traits of future generations of bulls and, thus, to significant losses for

the dairy industry in terms of reduced productivity.

2.3 Data Description

Improved genetics via the selection of bulls (sires) and cows (dams) is a crucial driver for productivity improvements in the dairy industry. Milk production in the United States has tripled in the last forty years, and over half of this increase is solely due to genetics (<https://uscdcb.com/impact/>). The United States is also the world's largest exporter of bovine genetics, with a share of 46.4% of the total value of exports in 2019. Every improvement in genetics in the United States will be quickly transmitted to the rest of the world dairy industry.

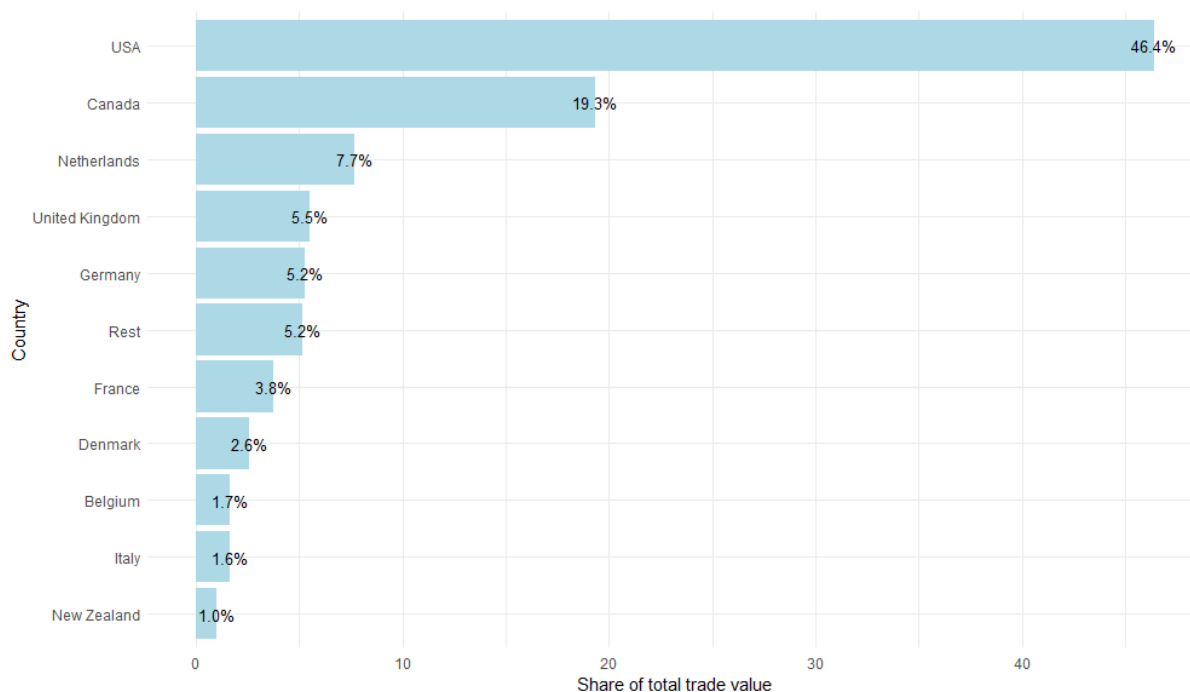


Figure 2.3. Share of trade Value of Bovine Semen exports by country (2019)

Source: UN Comtrade database

Traditional selection techniques relied on information from pedigrees and observed trait values to assess an animal's fitness. Genomic selection uses data from an animal's genome to accurately predict breeding values, particularly in cases where traits are expressed only in females. In such case, it was necessary to measure the traits of a large number of daughters from the same bull; for that to happen, these cows must have reached the age of sexual maturity (one year); hence, the entire process could take three to four years until enough information was gathered.

Instead, genomic selection uses the bull's DNA to detect gene combinations significantly different from the expected frequency if they were randomly associated. This process detects the presence of certain genetic traits that are highly correlated with specific gene combinations. It can be undertaken as soon as the calf is born, so by the time the animal has reached sexual maturity (one year old), it can be marketed immediately.

Figure 2.5 shows that genomic selection increased the number of animals released into the market since there is intense competition for supplying the most productive bulls; the relative value of a line is now higher

than ever. Since dairy farmers have a more extensive choice set of bulls to buy genetics from, lines act similar to “brands,” as they reduce the quality uncertainty from a product. Finally, this “loyalty” to established lines leads to breeders supplying more bulls from such families than from lesser-known lines that have similar or better traits.

Brand loyalty arises in this context due to the nature of animal genetics, dairy farmers demand a particular set of genetic traits related to the productivity and health of their cows, but they cannot be sold separately; they have to be acquired as a “bundle” at a single price (Melton et al., 1994; “Prices and Demands for Input Characteristics”, 1976) that is a linear function of the entire set of genetic traits, which makes a comparison between any two or more animals extremely difficult. This implies that there are very high search costs to farmers when deciding which animal to buy genetics from, Stigler (1961) and Farley (1964) argue that it is reasonable for an economic agent to seek further information on a single purchase as long as the expected reduction in price per unit as a result of additional search times the quantity purchased is equal to the marginal cost from searching for lower prices (or higher qualities.)

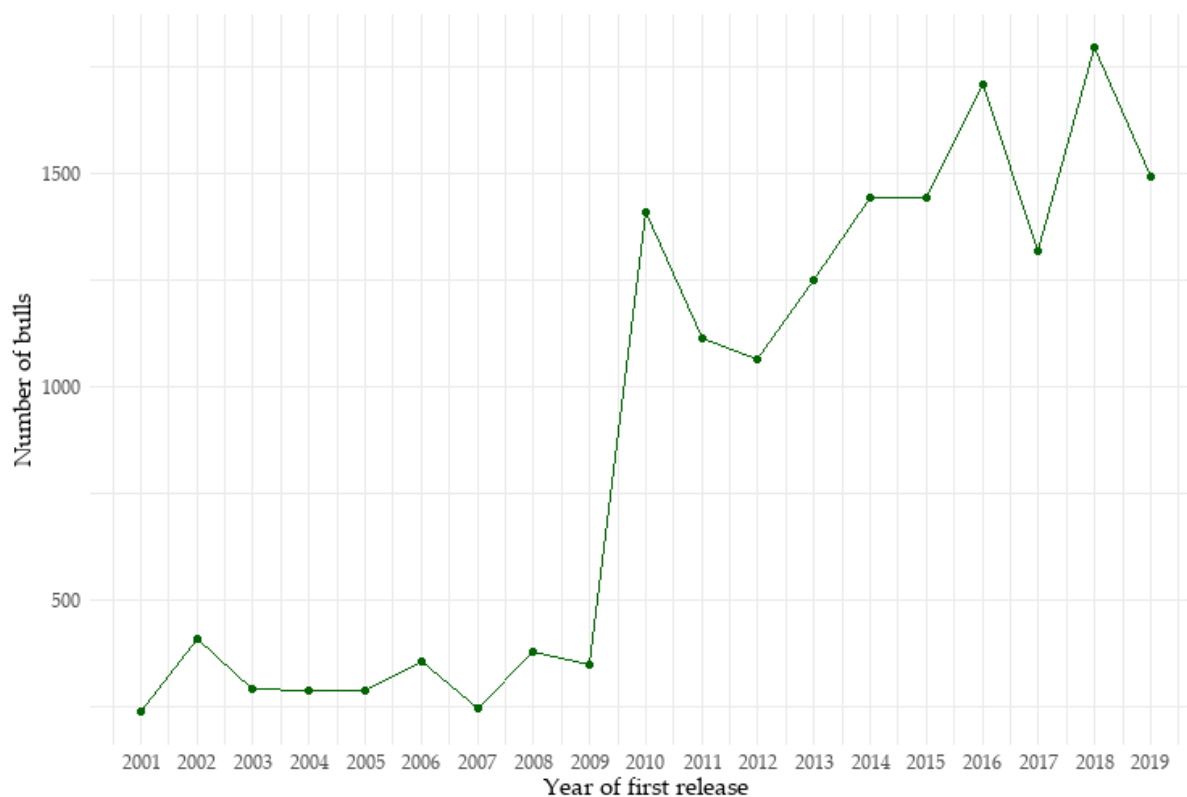


Figure 2.4. Total number of new bulls released in the market **Source:** NAAB

Genomic selection increased the number of new bulls released into the market, as shown in Figure 2.4. Consequently, search costs for dairy farmers increased as well, thus promoting greater brand loyalty for “superstar” lines, that is, bulls descended from a notorious ancestor (e.g., bulls that fathered many daughters). The ultimate consequence of this process is an increase in inbreeding for the entire population of dairy bulls; the more related bulls are, the more likely their offspring is to preserve their genetic traits. Higher inbreeding levels lead to a higher likelihood of observing **inbreeding depression**, a decrease in the performance of inbred animals due to lower fertility, higher incidence of rare diseases, and reduced longevity.

The National Dairy Herd Improvement Program (NDHIP) collects, manages, and analyzes data from American dairy cattle (Hutchins and Hueth, 2023). The information it collects assesses the industry’s productivity and estimates the paternal contribution to these traits. The raw data collected from farms is used to estimate a statistical model to predict the genetic contribution of an animal as a sire (or dam) after controlling for environmental variables such as herd, age, pedigree, and season effects. The predicted values for each trait from this model are commonly called the **predicted transmitting ability (PTA)** of a trait.

Animal breeders and genetics companies compete to offer an animal with the most significant possible improvement over the cohort’s average. As a consequence, firms will try to produce bulls with the highest productivity traits; similarly, this will lead to higher rates of inbreeding since the best way to ensure high productivity is through linebreeding, that is, by building a pedigree that has a relatively stable degree of relationship relative to a supersire. We call “large” companies those with the most significant average number of bulls across the entire 2000-2020 period in the North American genetics market. Similarly, a firm is called “preexisting” if it participated in the bull genetics markets before 2009; otherwise, they are considered “new.”

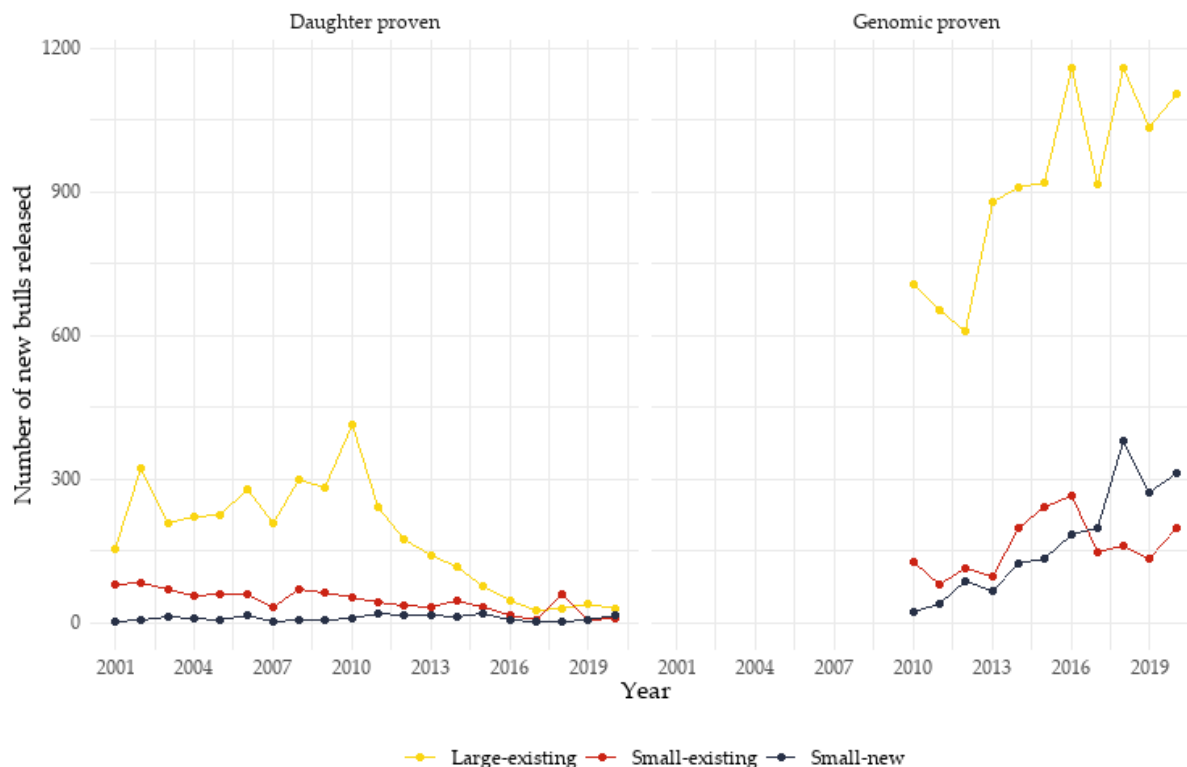


Figure 2.5. Number of genotyped bulls by firm size
Source: NAAB

Genomic testing allows genetic companies to detect suitable bulls earlier; thus, they can release those animals into the market as soon as they reach sexual maturity. Figure 2.5 shows the total number of new bulls released in the market by year and type of firm.

In the cattle genetics industry, genomic testing has been adopted remarkably fast. Figure 2.6 shows how fast genomic testing was adopted in our sample. The high starting costs for genotyping help explain this since large companies were the ones that could initially afford it. Still, more companies have genotyped their

calves since genomic selection was released. Larger firms genotyped a more significant percentage of their newborn calves in 2010, and the share has grown steadily, while smaller firms followed them at a slower rate.

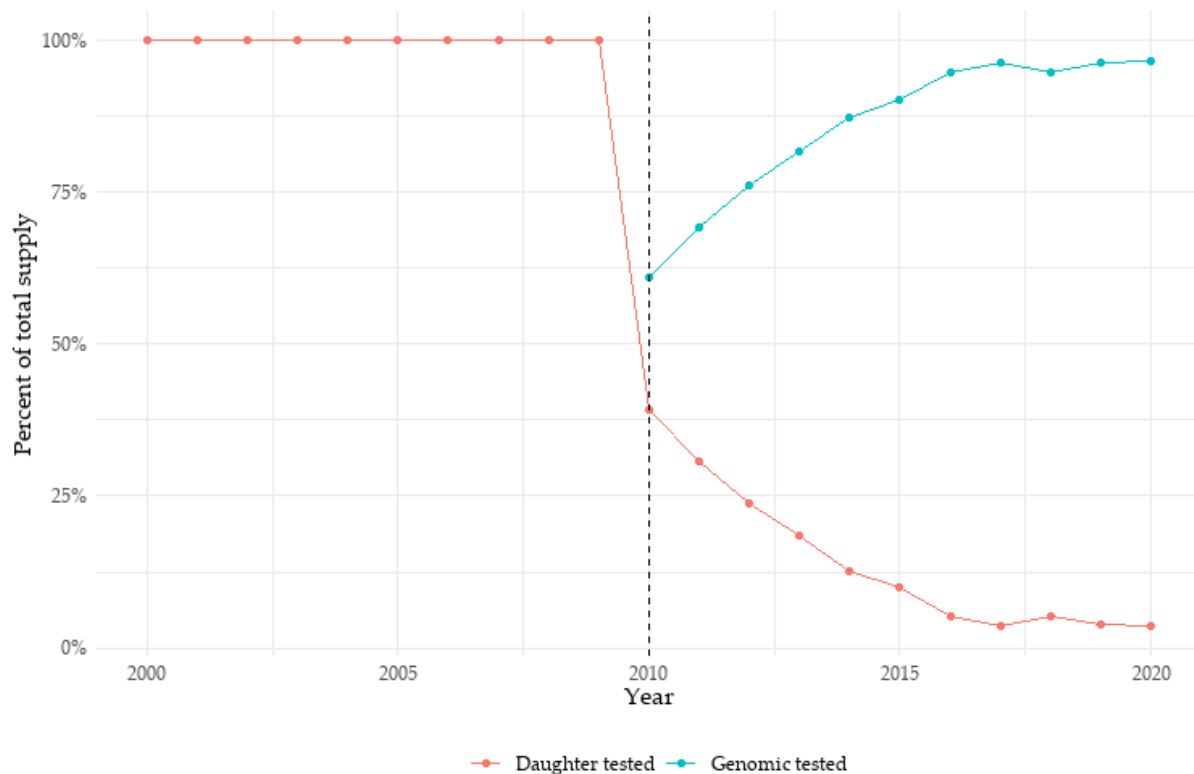


Figure 2.6. Share of each status of bull by year

Starting in 2010, large firms started genotyping their bulls, and consequently, the number of new bulls increased threefold in that year, when all bulls aged 1 to 4 were released simultaneously. The staggered adoption of genomic selection has further consolidated the market and led to more competition for superstar lines. However, the more inbred selected animals are, by definition, also increases the inbreeding of an average animal in any cohort, thus making inbreeding depression more likely within a herd. Figure 2.5 shows the total number of new bulls released in the market by year and type of firm. We call “large” companies with the most significant share of the North American genetics market. Similarly, a firm is called “preexisting” if it participated in the market of bull genetics before 2009; otherwise, they are considered “new.”

2.4 Empirical Framework

We want to test the hypothesis that the introduction of genomic selection led to higher inbreeding rates because of increased demand for well-known pedigrees fueled by an increment in dairy cattle productivity. We want to know the causal effects of introducing this technology; ideally, we need a treatment group affected by genomic selection and a treatment group that was not.

We must identify families whose founders had similar trait levels as superstars but had less “momentum”; that is, a smaller number offspring. This is a comparable group because breeders could have invested in them but did not because they were not demanded as highly as superstars. We can track these lines over time.

If their inbreeding rates did not differ substantially from those of superstars, we could argue that genomic selection allowed breeders to identify new, less-inbred bulls independent of their ancestry.

Figure 2.6 shows that the pace of adoption was very high from the beginning; this makes it unlikely to use non-genomic proven bulls as a counterfactual, mainly because we don't know why breeders choose not to genotype certain bulls after 2009. Similarly, we have very few foreign-born bulls to use as a control group; we also don't know the selection criteria to import genetics from those animals. Both alternative control groups suffer from an additional issue: they will have a significantly smaller offspring, thus leaving us with little statistical power to detect any relevant treatment effect. To overcome these issues, we use propensity score matching (J. A. Smith and Todd, 2005) to define a control group that resembles the treatment in all observable characteristics (except for our outcome: inbreeding rates.)

We define a **line** as the set of direct (male) descendants from a particular bull, which, in turn, can be described in terms of his number of male descendants; a bull is a superstar if his number of sons is in the 95th percentile of the distribution of offspring. Breeders are under high competitive pressure to supply bulls with the highest PTAs. However, the best way to do so is by linebreeding a family of bulls so that the relative distance to an ancestor remains constant across generations. Steyn et al. (2022) and Steyn et al. (2023) show that younger sires capture a more significant share of present-day population genetic variation; hence, genotypes of older sires can be discarded from (genomic) evaluations. As generations begin to accumulate, the contribution of a particular sire decreases (on average) exponentially, so we can disregard the impact of distant ancestors.

In the sample, we include all animals whose sires were born between 1998 and 2004, and we call those sires the “founders” of each line. Our definition depends critically on when the founders of such lines were born; if we choose an earlier date, we cannot find a statistically significant effect due to the high relatedness of all bulls. The reason behind this is that bulls are advertised only with their sire and grandsire; any ancestry beyond that is not taken into account by dairy farmers. If we choose a later date, both treatment and control groups will have little difference in inbreeding rates because there is not enough time for the control group lines to develop. Figure A1 in the appendix summarizes the sample selection mechanism; the treatment and control groups are comprised of the descendants of two groups: supersires (treatment) and matched sires (control).

After identifying the sires, we can recursively build their pedigrees by matching bulls with their sires until we find the latest animal in the sample. We want to measure the causal effect of introducing genomic testing in 2009 on the average inbreeding in extant lines; the first problem we must tackle is the nature of a valid comparison group. The introduction of genomic testing caused animal breeders to supply more animals from well-known lines, that is, descended from superstar sires; our control group must be the lines descended from other non-superstar sires.

First, we need to match the founders of such lines with other sires that did not father as many sons but had comparable genetic traits. Second, we find all descendants from both groups and compare the evolution of their inbreeding rates through time. The idea behind this procedure is that when those bulls were in the market, any reasonably similar sire could have been chosen but was not, and thus, can be used to construct a counterfactual line.

For every family, we must construct a counterfactual family with new bulls being born every year, each with an inbreeding rate that we can average across cohorts. Then, the **treatment group** are all lines of animals comprised of direct descendants of superstar sires. Similarly, the **control group** includes descendants from comparable animals matched to superstar bulls by their genetic trait levels.

Our next objective is to identify the counterfactual lines by matching superstar sires with alternative lines. To do so, we use a propensity score matching estimator, let $p(X) = \Pr(T_j = 1|\mathbf{X})$ where T is a dummy equal to 1 if sire j 's offspring is on the 95th percentile of the distribution and X_j be a vector comprised of the following:

- Predicted Transmitting Ability for production traits (milk, fat, and protein), health traits (mastitis, metritis, productive life), and reproductive traits (daughter pregnancy rate, effective calving, gestation length).
- Dummies for the absence of certain recessive genetic disorders (Foundation, 2021) such as *brachyspina*, cholesterol deficiency, and mule foot.
- Dummies for the presence of genes associated with specific qualitative traits (polled, red coat color, complex vertebral malformation).
- Dummies for the presence of haplotypes (specific sequences of DNA at different locations in the chromosome) that affect fertility (Foundation, 2018) when present in homozygous form (likely to be expressed when two carrier animals are mated).

We then estimate the propensity score (Cameron and Trivedi, 2005, p. 873), the probability of being a superstar given the level of PTAs and the presence of the entire set of variables listed above. The validity of our method relies on a series of assumptions (Heckman et al., 1997):

Assumption 1 (Stable Unit Treatment Value Assumption): $(inb^{treatment}, inb^{control}) \perp\!\!\!\perp T/\mathbf{X}$.

The average inbreeding rates of any two lines do not vary with the treatment assigned to other lines. For each unit, there are no different forms of each treatment level (Imbens and Rubin, 2015), conditional on a set of covariates. This assumption implies that conditional on $\mathbf{X}$, the distribution of inbreeding rates for both groups will be the same for all founders, but it must differ for their offspring. Figure A2, in the Appendix, shows the extent to which the inbreeding rates distribution of our matched sample resembles that of the treatment group. Unlike unmatched sires, their inbreeding distribution has a similar mean and variance conditional on the set of genetic traits.

Assumption 2: $0 < \Pr(T_j = 1|\mathbf{X}) < 1$ for all $\mathbf{X}$

This assumption implies that we can define a propensity score for every value in the treatment group. Assumptions 1 and 2 together are the “strong ignorability” assumption in Rosenbaum and Rubin (1983), and they are necessary in conjunction to identify causal parameters. The set of matching variables above exhausts the entire set of possible selection criteria for breeders; no unobserved variable can influence mating decisions. Figure A3 shows the degree of overlapping between superstars and alternative sires; since there is a considerable difference in the size of the groups, superstars make up only a tiny fraction of bulls; however, the overlapping is good in the range 0.05 – 0.50, but there exist a few superstars that few close matches in the sample.

A central part of our argument is that treatment and control groups should have similar inbreeding levels before genomic selection. To test that assumption, in Table A2, we calculate the difference in averages of treatment and control groups for inbreeding rates and a few selected traits. There is no statistically significant difference for most traits between treatment and control groups, so our matching method ensures that the difference between both groups is due to the “prestige” of superstar lines because the bulls in the treatment group had a larger offspring than those in the control group. Each year these bulls will have a new set of

descendants being born, each with its inbreeding rate and genetic traits; if our hypothesis holds, then the inbreeding rates of the descendants of supersires should be significantly higher than those of non-superstar descendants.

We then estimate the differential impact of genomic selection (our treatment) on the two types of lines, those from supersires and those from alternative sires. Model (2) includes leads and lags of the treatment to check to which degree a pre-treatment balance exists in both groups. This is not a test for parallel trends on our differences-in-differences specification (Cunningham, 2021), but it helps us ascertain how comparable both groups are. Table A2 shows that there is no significant pre-treatment difference in both groups and that coefficients from 2010-2011 are not statistically significant either, which coincides with the values from Figure 2.1 where there is a significant increment in inbreeding rates starting in 2012.

We observed the average inbreeding rates of the treatment and matched control groups before and after introducing genomic selection. If we divide the 2005-2018 period into two groups, pre-intervention (2005-2009) and post-intervention (2010-2018), denoted as 1 and 2 and let T_j denotes whether a bull from line j belongs to treatment ($T_j = 1$) or control ($T_j = 0$) groups, then:

$$inb_{jt} = T_j inb_{jt}^{treat} + (1 - T_j) inb_{jt}^{control}$$

Where $j = 1, \dots, J$ are the lines, each comprised of n_j bulls such that $i = 1, \dots, n_j$. Our causal estimand of interest is the Average Treatment on the Treated (ATT) in the post-intervention period:

$$ATT_{post} = E [inb_{post}^{treat} - inb_{post}^{control} | T_j = 1]$$

This quantity measures the average causal effect on the treated lines in the post-treatment period (Roth et al., 2023). To correctly identify causal effects, a series of assumptions must hold (Sun and Abraham, 2021):

Assumption 3 (Parallel trends):

$$E [inb_{post}^{control} - inb_{pre}^{control} | T_j = 1] = E [inb_{post}^{control} - inb_{pre}^{control} | T_j = 0]$$

We assume that the average inbreeding rates for both groups would have evolved in parallel without introducing genomic selection. This assumption is not testable because we cannot observe untreated inbreeding rates after 2009, but it constrains our model specification to a Two-Way Fixed Effects specification we call Model (1):

$$inb_{ikt} = treat_i + (treat_i * post_t)\beta + post_t + \alpha_k + \varepsilon_{ijt} \quad (\text{Model 1})$$

Where inb_{ikt} is the inbreeding rate of bull i born from a line that belongs to either the treatment or control group born in year t , supplied by firm k , β_t is a year of birth fixed effect, and α_j is a firm fixed effect.

Assumption 4 (No anticipatory effects): $inb_{pre}^{control} = inb_{pre}^{treat}$

This assumption ensures that both groups were comparable before the treatment; inbreeding rates of superstar lines did not increase before the introduction of genomic selection; we use matching methods to ensure that this condition holds.

Under Assumptions 3 and 4, we have that:

$$ATT_{post} = E [inb_{post}^{treat} - inb_{post}^{control} | T_j = 1] - E [inb_{pre}^{treat} - inb_{pre}^{control} | T_j = 0]$$

The ATT is identified as the difference-in-differences of observed outcomes in the pre and post-intervention periods. All elements in this equation are observed and, therefore, estimable. Suppose superstars had superior traits to our matched control group. In that case, this condition is violated, and we would not be able to parse out the causal effect of inbreeding from the effect of the bull’s genetic superiority. Assumptions 3 and 4 ensure that the OLS estimates $\hat{\beta}$ from Model (2) are consistent and asymptotically valid (Roth et al., 2023).

Assumption 5: Let $W_j = (inb_{j,pre}, inb_{j,post}, T_j)'$ be a vector of outcomes and treatment status for unit i . Then, we observe a sample of N independent and identically distributed random draws $W_i \sim F_W(w)$ for a distribution F satisfying parallel trends.

Under Assumptions 3, 4, and 5: $\sqrt{N}(\hat{\beta} - ATT_{post}) \xrightarrow{D} N(0, \sigma^2)$, then we can consistently estimate the variance σ^2 clustering at the line level (Roth et al., 2023) asymptotically ($N \rightarrow \infty$ and T fixed).

We can also decompose the variations in treatment time since every year, a certain number of bulls will be born from lines in both groups, each with its inbreeding rate that can be averaged across treatment and control groups. Models 2-4 show the complete specification with leads and lags of the treated group. These models decompose the estimated ATT relative to the treatment average by birth year. Sun and Abraham (2021) call this specification “Dynamic specification.” However, there is no staggered treatment since genomic selection affects all lines in the same year, but inbreeding rates evolve independently after the treatment.

This assumption states that the treatment status of an individual from any given line is not affected by the treatment status of others. In our setting, the treatment (genomic selection) affects all individuals, so there is no selection into treatment. To measure the effects of genomic selection on inbreeding, we estimate the differences-in-differences model:

$$inb_{ikt} = treat_i + treat_i * \sum_{\substack{t=2005 \\ t \neq 2009}}^{2019} \beta_t \mathbf{I}(yob = t) + \alpha_t + \alpha_k + \varepsilon_{ijt} \quad (\text{Model 2})$$

$$inb_{ikt} = treat_i + treat_i * \sum_{\substack{t=2005 \\ t \neq 2009}}^{2019} \beta_t \mathbf{I}(yob = t) + \alpha_t + \alpha_k + \mathbf{X}_{it} \gamma + \varepsilon_{ijt} \quad (\text{Model 3})$$

$$inb_{ikt} = treat_i + treat_i * \sum_{\substack{t=2005 \\ t \neq 2009}}^{2019} \beta_t \mathbf{I}(yob = t) + \alpha_t + \alpha_k + \mathbf{X}_{it} \gamma + (\mathbf{X}_{it} * post_t) \rho + \varepsilon_{ijt} \quad (\text{Model 4})$$

Model (2) is a specification that does not contain any covariates (PTAs), and it decomposes the ATT on inbreeding rates by year of birth cohort. Model (3) incorporates PTAs ($\mathbf{X}_{it}$) and firm-level fixed effects, and Model (4) includes interactions of PTAs and the post-treatment dummy. We include the last specifications to account for trait-level trends. We also know that the treatment affects both inbreeding rates and PTAs, but we do not have a structural model for their impact (Lewbel, 2019), so we include them in a non-linear fashion as both levels and interactions to reduce the uncertainty about the functional form of our controls.

2.5 Results

Figure 2.7 shows treatment and control groups’ average inbreeding rates and confidence intervals across years. Genomic selection did not increase inbreeding rates immediately; both curves began to diverge substantially after 2011. Inbreeding rates, in turn, started growing two years later, and both curves diverged significantly from 2012 onwards.

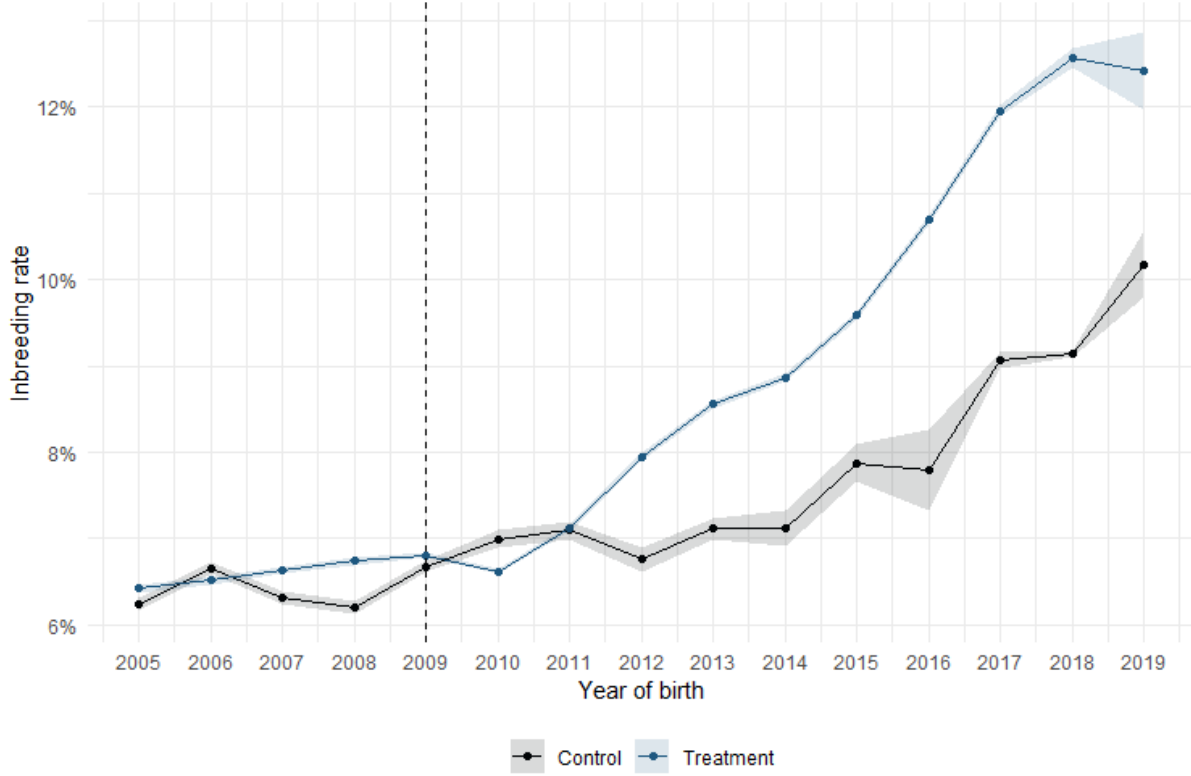


Figure 2.7. Average inbreeding rates in treatment and control groups

This plot closely matches the evolution of the population average inbreeding rates plotted in Figure 2.1. This behavior is consistent with our hypothesis that breeders prefer well-known over new lines. Our findings coincide with results reported by Guinan et al. (2023), who also look at the time trends of productivity and reproduction traits in US dairy cattle. The authors also document a decrease in the generation interval and an increase in inbreeding rates of Holstein and Jersey cattle but do not consider the latter an issue because the expected inbreeding rates of future offspring adjust genetic trait values.

Table A3 shows the results of three models; model (1) is the analog to Figure 2.7, including only the interactions between treatment and year of birth. Model (2) splits the treatment effect by year of birth averages, Model (3) includes the complete set of genetic traits and firm-level fixed effects, and Model (4) is the model with traits and a post-treatment dummy. In all specifications, standard errors are clustered at the line level to better deal with the inconsistency generated by the autocorrelation of inbreeding rates between different generations (Bertrand et al., 2004).

Model 1 is the simple differences-in-differences estimation where $post_t$ is a dummy equal to 1 if the bull was born after 2009. The results show that bulls descending from supersires increased their inbreeding rates by 1.46 percent more than those who descend from less popular bulls, on average. Since the impact of genomic testing on inbreeding rates only became noticeable after 2011, the coefficient underestimates the ATT of genomic selection. Model 2 shows that the treatment-year of birth interactions are positive and significant from 2012 onwards, with the 2018 effect being the highest.

Models 3 and 4 in Table A3 show that there is a positive and significant incremental variation for the treated versus control groups in the years 2012-2018; however, the addition of controls in either levels

or interactions attenuates the coefficients for 2014 and 2015 to the point of reducing the significance of interactions between treatment and year of birth. Still, in all instances, the coefficients are positive. Regression models show a positive and increasing inbreeding rate in the treatment group comprised of all descendants from prestige sires relative to the control group of descendants of comparable non-superstar sires.

The last period in the sample is 2018 instead of 2020 because we cannot make inferences after 2018 due to a reduction in the number of animals in the control sample; the number of comparable descendants is too small for us to consider them a valid comparison group. This drastic reduction is a consequence of lower fertility rates in the control group; since those sires fathered fewer sons, their sons will beget even fewer grandsons, and after a few years, this group collapses.

Table A3 also includes the results for the F-test for joint significance of all pre-treatment differences in trends for the interaction of treatment and year of birth. The p-values show no significance in two of the three specifications and a 5% significance for the pre-treatment coefficients of Model 4. This is not a test for parallel trends since the assumption relies on unobservable counterfactuals, letting $\hat{\beta}_k$ be the coefficient for any of the post-treatment periods, then:

$$E[\beta_k] = ATT_k + (E[inb_{ik}^{control} - inb_{i0}^{control} | T_i = 1] - E[inb_{ik}^{control} - inb_{i0}^{control} | T_i = 0])$$

The parallel trends assumption implies that the term in parentheses is equal to zero, then the coefficients correctly capture the ATT for any post-treatment period $k = t - 2009 > 0$. On the other hand, its pre-treatment analog is identified as:

$$E[\beta_{(-k)}] = E[inb_{i(-k)}^{control} - inb_{i0}^{control} | T_i = 1] - E[inb_{i(-k)}^{control} - inb_{i0}^{control} | T_i = 0]$$

We have shown that these coefficients are not significantly different from zero for all pre-treatment periods, individually and jointly. It is usually considered evidence in favor of the parallel trends assumption. However, this is not necessarily the case since this test has low power against a linear or quadratic pre-trend alternative (Roth, 2022). In our case, we can dismiss this concern; since our matching method ensures the fulfillment of the condition for the founders, the entire effect must be due to their sons and grandsons.

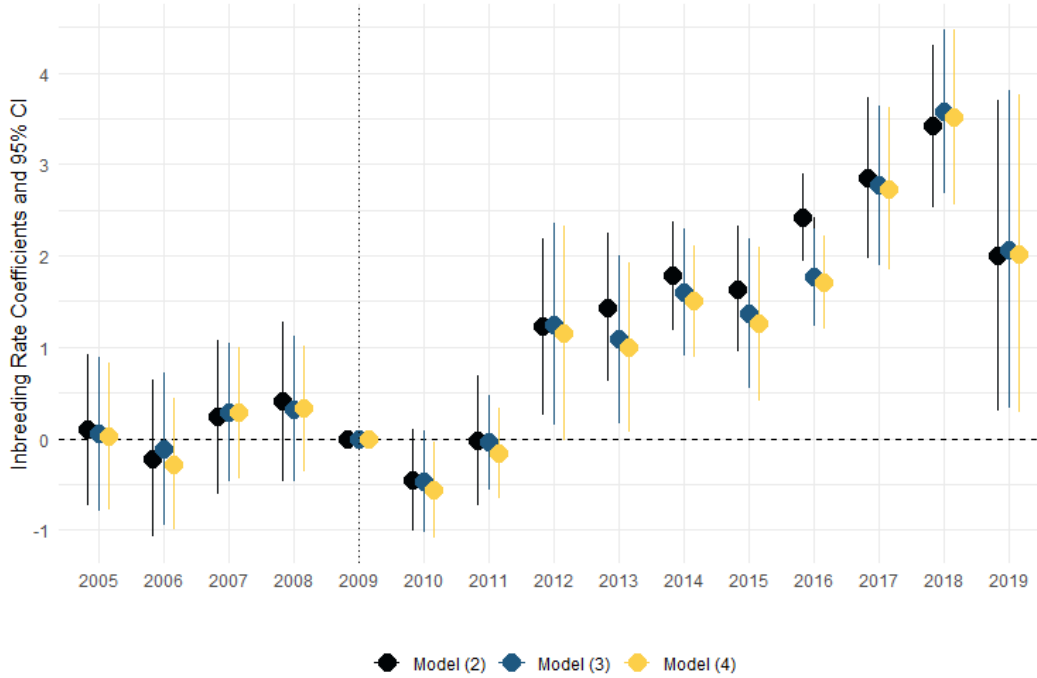


Figure 2.8. Coefficient plots for different specifications

Figure 2.8 shows the coefficients from the interaction of year fixed effects and treatment status to assess the impact of inbreeding per year correctly; every year after 2012, the inbreeding rate of the treatment group is between 1 and 2.8 percent points higher than the control group. The coefficient values decrease slightly after controlling for their genetic traits but remain statistically significant. The figure also shows that all pre-treatment coefficients are not significant; we interpret this as evidence in favor of the reasonability of the research design. Figure 2.8 uses the *sup-t* uniform confidence intervals for the pre-treatment periods of Montiel-Olea and Plagborg-Møller (2019) as suggested by Freyaldenhoven et al. (2019) instead of the standard pointwise confidence. These confidence intervals contain the actual path of the coefficients 95% of the time and can better help understand which types of pre-trends are consistent with the data.

So far, we have discussed our results regarding inbreeding rates only, but we can attach a monetary value to increased inbreeding rates. L. A. Smith et al. (1998) shows that every additional inbreeding percent is associated with 24.45 USD less income by milk sales. The total cost depends on what we consider the average inbreeding; we can use the control group rates as a counterfactual and calculate the cost from increased inbreeding of the treatment group, multiplied by the total stock of Holstein cattle; in that instance, we estimate a total cost from increased inbreeding of 4 billion dollars for the period 2011-2019. Alternatively, by industry standards, the maximum recommended inbreeding rate is 6.25%¹, so we can calculate how much is the cost from inbreeding levels greater than such value, we estimate the total cost for the same period as 7.4 billion dollars for the entire dairy industry.

The benefits from the introduction of genomic selection to the dairy industry can also be calculated using an index called **Net Merit**, which measures the net profit over the lifetime of the animal's average daughter, expressed in U.S. dollars. Net merit is calculated by multiplying the values of PTAs by a set of dollar values, intended to capture the "market value" of all traits. Table A1 shows what are the magnitudes

¹<https://www.agproud.com/articles/36810-the-impact-of-genomics-on-rapid-increase-of-inbreeding-of-holsteins>

of such weights (Van Raden et al., 2021), note that the calculation involves the usage of subindexes for health and reproduction traits, as well as an estimation of residual feed intake per animal.

Across the 2011-2019 period, this value is 90.81 billion US dollars, and the cost/benefit ratio for Genomic Selection ranges from 4.4% to 8.1%, depending on which cost calculation we employ.

Alternatively, we can run Model (1) using Net Merit as the outcome to estimate the impact of genomic testing. We have that the net Merit increases by 102.2 USD per animal on average across the period. Using the 24.45 USD per inbreeding percent, we can convert the 1.4 additional points of inbreeding in USD dollars and subtract it from the gains in Net Merit; this equals 74.25 USD per animal on average across the 2011-2019 period for the treatment versus the control group in our sample.

2.6 Robustness Checks

2.6.1 Total number of descendants

In our analysis, we defined a “supersire” as a bull whose total number of descendants was among the 95th percentile of their distribution by 2021. This calculation is also influenced by the cutoff point since genomic testing leads to an expansion of the supply of bulls, then the size of any line is endogenous to genomic selection. It can be argued that selecting bulls on the ex-post number of sons is tautological to our argument.

Figure 2.9 plots the coefficients for the different differences-in-differences specifications, but this time, the number of descendants is calculated only up to 2010, so we assume that genomic selection is just being introduced, and we are trying to guess which lines will be more successful. Coefficients for 2014 and 2015 are now substantially reduced and become insignificant, but the overall result remains unchanged, we still have that the treatment group has notoriously higher inbreeding rates than the control group, as seen in the 2×2 model, where the coefficient is still positive and significant for the entire period, as seen on A4.

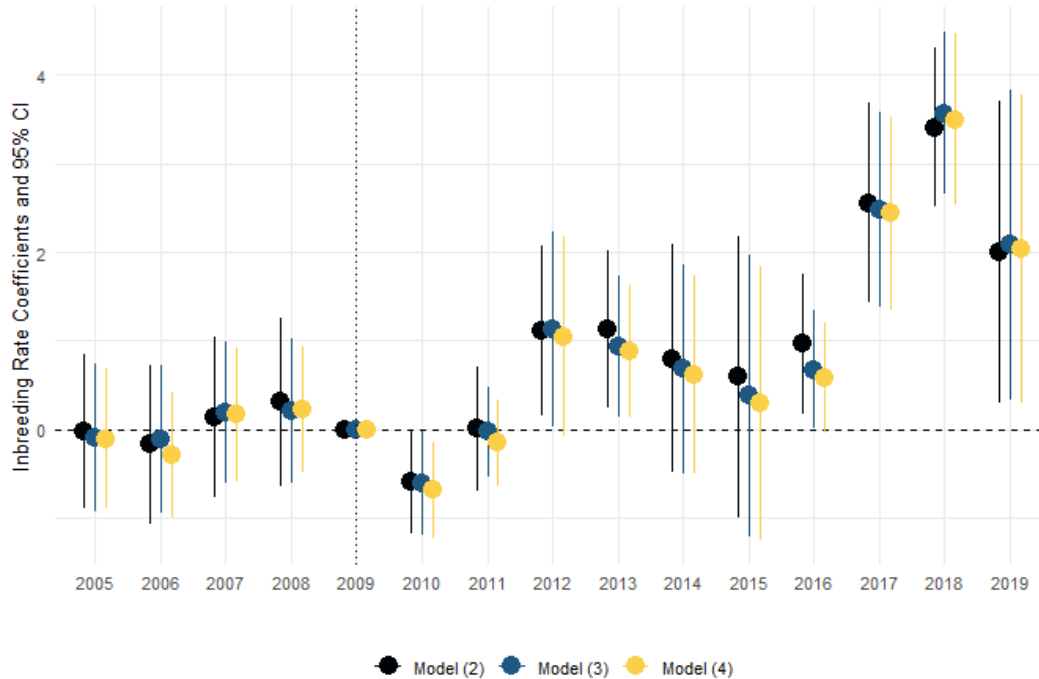


Figure 2.9. Coefficient plots for an alternative definition of supersire

2.6.2 Quarter of birth effects

Since we have the exact date of birth of every animal in the sample, we can group them by a quarter of birth as well; Figure 2.10 shows the results from Models 2-4 but replacing year for a quarter of birth interactions with the treatment. The signs and magnitudes of coefficients do not differ substantially from the original formulation. The increase in coefficient values is monotonic and statistically significant since 2012, but having a smaller sample size per quarter makes confidence intervals wider.

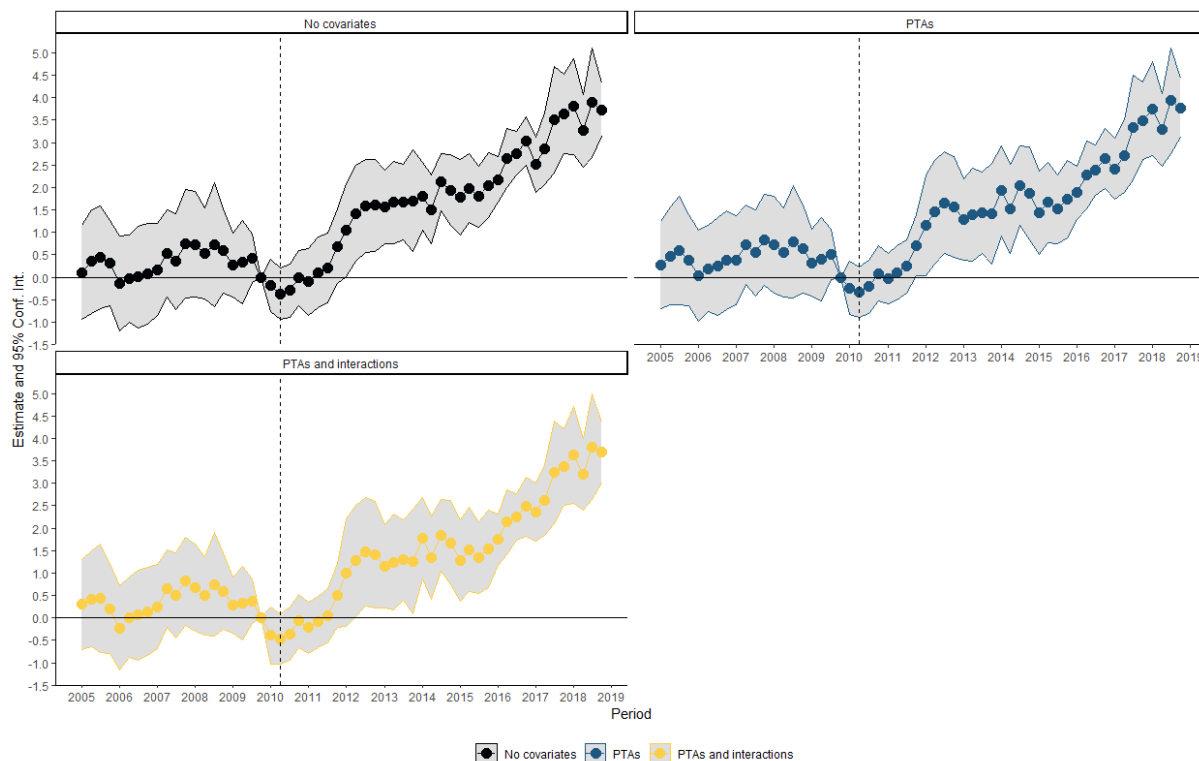


Figure 2.10. Coefficient plots for quarter of birth effects

This article’s results depend heavily on a somewhat arbitrary definition of a “superstar” bull in terms of the size of his line. What would happen if we changed that from the 95th to another quantile of the line size distribution? Figure 2.11 shows the coefficient plots for different definitions of “superstar”.

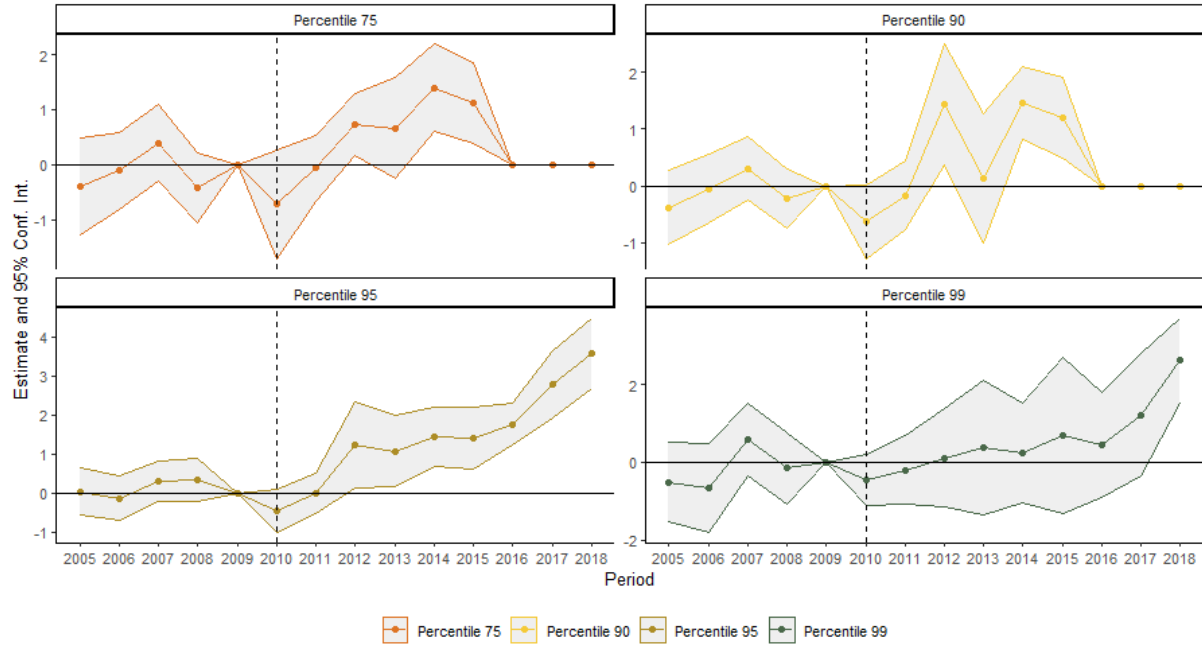


Figure 2.11. Robustness test: changes in the definition of “superstar”

If we change the cutoff point from the 95th to the 99th percentile, the sign, and magnitude of our coefficients do not change much. Still, the confidence intervals widen significantly because our treatment group comprises such specific animals that the matching algorithm cannot find suitable matches. As a consequence, the significance of the difference-in-differences coefficients decreases significantly. Going in the other direction, we can find marginally better matches for the superstar group. However, the control group ended abruptly in 2015 because most of the original group’s control lines are now included in the treatment, making the control group short-lived. We conclude that our results depend critically on a set of lines that allow for a good comparison when included in the control group; otherwise, the analysis yields insignificant results.

2.6.3 Date of birth of “superstars”

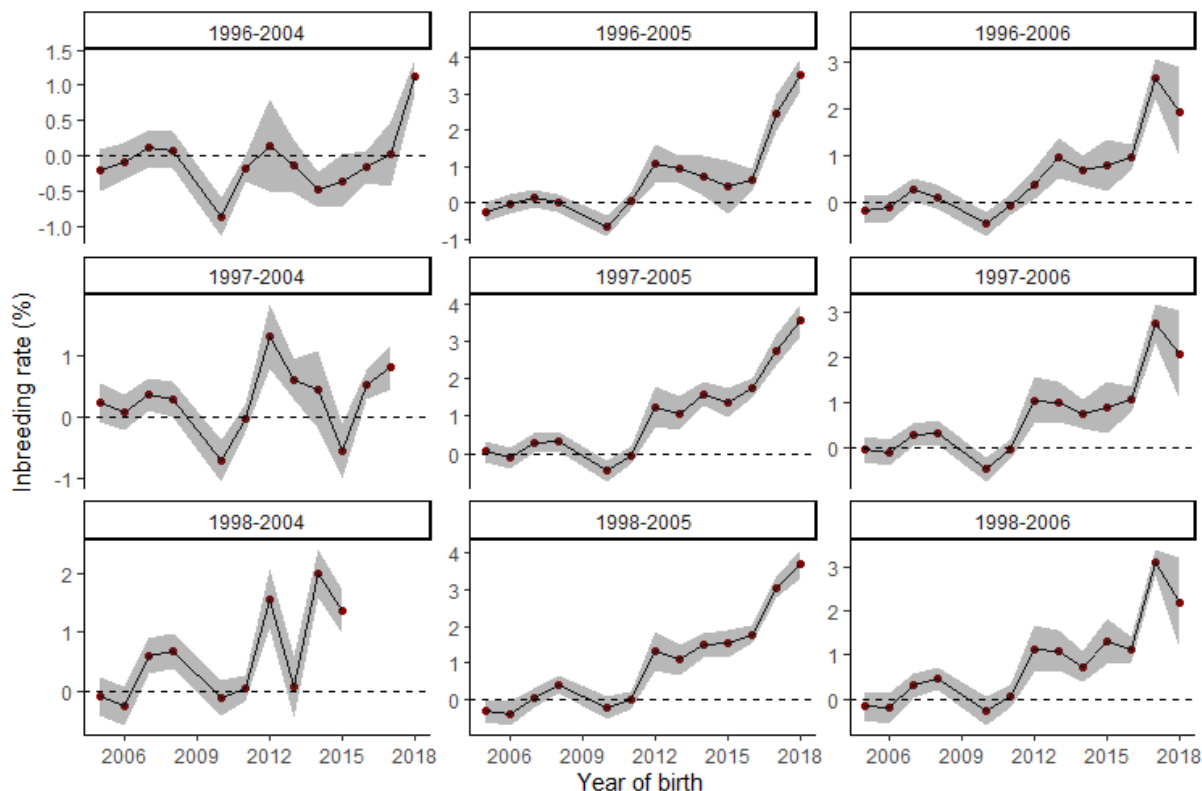


Figure 2.12. Coefficients from model (2) per different definitions of founders’ date of birth

Finally, Figure 2.12 shows coefficient plots by different definitions of the founders’ date of birth. We defined founders as all bulls born between 1997 and 2005, but such a definition is also arbitrary. If we adopt different cutoffs for the period where the founders were born from 1996/1998 to 2004/2006, we have $3 \times 3 = 9$ distinct combinations for starting and end date. The coefficient plot shows that the bulk of the effect can be attributed to sires born from 1996 to 2006, if we extend further the selection period of sires, we will have more endogeneity issues since, by this date, it was known that the bovine genotype was being sequenced, and thus, that there would be a significant effect on the market.

2.7 Discussion

This article investigates the consequences of genomic selection on the American dairy industry. On the one hand, this new technology significantly improved productivity, but on the other hand, it increased the average inbreeding rate of all bulls in the market. Such increment in productivity was fueled by two effects: a decreased generation interval motivated by a new technology that accurately estimates a bull’s genetic traits as soon as it is born and more intense competition between animal breeders to produce high-performing bulls. Faster access to information allowed breeders to select better bulls faster, and consequently, genetic gains per generation are now much higher than they used to be in the past.

The most popular method to breed highly productive bulls is by breeding lines that descend from a famous ancestor; we call these animals supersires and show that breeding on these well-known lines is the

most critical driver of a greater degree of inbreeding among bulls. As the average inbreeding rate increases, every new generation of animals experiences an increased risk of experiencing inbreeding depression, which in turn leads to lower fertility and a higher prevalence of genetic diseases. Inbreeding depression is a cost that breeders impose on others, a negative externality that is a function of the number of breeders that release closely related bulls into the market. Since pedigrees transmit the externality, it is a network externality that affects all breeders and dairy farmers.

It is impossible to directly test the impact of genomic selection on inbreeding rates because there is no natural control group and no lines unaffected by genomic selection. We used a propensity score matching method to find equivalent families based on the founder’s genetic traits but had fewer offspring; breeders could have invested in them but did not. We then track those families across years and observe the average inbreeding rates of their offspring. We show that bulls from the treatment group (those who descend from famous ancestors) have consistently higher inbreeding rates than those from less-known ancestors from 2012 to 2017. Our analysis shows that the cost of this increased inbreeding amounts to around 500 million dollars.

Inbreeding rates can be reduced significantly by importing bulls from other countries that are not (or less) related to supersires. Another possible solution is to adjust the values of genetic traits by inbreeding; the Holstein Association of the United States adjusts the values of PTAs according to inbreeding levels to penalize breeders that select bulls with high inbreeding rates to account for the higher reproduction costs of such animals. The value of this penalty is a crucial element in reducing inbreeding by incentivizing dairy farmers and animal breeders to breed lines that are less related on average by, for example, importing bulls from abroad or by crossbreeding Holsteins with other high-yield milk breeds.

CHAPTER 3

ECONOMIC VALUE OF BRANGUS CATTLE TRAITS IN ARGENTINA

3.1 Introduction

Cattle adaptation to new environments is a possible pathway for increasing productivity and meat supply. Reca et al. (2010) argues that in Argentina, changes in relative prices between agriculture and livestock pushed cattle ranching out from the core of the Pampean prairie to other regions of the country. The demand for adaptation in less productive areas, such as subtropical regions, created the economic opportunity to develop new synthetic breeds, like Brangus and Bradford. These breeds preserve the meat quality of Angus and Hereford bulls while improving the animals' ability to withstand heat stress, mainly by crossing these breeds with Brahman bulls. Estimating the value of bull characteristics will help ranchers make investment decisions in genetics.

While there is evidence of the adoption of crossbreeds, there is little to no information on the economic value of crossbreed traits in Argentina. Argentina's relevance in beef markets is two-fold. First, Argentina is a historical producer and exporter of beef. In 2022, Argentina was the sixth producer and fifth exporter country in the world (USDA, 2023). Like other producing nations from the developing world, its modal production system relies predominantly on pastures rather than feedlots (Greenwood, 2021). The latter is relevant because there is potential for productivity and efficiency gains from introducing changes to genetics in harmony with sustainable production practices. Second, because of its location, Argentina has the ecosystems to adapt breeding options for mild and subtropical weather, with the potential of developing and exporting cattle genetics to different parts of the world.

This article aims to estimate the economic value of genetic and non-genetic traits for Brangus bulls in Argentina. Ranchers select bulls according to their preferences for traits. A trait is any characteristic of the animal that has an intrinsic value. They can be phenotypical, observed, such as the hide color or sex, or genotypical (estimated) based on the Expected Progeny Differences (EPDs). EPDs are measures of the genetic contribution of an animal to the set of genetic traits of its progeny. These indicators provide helpful information regarding a bull's likelihood to father calves with superior attributes relative to its ancestors. Examples of EPDs are birth weight, weaning weight, and scrotal circumference (Dhuyvetter et al., 2005). For instance, a farmer looking to reduce birth complications would select a bull with a negative birth weight EPD. Incorporating that characteristic in the herd is intended to reduce the time of vigilance of calving cows.

We hypothesize that EPDs' economic value is affected by farmers' objectives and characteristics. The use of EPDs for bull selection is a new tool, and its value comes from improving precision for genetic information and investments. If EPDs are superior technology compared to bulls' visual aspect and own data, we should have observed widespread adoption; however, this is not the case. Foundational research in agricultural

economics shows that technology is not a random process, and availability does not necessarily imply adoption (Hayami and Ruttan, 1985; Schultz, 1983; Sunding and Zilberman, 2001). At some point, top farmers adopt first, and other farmers become followers. Innovative farmers are less risk-averse and have lower adoption costs, giving genetic information relevance in bull selection. Traditional farmers’ choices are based on a combination of visual inspection and phenotypical bull data. It takes more time than top farmers to adopt new tools. Hence, we should expect a coexistence of farmers with different levels of technology adoption at any given point in time.

We use a hedonic price model to estimate the value of EPD information. We follow the “Prices and Demands for Input Characteristics” (1976) approach, which asserted that the value of an input can be decomposed into the value of each one of its characteristics. We use data from Brangus bull auctions from 2015 to 2022. Our dataset contains information for bull sales price, their phenotypical attributes, and EPD information. Our results show that phenotypical traits are more relevant than EPDs as decision variables for cattle ranchers. We find that cattle ranchers prefer observed traits such as weight, scrotal circumference, and coat color, while EPDs have secondary importance after controlling by sire, stud, and pedigree. In general, EPDs seem less relevant for the price-setting process than phenotypical traits; however, they are more important by groups. Production and reproductive traits taken together are statistically significant. We argue that the preference for red-coated bulls is a means to adapt to subtropical climates. Conversely, the weak association between bulls’ prices and EPDs may be related to a lack of information about the nature and uses of genetic traits.

Our results contribute to understanding the process of technology adoption in cattle production. We provide evidence that, while genetics tools and information are available, they are still not fully adopted. Our results indicate that there is still room for breeding programs to work on extension and education about new traits and breeds. Moreover, we generate evidence regarding the economic value of attributes of the Brangus breed, which has yet to be explored in the literature. Producers could benefit from obtaining information on animal genetic attributes to increase productivity per hectare, feed efficiency, and livestock quality. Likewise, the industry could benefit from cattle with genetic improvements resulting in better meat quality and quality that facilitates processing.

3.2 Beef Cattle Genetics Market in Argentina

3.2.1 Supply of Beef Cattle Genetics

The supply of cattle genetics is comprised of studs that develop service bulls and heifers, semen and embryos and sell them to commercial ranches, who demand such genetics (Etcheverry, 2009). A series of factors influence beef supply; chiefly among them is the existence of diverse breeds. Crossbreeding is a method for increasing the productivity of cattle genetics, which constitutes a climate-induced innovation by the demand for beef and other factors, such as grassland and pasture availability (Espasandín and Ducamp, 2004). Cattle breeders and feeders demand high-productivity bulls that increase the efficiency of their herds and sell their meat to domestic and foreign markets (Marcantonio, 2017).

Beef cattle genetics in Argentina have been historically based on European breeds such as Aberdeen Angus and Hereford. Because of the demand for adaptation, some studies have developed Brahman genetics. Ranches that breed Brahman bulls are mainly located in the northeastern region of Argentina, which observes subtropical conditions that are less suitable for Angus or Hereford. According to the 2018 National Agricultural Census, there are 788 bovine studs. Studs that keep bulls from these breeds are located mainly

in the temperate Pampas region and constitute around 75% of the supply. Brahman studs constitute 2.4% of the total, but they are essential to improving the adaptability of Angus and Hereford breeds to subtropical climates.

The different climates strongly influence bovine genetics supply spatial distribution (Figure 3.1). Firms aim to breed pedigree animals adapted to the ecosystem where they are located. Ranchers in the area might be interested in buying bulls adapted to the productive environment. This is important because the bull will be responsible for 50% of the genetics of the future herd. If that herd receives traits that are not suitable, it may harm productivity and ranchers' future income. The two main crossbreeds are Brangus (Brahman $\times$ Angus), and Braford (Brahman *times* Hereford) expanded beef cattle production to new areas. Figure 2 shows examples of three different breeds and how much a Brangus bull resembles a mixture of its original breeds. Brangus and Braford breeds constitute 31% of stud farms in Argentina. The geographical distribution of both breeds shows they are in areas where Angus and Hereford bulls are not the first choices.

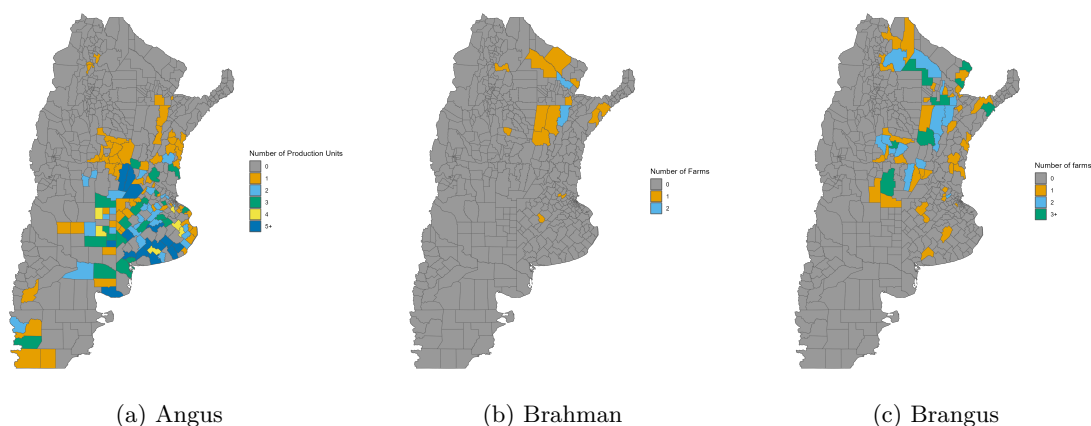


Figure 3.1. Spatial distribution of farms by breed

Note: Author's calculations from 2018 National Agricultural Census data.



Figure 3.2. Typical visual characteristics of all three breeds

¹<https://domesticanimalbreeds.com/angus-cattle-breed-everything-you-need-to-know/>

²<https://brcutrer.com/health-and-wellness-with-brahman-cattle/>

³<https://www.bovine-elite.com/shop/beef-semen-sales-registered/brangus-black/mr-new-blood-50h/>

3.2.2 Demand for Cattle Genetics

The demand for cattle genetics comes from cattle farmers who purchase bulls for crossbreeding and trait selection. Crossbreeding is a response to long-term demand trends from cattle farmers, and its main objective is to take advantage of heterosis, an increase in the performance of hybrids over that of purebreds caused by specific gene combinations (Bourdon, 2000, p. 29). In the case of Brangus animals, they have better survivability in hotter climates than Angus and can gain weight at a faster rate than Brahman bulls. Crossbreeding takes several generations; for example, Brangus and Braford crossbreeds require around ten years to stabilize their phenotypic and genotypic traits due to small initial herd numbers. Once traits are stable, breeders could use them in a genetic plan to select attributes they want to improve in their herd. For instance, Brangus has received increased attention in Argentina because of its adaptation to different climates. The Brangus herd book, which approximates the breed presence in the national herd, shows an increase of 5.9% on average in the number of calves under the Brangus genetic evaluation in the last 20 years (Figure 3.3).

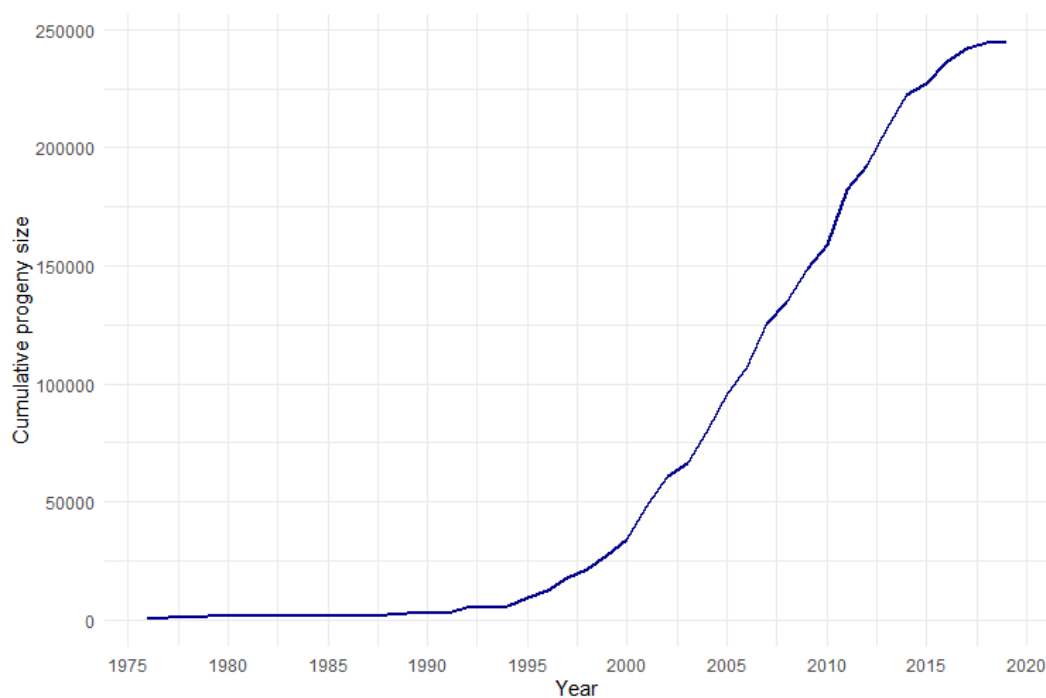


Figure 3.3. Number of calves by bulls' birth year present in the Brangus herd book

In Argentina, cattle studs sell bulls through auctions or direct sales to commercial ranches. Breeders' Associations sponsor the auctions to signal the quality of bulls. Further, two types of bull are sold on whether they have complete or incomplete genetic data. Those with complete ancestry and pedigree data are purebred animals registered in the breed registry of the corresponding association. The main buyers of purebred bulls are other studs interested in reproducing them and offering more bulls for commercial ranches. The other type of bull offered in sponsored auctions are animals with incomplete data that have passed a visual inspection and gene testing; these are marketed as pure controlled animals. These bulls are generally cheaper than

³Source: Own calculations based on the 2023 Brangus Association herd book. Available at: <https://brangus.org.ar/programa-erbra/>

purebred animals and the primary alternative for cow-calf operators interested in investing in genetics.

Conditional to having or choosing a breed for the herd, farmers would select the bulls according to their preferences. Service bulls are the most common option for commercial ranchers interested in investing in genetic improvement. Several attributes matter in selecting the best bull. We can divide these aspects into (i) visual characteristics that describe bulls’ breeding ability associated with their biological type and age; (ii) phenotypical traits are observed characteristics of bulls that are generally associated with their productive and reproductive abilities; and (iii) genetic traits which are objective measurements performed on pedigree cattle with genomic tests to determine their ability to pass certain traits to future generations.

Genetic traits are called Expected Progeny Differences (EPD) and are the genetic value of an animal as a parent, the value of an animal’s genotype due to independent and transmittable gene effects (Bourdon, 2000). The intuition behind these estimations is that after controlling for all possible environmental factors that affect the characteristics of an animal, we end up with a portion of the variation in the value of each trait that can only be attributed to the inheritance of the animal’s parents. These values are calculated using large-scale genetic evaluations that breed associations typically carry out periodically; results from these evaluations are reported in the sire summary, a list of animal traits relative to the average sellers use to market the bull. Examples of EPDs match phenotypical data of the bull, such as birth weight, weaning weight, and scrotal circumference, or meat carcass attributes, such as ribeye area (Dhuyvetter et al., 2005; Mitchell et al., 2018). For instance, selecting bulls with a negative birth weight EPD increases the probability of having lower birth weight in calves, which could facilitate births and reduce costs for farmers. Further, its correlation with a high EPD weaning weight would increase the chances of increasing calves’ weight-gaining productivity.

3.3 Theoretical Framework

3.3.1 Hedonic pricing model

We introduce the hedonic pricing model, suitable for estimating the demand for cattle traits as an input of the commercial ranchers’ production function. The hedonic pricing model asserts that the price of a good or service (livestock genetics in this case) is a function of its intrinsic characteristics Rosen (1974). Similarly, the Input Characteristics Model developed by “Prices and Demands for Input Characteristics” (1976) is the theoretical basis for adapting the hedonic pricing framework to input demands.

According to this model, the price paid for a unit of the input depends on its set of characteristics; consequently, the value of such characteristics can be estimated using a linear regression of the price on their quantities. Dhuyvetter and Schroeder (2000) adapted this framework to the particular case of feeder cattle pricing, similar to our case, where the price of any lot of cattle depends on the attributes of the animal, which can be observed (weight, age, color, or breed) or estimated (such as the Expected Progeny Differences or selection indexes), as well as market factors such as expectations about input prices, interest rates or beef prices. In other words, the price of a bull i at time t on market h is the n:

$$price_{it} = f(X_{it}, Z_{ht}) \quad (3.1)$$

Where X_{it} are the animal’s characteristics and Z_{ht} are the factors affecting market h .

The literature on hedonic pricing has looked at the market value of cattle traits. (Dhuyvetter et al. (1996) is the first article that includes estimated and observed genetic traits in the pricing equation and additional

variables such as the presence of a picture in the bull catalog, sale location, and percentage of semen rights kept by the seller. The authors find that all variables related to the animal’s weight are significant (birth weight, weaning weight, and weight EPD), as well as those related to the visual aspect of the animal (color, polled, muscling, and conformation). Concerning market factors, sale location, picture in the catalog, order of sale, and location have significant effects on the price of the animal. Jones et al. (2008) improve on earlier articles by expanding the set of EPDs and market factors; the most important finding is that weight EPD had a higher value than the observed weight in their sample and predictors of carcass quality. This article also incorporates pedigree into the analysis using sire fixed effects and marketing factors such as order of sale, picture of the animal, and season. In dairy markets, Richards and Jeffrey (1996) elucidate the impact of production and health traits on the price of bulls’ semen. This article finds that production traits (milk yield, protein, and fat content), general conformation (also called “type”), body capacity, and bull popularity significantly impact the animal’s price.

The literature uses data from cattle associations or auctions in different places in North America. Walburger (2002) showed that the most critical traits for Canadian breeders are sale weight, birth weight, scrotal circumference, ribeye area, and weight gain; that is, production and reproductive traits, but the former have increasing importance. These findings are replicated in other segments of the market. For instance, Boyer et al. (2020) studied which factors affect the price of bred heifer prices, particularly months of pregnancy, lot size, heifer price, and timing of purchase affect their price: the authors also found that the effect of lot size is non-linear on the logarithm of bred heifer price.

Another strand of literature has investigated the demand side of cattle markets using surveys or choice experiments. Sy et al. (1997) ran a survey from different segments of the market (purebred breeders, commercial cow-calf producers, and cattle feeders). Each segment prefers one set of traits over another due to its distinct profit maximization objectives. Purebred breeders place more weight on milking ability and weaning weight, cow-calf operators value calving ease and temperament, while feeders prefer animals with higher slaughter weight and feed efficiency.

Recently, attention has shifted toward evaluating the impact of video cattle auctions, such as the Superior Livestock Auction Zimmerman et al. (2012), which allows buyers from different locations to participate in the auction via the Internet. Video auctions created a demand for specific management practices such as age and source verification, vaccination protocol compliance, and weight variation certification. Similarly, Martinez et al. (2021) show that feeder cattle were sold at a premium if the animal has been tested for bovine diarrhea virus but also finds a significant impact of corn future prices on the valuation of animals.

3.3.2 Input Characteristics Model

We use the Input Characteristics Model (“Prices and Demands for Input Characteristics”, 1976) to explain the demand for cattle traits. The model assumes that a firm uses a set of n inputs used to manufacture one unit of a good q that is sold at a price p , let r_i be the price of a unit of the input, which requires a fixed amount of a characteristic, such that ω_{ji} is the amount of characteristic j required to produce one unit of input i , then:

$$\omega_j = \sum_{i=1}^n \omega_{ji} x_i \quad (3.2)$$

The production function depends, in turn, on the entire set of characteristics, such that:

$$q = F(\omega_1, z_2, \dots, \omega_m) = F\left(\sum_{i=1}^n \omega_{1i}x_i, \sum_{i=1}^n \omega_{2i}x_i, \dots, \sum_{i=1}^n \omega_{mi}x_i\right) \quad (3.3)$$

The profit function can be written as:

$$\pi = pF(\omega_1, \omega_2, \dots, \omega_m) - \sum_{i=1}^n r_i x_i \quad (3.4)$$

The first-order conditions for $i = 1, \dots, n$ are:

$$\begin{aligned} \frac{\partial \pi}{\partial x_i} &= p \sum_{j=1}^m \frac{\partial F}{\partial \omega_j} \frac{\partial \omega_j}{\partial x_i} - r_i = 0 \\ &= p \sum_{j=1}^m \frac{\partial F}{\partial \omega_j} \omega_{ji} - r_i = 0 \end{aligned}$$

Solving for r_i , we get:

$$r_i = p \sum_{j=1}^m \frac{\partial F}{\partial \omega_j} \omega_{ji} = \sum_{j=1}^m \tau_j \omega_{ji} \quad (3.5)$$

Then, the price of input i is a linear function of the number of characteristics weighted by the value of their marginal product, τ_j .

3.3.3 Empirical Specification

We estimate the input characteristics hedonic model for identifying where the price of a bull i in year t can be described as:

$$\log(\text{price}_{it}) = \mathbf{X}_{it}\gamma + \mathbf{Z}_{it}\beta + D_i\gamma + \sigma_i + \rho_t + \varepsilon_{it} \quad (3.6)$$

Where price_{it} is the price of the bull, $\mathbf{X}_{it}$ is a set of phenotypical traits, $\mathbf{Z}_{it}$ is a vector of genetic traits. In our model, $\mathbf{X}_{it}$ includes the animal's weight, age, or color, which indicates the bull's condition and characteristics. The second group of variables, genetic traits ($\mathbf{Z}_{it}$), are statistical estimations calculated from the performance information of an animal relative to its relatives, past and present, controlling for pedigree, age, breed, season, and environmental factors. These data represent the potential of the bull to pass features to its progeny. Since EPD values do not observe significant variation within a stud and are strongly correlated, we include them as dummy variables that take a value of 1 if the EPD is above the median and 0 otherwise. In addition, we incorporate a vector of dummy variables to control differences in stud characteristics (stud_i) and the genetic quality of the bulls (PP_i). Specifically, the latter takes a value of 1 if the bull is Pure Pedigree (PP) and 0 if it is a Pure Controlled (PC) bull. Lastly, to control potential differences in auctions by year, we include a set of fixed effects, including sire (σ_i) fixed effects, and year (ρ_t) fixed effects, and, ε_{it} is the error term⁴

⁴See Appendix A for more details.

3.4 Data

Our data consists of a list of bulls sold in auctions in August of every year from two studs in the Northwestern districts of Córdoba, Argentina. One has records from 2015 to 2022, while the other spans the years 2018 to 2022. Both studs have more than 20 years of auctioning bulls and heifers to ranchers from the northwest of that province. They sell Brangus bulls suitable for ranchers in the region, but also for ranches in the northern provinces of Argentina, Paraguay, and south of Brazil.

We used the price and their phenotypical and EPD information for each bull. The phenotypical variables available are coat color, which takes a value of one if it is red, 0 if it is black, weight of the bull in kilograms, scrotal circumference in centimeters, and age in years. The estimated genetic data contains two indicators: productive and reproductive EPDs. The first group indicates the potential growth attributes to be passed to the calves of the bull, which includes birth, weaning, and final weight progeny potential. The second group indicates the ability of the bull to transmit reproductive features, and it contains the scrotal circumference and maternal aptitude of heifers. The scrotal circumference EPD is associated with improving male progeny fertility and early precocity in female progeny. Maternal EPD indicates the potential for producing more kilograms at weaning because of cows' milk production.

Table 3.1 presents summary statistics for 723 Brangus bulls auctioned from 2015 to 2022 and a comparison with the EPDs of the base population of the Brangus herd book⁵. Descriptive statistics show that the weight of the bulls is 750 kilograms, they are 2.2 years old, and the scrotal circumference is 39 centimeters. In our sample, 33% of the bulls have a red coat, while the remaining animals are black Brangus. We use the EPD herd book averages to show whether auctioned bulls are representative of the national supply of Brangus genetics. We can interpret EPDs as deviations from the herd book-based population. Specifically, these bulls give calves with birth weights about the breed's average (+0.08). Calves may gain 9 kilograms on average at the weaning stage, while the final adult weight would be 19.9 kilograms on average. The reproductive EPDs show that bulls have good qualities regarding scrotal circumference (+0.02 cm) and weaning weight gains because of milk production from the maternal EPD (+0.51 kg). On average, auctioned bulls have birth weight EPDs (+0.23) above the national average (+0.13).

⁵The herd book dataset has the EPD for all the registered bulls of the breed in Argentina; however, the dataset has no prices.

	Herd book average	Mean	SD	Min	Max
Bull Price (USD)		3029.75	1193.52	1274.36	10590.06
Weight (kg)		750.04	86.92	557.00	1035.00
Scrotal Circumference (cm)		39.31	2.33	34.00	48.00
Age (years)		2.24	0.37	1.65	3.34
Color (= 1 if red, = 0 if black)		0.33	0.47	0	1
Birth Weight (EPD)	0.13	0.21	0.85	-3.19	3.25
Weaning Weight (EPD)	8.20	9.54	3.93	-2.76	24.59
Final Weight (EPD)	17.59	19.93	7.26	-2.78	49.34
Scrotal Circumference (EPD)	0.91	0.93	0.60	-0.91	3.97
Maternal (EPD)	2.85	3.36	1.59	-4.13	9.02
$Stud_i$ (=1, if stud 2)		0.50	0.50	0.00	1.00
PP_i (=1, if Purebred)		0.86	0.34	0.00	1.00
Observations		723			

Note: Brangus population averages are extracted from the 2022 herd book.

Available at: <https://brangus.org.ar/programa-erbra>

Table 3.1. Descriptive Statistics

3.5 Results

This section presents the results from a hedonic pricing analysis of Brangus bull characteristics. Table 3.2 shows the coefficients from the hedonic regressions; we use four specifications with varying numbers of predictors. From columns (1) to (5), we add the dummies for studs, purebred year, and sire fixed effects. All the estimations have heteroskedasticity-robust standard errors. We show the R^2 and the adjusted R^2 ⁶. Column (5) is the complete and preferred specification because it has the highest R^2 .

We find that coefficient estimates associated with bulls' traits are not significantly affected by the incorporation of fixed effects. Weight, scrotal circumference, and coat color positively impact prices, which remain statistically significant after incorporating fixed effects. An additional kilogram increases bulls' price by 0.1%. The age coefficient implies that an extra year implies 7.9% lower prices per bull. An additional centimeter of scrotal circumference leads to 2.7% higher prices. Lastly, having a red coat implies 13% higher prices compared to a black Brangus bull. This shows that phenotypical characteristics are associated with bull prices individually.

In columns (2) to (5), the addition of fixed effects affects the statistical significance of EPDs. Results show that these attributes are sensitive. Specifically, an above-the-median birth weight EPD is associated with a bull price 4% lower than bulls with EPD values below the median. The negative association in this case is expected because a higher-than-average birth weight increases birth problems for ranchers. A higher

⁶We show the correlation matrix between explanatory variables and the Variance Inflation Factor from our estimates in Appendix A

weaning weight EPD is expected to be positively associated with price because cow-calf operators want to obtain heavy calves. The weaning weight EPD median coefficient shows that bulls with these values have 6% higher prices than bulls below the median. Genetic traits related to reproduction are statistically significant and imply a price 5% higher than below-the-median bulls, respectively (scrotal circumference and maternal EPDs).

	<i>Dependent variable:</i>				
	<i>log(price)</i>				
	(1)	(2)	(3)	(4)	(5)
Weight (kg)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Scrotal circumference (cm)	0.033*** (0.006)	0.033*** (0.006)	0.029*** (0.005)	0.025*** (0.005)	0.022*** (0.006)
Age (years)	-0.102** (0.040)	-0.119*** (0.040)	-0.121*** (0.038)	-0.146*** (0.043)	-0.105** (0.044)
Color (=1 if red)	0.171*** (0.023)	0.133*** (0.023)	0.185*** (0.021)	0.174*** (0.021)	0.110*** (0.039)
Birth weight EPD (= 1 if above median)	-0.075*** (0.024)	-0.076*** (0.023)	-0.056*** (0.021)	-0.052** (0.021)	-0.040* (0.022)
Weaning weight EPD (= 1 if above median)	0.061** (0.027)	0.068*** (0.025)	0.055** (0.023)	0.059*** (0.022)	0.055** (0.025)
Final weight EPD (= 1 if above median)	0.008 (0.028)	-0.008 (0.026)	-0.006 (0.023)	-0.013 (0.023)	-0.010 (0.026)
Scrotal circumference EPD (= 1 if above median)	-0.055** (0.024)	-0.070*** (0.023)	-0.032 (0.020)	-0.026 (0.021)	-0.038 (0.023)
Maternal EPD (= 1 if above median)	0.123*** (0.021)	0.091*** (0.021)	0.058*** (0.020)	0.044** (0.020)	0.031 (0.025)
$Stud_i$		0.153*** (0.022)	0.047* (0.025)	0.028 (0.026)	0.037 (0.055)
PP_i			0.116*** (0.027)	0.110*** (0.023)	0.084** (0.041)
Lot order FE				X	X
Sire FE					X
Observations	723	723	723	723	723
R ²	0.288	0.336	0.348	0.495	0.631
Adjusted R ²	0.279	0.326	0.338	0.482	0.535

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 3.2. Brangus bulls price analysis

Joint F tests show that while some traits might not be significant individually, they are relevant by group. Table 4 presents hypotheses testing to assess whether groups of traits are statistically significant using the estimates in column 5 from Table 3.2. This test is performed to verify whether the attributes have a joint association with prices. The results show that the coefficients related to phenotypical traits and EPDs are statistically significant by group at different p values.

	F-Stat	p-value	Result
Phenotypical traits	39.54	0.00	Significant at 1%
Production EPDs	2.22	0.08	Significant at 10%
Reproductive EPDs	3.85	0.02	Significant at 5%
All EPDs	2.61	0.02	Significant at 5%

Table 3.3. F-test by groups of coefficients

3.6 Discussion

Our results suggest different levels of analysis regarding the value of genetic traits. First, one feature of our analysis is that phenotypical traits are more relevant as decision variables than genetic traits for purchasing a bull. When taken individually, EPDs are generally less relevant for the price-setting process than phenotypical traits. The latter could be due to three reasons. First, the co-existence of ranchers interested only in phenotypical attributes and those looking at both traits. Second, the variability of EPD values is small, i.e., often estimated with low precision, which may mask the fact that the value of genetics may come from the stud’s reputation and not from specific traits. Third, there are different types of buyers, characteristics, and objectives, but our data does not contain information about them.

The first reason is related to investments in human capital. More sophisticated ranchers may emphasize reading EPDs before making decisions, while others may only look at visual characteristics and the bull’s behavioral data. The more dimensions or traits to look at before choosing bulls, the more investment in human capital is required. For instance, buying a bull with the potential of reducing the average weight of calves may come with correlations with other genetic traits that may express themselves in calves. For instance, a bull with birth weight EPD lower than the average could come with low weaning weight EPD. The investment might be affected negatively if the rancher had good weaning weight. Hence, buying bulls requires understanding complex genetic correlations.

Further, EPD traits do not vary significantly within studs, and their low precision could affect their value. The value of traits looks relatively small because the demand is conditional on the studs’ reputation. It is reasonable to think that farmers understand that studs have good-quality bulls. Conditional to the decision to participate in the auction, farmers would look at phenotypical and visual attributes because they already know that these studs have good genetics. K. V. Smith et al. (2023) studied Genomically-Enhanced EPDs (GE-EPDS), traits estimated using genomic tests on newborn calves, pedigree, and performance data to increase the accuracy of regular EPDs. They observe that all “traditional” EPDs positively and significantly affect the willingness to pay for genetic traits⁷. In contrast, GE-EPDs do not appear to convey any additional value. The authors argue that this result results from buyers deciding first on a specific value for certain EPDs to choose bulls. A possible explanation is that there is a lack of information on interpreting the accuracy of EPDs while such values are reported as necessary for ranchers’ decision-making process.

⁷Calving ease, weaning weight, docility, milk, mature and carcass weight, ribeye area, and weaned calf value.

Several articles have dealt with the limited statistical significance of EPD valuations on bull prices, at least in the American context. Boyer et al. (2019) used panel data of bull auction sales and EPDs from Tennessee and found that producers value traits directly related to their profits. Specifically, calving ease and a measure of projected growth, i.e., weaning weight EPD minus birth weight EPD, have a positive and significant effect on the bull’s price. However, the bulk of the impact on prices is explained by two observed traits: average daily gain and frame score.

Tang et al. (2020) showed evidence of grid pricing allowing for EPDs to signal the quality of a bull, but it is not conclusive since the sign of such changes is not monotonic across the period. The authors examine buyers’ marginal valuations of bull attributes over time as a response to grid pricing, a pricing scheme where a bull’s base price is subject to premiums or discounts for carcasses above and below a base or standard set of quality specifications. If buyers become more familiar with EPDs for specific traits, their valuations must increase accordingly. They show that birth weight EPD is the only trait systematically significant across periods (but with a negative coefficient). At the same time, maternal milk EPD is substantial mostly in the latter years, while marbling and ribeye area are primarily insignificant, particularly for latter periods.

Similarly, Thompson et al. (2022) looked at the determinants of bulls’ prices in traditional vs. genomic EPDs and observed characteristics with auction data. The influence of genomic traits is measured through an interaction term. In this data set, authors find a similar result: genomic-enhanced EPDs do not significantly affect the price beyond that of traditional EPDs. Moreover, only birth weight and ribeye area EPDs are found to have a statistically significant effect on the price of a bull. Also, the genomic-enhanced weaned calf value is statistically significant. The authors find that observed traits such as weight, age, scrotal circumference, and intramuscular fat explain most of the variation in bull prices.

The demand for traits is related to ranchers’ characteristics, which is a data limitation of our study. We cannot separate between breeding, cow-calf operators, or feeder cattle ranchers. The three types of buyers might have different preferences. For instance, while a cow-calf operator might be interested in low birth weight to reduce costs, feeder cattle ranchers might look at attributes related to rapid weight gain. We do not have demographic and production characteristics that might affect the demand for EPDs. With that data, we could further investigate and better understand the effect of farmer types and their preferences concerning genetic investment decisions.

3.7 Conclusion

In this paper, we investigate the determinants of Brangus bull prices in auctions from Northwestern Córdoba to determine which attributes are more highly valued by cattle ranchers. Understanding these relationships would help to know the potential for breeding and trait selection in areas where production conditions are arid or semi-arid. Moreover, it can increase beef production by incorporating new lands into production and productivity through technical efficiency. Our results show that observed traits have a statistically significant effect on the price of an animal, while Expected Progeny Differences mostly do not.

One possible interpretation of this result is that ranchers are still unfamiliar with EPDs and do not use them to base their breeding decisions. Similarly, it may also be the case that genetic traits are not as relevant for beef cattle as they are for dairy cattle; in the dairy industry, all relevant production traits are expressed in females only, and an estimation of the transmitting ability of such traits is needed to determine which traits will the bull’s daughters inherit. For beef cattle, these traits are expressed in both sexes, so a visual evaluation of the bull may provide a sufficient evaluation of its characteristics.

Coat color stands out among the set of observable traits; we argue that the demand for coat color could be associated with adapting to hotter and dryer climates in northern Argentina. Due to the competition for land with agricultural commodities, cattle have been adapted to these areas. However, from our results, we see that productive traits have heterogeneous demand, indicating that unobservable characteristics of the demand might be affecting the results.

A follow-up of this study will consist of a choice experiment with cattle ranchers to determine whether additional information about EPDs would change their decision when deciding which animal to buy. This experiment will also survey individuals about their knowledge of genetic traits and how to use them. Generating primary data would help us identify whether genetic trait information increases ranchers' willingness to pay and their perceptions about them.

CHAPTER 4

THE PRICE OF FAME: THE EFFECT OF AWARD-WINNING PEDIGREES ON THE PRICE OF LIVESTOCK GENETICS

4.1 Introduction

Dairy farmers periodically make breeding decisions to change the genetics of their herd by choosing a bull from the vast market of bull genetics available through artificial insemination. Each bull possesses a series of traits that may matter to dairy farmers, ranging from production traits such as milk yield to health traits such as fertility. Most of the extant literature, for example, Sy et al. (1997) and Richards and Jeffrey (1996), assume that dairy farmers choose health or production traits to maximize farm profits, which motivates a hedonic approach. In these analyses, they find that, in addition to health and production traits, the popularity of the bull is correlated to a higher price, even independent of other characteristics.

This analysis examines an overlooked aspect of these models: famous pedigrees. It is challenging to consider pedigree to determine the price of a bull since it is not a quantitative or qualitative trait that can be included directly in a linear regression for the animal’s price. However, the importance of pedigree manifests in several ways in the bull genetics market, one major being awards at cow shows.

The other missing factor is the actual impact of cattle shows on the price of bulls; for lack of a proper measure of its impact on genetic prices, economists have not analyzed how it affects them. In this paper, we use the results of the world’s most important dairy cattle show contest to measure its impact on the price of winners and their immediate relatives.

The World Dairy Expo meets every year and selects a “premier sire” that meets various show cattle criteria, namely physical traits, referred to as “type.” While high “type” sires often have traits that correlate to traits that matter to farm profits, the sire’s award-winning status may be a significant driver of price, even independent of traits considered crucial to profit. Our analysis determines whether award-winning pedigrees can explain price independent of production and health traits in the market for dairy bull genetics.

Understanding the relationship between pedigrees and the adoption of genetics is crucial for the dairy industry; the influence of pedigree can lead to repeated adoption of only a few genetic lines, which can negatively impact productivity because of inbreeding depression (Bjelland et al., 2013; G. Wiggans et al., 1995). Inbreeding depression is the increased prevalence of recessive (i.e., low frequency) traits caused by the mating of closely related individuals, and it is a pressing problem in dairy for breeds like the Holstein. For example, around 99% of US Holstein pedigrees can be traced back to only two ancestors: “Pawnee Farm Arlinda Chief” (born in 1962) and “Round Oak Rag Apple Elevation” (born in 1965) (Yue et al., 2015).

Thus far, very little has been written regarding the role famous pedigree plays in cattle genetics pricing. Previous literature on dairy cattle pricing almost exclusively focuses on the impact of production-relevant traits on prices and demand for cattle (Carvalho et al., 2022; Richards and Jeffrey, 1996; Schroeder et al.,

1992; Sy et al., 1997). Nevertheless, Rosen (1981) was the first of many papers to illustrate the special role that fame can play in markets. In the dairy genetics market, cattle shows like the World Dairy Expo can create “superstar” bulls whose notoriety may impact their price. Before data existed to estimate genetic traits, the dairy cattle show was the primary way to obtain an objective measurement of genetic quality. We hypothesize that this particular institution still exerts influence on markets today.

In this article, we use the pedigrees of winners and runner-ups of the “Premier Sire” category at the World Dairy Expo to disentangle the extent to which being related to such animals impacts animal’s prices in the periods immediately after winning the price. Testing the effect of winning a dairy cattle show on the price of a bull and its relatives is a practical and tractable way to illuminate how critical famous pedigrees are to farmers when choosing bulls.

We use data on genetics traits, pedigree, and prices of bull semen from the Council on Dairy Cattle Breeders for the winners and runner-ups from the World Dairy Expo from 2001, 2002, and 2003. We chose these years because there was a very high turnout of top sires, which allows us to identify pedigree effects on prices. Our econometric analysis shows mixed results. Only in some years do we see a positive price impact on the winner’s relatives compared to the relatives of the runner-up. Our analysis is complicated by the fact that runner-ups may have equal amounts of fame and, because of inbreeding, they are often related to the winner in some way. Thus, there is little variation to use to distinguish one bull’s relatives from the other. Future analysis must find ways to overcome these data limitations to gain a clearer understanding of how fame influences dairy bull pricing.

4.2 Background and Conceptual Framework

4.2.1 The World Dairy Expo

The World Dairy Expo (WDE) is an annual event held in October in Madison, Wisconsin, United States. Since its creation in 1967, it has been one of the world’s largest dairy-focused trade shows and cattle exhibitions. The expo allows dairy farmers and industry professionals to come together to showcase the latest innovations in dairy farming, technology, and genetics.

The WDE includes several activities, including cattle shows for most North American dairy breeds (Holstein, Ayrshire, Brown Swiss, Jersey, and Milking Shorthorn). This article will concentrate only on Holsteins, the most popular dairy cow breed in the United States. Any cow is eligible to participate if it has an Animal Identification Number (AIN) provided by the USDA to correctly identify the animal and link it to its production records from the National Dairy Herd Improvement Association (DHIA), as required for some classes.

There are several classes (categories) depending on the animal’s age (yearling, two, three, four, five, and six-year and older old cow), production, offspring, and date of birth. In every category, all winning cows are ranked from best to worst, and their sires are given points proportional to the rank of their daughters, for example, the “Best March Calf” category has 20 entries, then the sire of the winner gets 20 points, while the sire of the runner-up gets 19 points and the total number of accrued points decreases linearly until the 20th entry. On the other hand, the “Best Fall Yearling in milk”¹ had only 8 entries, and the winner’s sire was awarded 40 points, while the rest were given a decreasing amount of points by 2 points less for every position in the rank. Interestingly, if two or more daughters enter the top 20 (or top 8) in any given category, the sire accumulates his daughter’s points.

¹One year old cow

We focus our attention on the top category for bulls: *Premier Sire*. A bull gets points as a function of the points earned by his daughters and their number of occurrences in the competition, which are counted and inputted to him. Consequently, a bull gets more points due to his daughter’s performance and how many compete in any given year. The Premier Sire does not participate directly in the WDE and can still win points even after passing away. For example, the bull “Braedale Goldwyn” won the Premier Sire category from 2008 to 2019, despite having died in 2008. This post-mortem winning streak is because his daughters won in almost every category (Steyn et al., 2023), Table 4.1 shows the winners and runner-ups for 2000 to 2010, there is a significant turnover from 2000 to 2004, but it disappears after 2004.

Year	Winner	Runner-up
2000	Donnandale Sky Chief	Boulet Charles
2001	Marcrest Encore	Duregal Astre Starbuck
2002	Lystel Leduc	Boulet Charles
2003	Lystel Leduc	Regancrest Elton Durham
2004	Regancrest Elton Durham	Lystel Leduc
2005	Regancrest Elton Durham	Lystel Leduc
2006	Regancrest Elton Durham	Regancrest Dundee
2007	Regancrest Elton Durham	Regancrest Dundee
2008	Braedale Goldwyn	Regancrest Dundee
2009	Braedale Goldwyn	Regancrest Dundee
2010	Braedale Goldwyn	Wilcoxview Jasper

Table 4.1. World Dairy Expo Premier Sire winners and runner-ups 2000-2010

Dairy farmers have two motives for weighing the importance of a pedigree in their adoption decision. First, the importance of dairy cattle shows in the dairy industry may implicitly bias farmer decisions towards cows that are likely to win shows, much in the same way, that consumers factor the popularity of a brand into their purchasing decisions.

Second, some dairy farmers selling heifers can increase their value by using pedigrees in high demand. Thus, the popularity of a pedigree can matter to farmers for both utility and profit, even independent of its production or health traits.

Third, given the popularity of cattle shows in general, and the WDE in particular, dairy farmers are more likely to trust the opinion of WDE judges more than a wide array of genetic traits that can be very difficult to interpret without prior knowledge of genetics.

4.2.2 Measurement of genetic traits

Unlike their beef counterparts, dairy farmers do not usually keep bulls within their herds; they routinely buy bull semen from genetics companies to impregnate their cows. Dairy cattle present an additional complication: all relevant genetic traits are expressed only in females, and it is not possible to directly measure the contribution of a sire (father) to the productivity of his daughters. This problem was circumvented using *large-scale genetic evaluations* (Bourdon, 2000, p. 262) that measure cow’s observed traits such as milk output, butterfat production, calving ease, or pregnancy rate are measured (evaluated) three times a year. After these data are gathered, the paternal contribution to observed values can be estimated using a statistical model incorporating the measured traits with the animal’s pedigree and other fixed effects.

The predicted values from such a model are called **Predicted Transmitting Ability (PTA)**; we call these values “genetic traits” and are public knowledge since the last measured values are usually reported when advertising a bull.

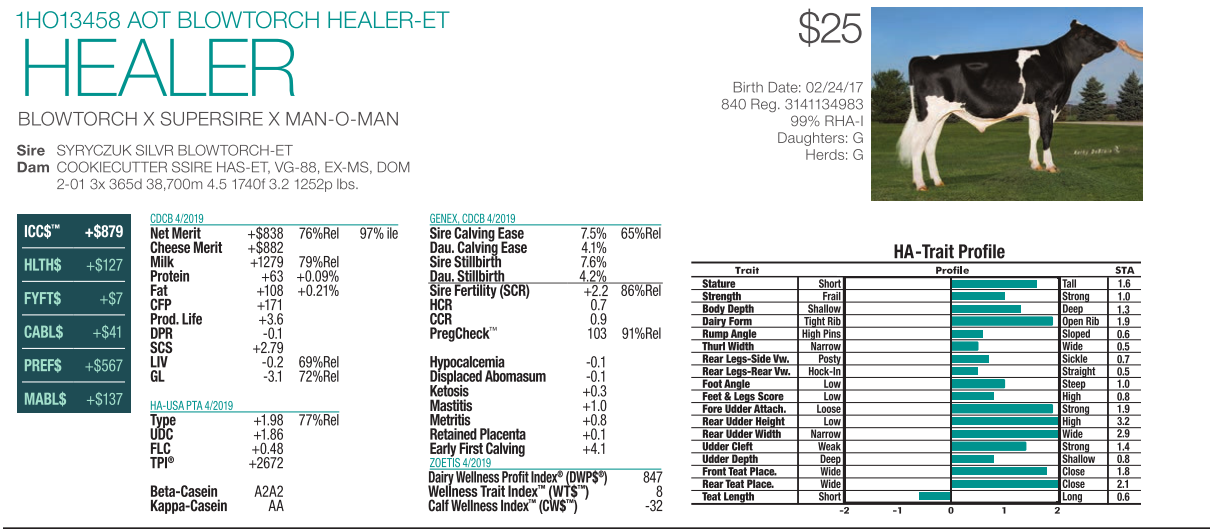


Figure 4.1. Sire Summary

Figure 4.1 is the summary for a bull called *Healer*; this document contains information on the animal’s pedigree, that is, the identity of his sire (father), and dam (mother.) It also contains the last estimated values for the genetic traits he can transmit to his daughters. In the United States, dairy cattle genetic evaluations are reported by the **Council on Dairy Cattle Breeding (CDCB)**, the date of measurement (April 2019 in this case.) The summary also contains a list of selection indexes (Melton et al., 1994), such as Net Merit, Cheese Merit, or Total Productivity Index (TPI); these are weighted averages of trait values measured in US dollars that capture the value of each animal in a single value that can be used to rank sires at any given time.

Pedigree plays a crucial role in the transmission of prestige. As long as a farmer knows an animal’s sire, it is easy to tease out the degree of relatedness to a World Dairy Expo winner. Similarly, since genetic traits are complex to contextualize by means other than selection indexes, it is the case that a dairy farmer values prestige over genetic traits as a proxy for genetic quality.

4.2.3 Input Characteristics Model

Bull genetics is a crucial input in the dairy industry; its value is a function of a set of characteristics (genetic traits) that are useful to dairy production (fat/protein content of milk) or help reduce costs (somatic cell score or udder health traits). These traits cannot be purchased separately; they are “bundled” in the bulls’ genotype; consequently, the price of genetics is a function of all relevant traits.

“Prices and Demands for Input Characteristics” (1976) and Melton et al. (1994) present the pricing framework relevant to this problem, called the **Input Characteristics Model**, an adaptation of the neoclassical profit-maximization firm where the output is a function of a set of inputs and their characteristics.

In this framework, a dairy firm produces one output (milk), denoted by y , which is a function of a set of n inputs p genetic traits that enter the production function as an index $z(q_1, \dots, q_p)$ such that:

$$y = f(z(q_1, \dots, q_p)) \quad (4.1)$$

Note that the firm cannot choose the composition of $z(\cdot)$ as it is supplied jointly by a genetics company. Let w_i denote the price of input i and r the price of one unit of genetic material (e.g., a straw of bull semen.) Since $z(\cdot)$ is assumed to be a linear homogeneous index:

$$z(q_1, \dots, q_p) = \sum_{j=1}^p \frac{\partial z}{\partial q_j} q_j \quad (4.2)$$

Richards and Jeffrey (1996) and Walburger (2002) argue that the price paid by a dairy farmer depends on the marginal contribution of a specific bull to the farm's genetic stock; hence, the price of a semen straw must be equal to the sum of the values of all its genetic characteristics. Breeding decisions in the present are made to maximize the present value of a farm's herd, then the marginal contribution of a bull can be measured through the increment of the herd's net present value.

The value of each trait q_j is determined by its marginal impact on the price associated with index z since the index is homogeneous of degree one, the price paid for a given index level is:

$$p_z = \sum_{j=1}^p p \frac{\partial f}{\partial z} \frac{\partial z}{\partial q_j} q_j \quad (4.3)$$

Where $p \frac{\partial f}{\partial z}$ is the value of the marginal product of z . The partial derivatives $\frac{\partial z}{\partial q_j}$ are assumed to be constant as a consequence of the breeding process. Animals are selected based on a weighted average of their traits, and these weights are the economic values attached to each trait (Hazel, 1943). There are several indexes frequently used by the dairy industry, such as the Total Performance Index (TPI), Net Merit (NM), and Grazing Merit (GM).

Then, a bull's value is equal to the sum of the values of all its traits, where the price of each characteristic is equal to its market value, defined as the price of the bull times the economic weight of each trait. Each r_k is the implicit price of trait k ; since there are no individual markets for each trait, there can be no explicit market prices.

Equation 4.3 shows that there is a linear relationship between the price of a bull's genetics and the level of his traits; following Schroeder et al. (1992), we can estimate the equation:

$$price_i = \beta_0 + \beta_1 q_{1i} + \dots + \beta_n q_{ni} + \epsilon_i \quad (4.4)$$

Where $price_i$ is the price of bull i , and q_{ki} is the k -th genetic trait, then each β_k is the implicit price of all traits. Most applications of hedonic pricing for genetic traits have focused on estimating equation 4.4 with cross-section data. Schroeder et al. (1992) is the earliest example of an article that uses the "Prices and Demands for Input Characteristics" (1976) framework to derive a hedonic pricing scheme using 213 Holstein sires for the January 1990 evaluation. This paper uses 16 traits, two aggregate indexes, and firm fixed effects; the authors find that all productivity traits (milk, protein, fat, and reliability) are significant, as well as two udder type traits and leg width.

Another early example is Madalena et al. (1996), which uses data on semen prices and traits for a sample of 32 *Gyr* and 73 *Nellore* sires from Brazil in 1993. This paper includes two other elements not present in most of the related literature: pedigree and dummies for winning cattle shows; the former is included via regressing the price on the degree of relatedness to the "premier sires" of each breed to control for the effect

of winning cattle shows a dummy equal to 1 for sires that were either winners or runner-ups of both national and state shows are included. The authors find that most of the effect is driven by the animals’ pedigree and by the winner dummy rather than their breed, weight, or body measures; unfortunately, this article does not measure PTAs for productivity traits.

Carvalho et al. (2022) use a sample of 554 *Gyr*, *Guzerá* and *Girolando* dairy bulls from Brazil that were marketed in June 2019; they use PTAs for Milk yield, age at first birth, fat, protein, and solids, they also include genetic company, breed, and sexed semen dummies. Across breeds, only the milk PTA has a significant coefficient; concerning other variables, company, breed, and sex dummies are also significant.

Another line of research similarly uses surveys to choice experiments, Sy et al. (1997), for example, surveyed 253 dairy farmers in Manitoba, Canada, in August-October 1993. This method does not require actual observed semen prices, but instead of providing a partial effect of traits on prices, it allows the researchers to estimate a partial ordering (ranking) of preferences using an Ordered Probit model. Researchers were able to separate the effects on different segments of the cattle market (bulls and steers, breeders and feeders); in this sample, weaning weight, temperament, and calving ease are significant in determining rankings.

4.3 Data and Empirical Approach

4.3.1 Data description

Our data comes from the **National Association of Animal Breeders** (NAAB), a trade group of dairy genetics companies. Over the last 20 years, NAAB has reported the entire set of genetic traits for the dairy bulls marketed by its member companies and a posted price for advertising (Hutchins and Hueth, 2023).

To market a bull in the United States, the animal must be registered and tested (“evaluated”) for the presence of a set of around fifty different genetic traits. The Council on Dairy Cattle Breeding (CDCB) calculates and publishes these traits thrice yearly (May, August, and November). Most U.S. traits are expressed as *Predicted Transmitting Abilities*, or PTAs, that measure the predicted difference of an animal’s offspring from the average population. In most cases, PTAs are expressed in the units used to measure the trait. It is usual to summarise the data on a set of traits into an index, a weighted average of traits used to rank bulls (Lush, 1937.)

We use these evaluation and price data for 3,246 bulls from 2001 to 2003, totaling 531 observations. The bulls in our data set are evaluated for 22 production, health, type, and reproductive traits (PTAs). Production traits predict future milk, fat, and protein yields of mature daughters in pounds compared to cows of the same breed born in the same base year. Health traits measure the heritability of a group of measures of life expectancy of cattle. Fertility traits capture a cow’s ability to conceive as well as the likelihood of experiencing spontaneous abortions. Finally, type traits measure a series of visual traits of cows, which can be associated with the “beauty” of the cow and, therefore, highly regarded in cattle shows.

Our data set also contains information about each bull’s male line ancestry, some dating back to the mid 20th century (10 generations); this allows us to build up the pedigree for each bull in the sample. It is relatively easy to build up the pedigree by linking sons to sires and grandsires until the last recorded generation is found, however, this leaves us without a way to link animals to their brothers, (paternal) cousins, and half-brothers. We use the *gggroups* pedigree building R package (Nilforooshan and Saavedra-Jiménez, 2020) to get around this issue.

Finally, the data on World Dairy Expo winners from 2000 is publicly available on the expo website; we use the data for the **Premier Sire** category. For a bull to earn high enough points, he must either be the

sire of a cow that earns top positions in several categories or have a large enough number of daughters in a similar position.

4.3.2 Empirical framework

Our objective is to determine the impact of an award-winning pedigree on the price of bull genetics, conditional on the levels of their genetic traits. Similarly, we can think of an alternative outcome: the number of daughters, unlike the asked price, proxies the effect of the WDE on the demand for animal genetics.

To do so, we will exploit the high turnover rate between winners and runner-ups between 2001 and 2003, but such a situation did not repeat after 2004, due to a long reign of two “supersires”: *Regancrest Elton Durham* and *Braedeale Goldwyn*, the latter extended his winning streak into 2019. Similarly, our hedonic pricing model requires that most bulls have a reported price included in their sire summary, but after 2010, most companies stopped reporting such asked prices in the summary.

Incorporating an animal’s pedigree into a linear model is problematic because it cannot be easily converted into a continuous or categorical variable. We propose a different approach since the World Dairy Expo is a cattle show where cows are evaluated regarding their observed traits. Then, a fraction of their total points are inputted to their sires, and the price difference between the relatives of the winner and the runner-up’s relatives, in a context of high turnover, must be explained chiefly by random factors.

Using genetic traits in a hedonic pricing model presents several issues; the most important one is collinearity (Thompson et al., 2022); most traits are highly correlated with others in the same class (production, health, and longevity); as a consequence, coefficient estimates are not reliable and standard errors are inaccurate.

Figure C1 shows the degree to which all traits correlate with each other. The price of a bull is weakly correlated with most traits, except for the “type”. Similarly, all production traits (milk, fat, and protein) and reproductive traits (daughter pregnancy rate, heifer, and cow conception rate) are highly and positively correlated with each other to a high degree. The “type”, also has a positive correlation with all physical traits and udder size measures, as expected. On the other hand, the expected number of daughters is not correlated to any other measure to a significant degree.

Conversely, Figure C1 shows that there are at least four groups of highly correlated traits:

1. **Production traits:** milk, fat and protein yields.
2. **Reproductive traits:** daughter pregnancy rate, heifer and cow conception rates, sire and daughter stillbirth.
3. **Health traits:** productive life, somatic cell score, livability.
4. **Visual aspect traits:** type, stature, strength, dairy form, leg, and udder characteristics.

To deal with this issue, we run a Lasso selection model, assuming that the actual model is sparse, that is, most coefficients can be set to zero with little effect on the R^2 of the model. The Lasso (Hastie et al., 2015) is a generalization of the least squares estimator that minimizes the mean squared error of the objective function subject to a L^1 penalty of the form:

$$\min_{\beta \in R^p} \sum_{i=1}^N \left(y_i - \sum_{j=1}^p x_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^p |\beta_j| \quad (4.5)$$

Where the parameter λ is selected by 10-fold cross-validation. Figure C2 shows the optimal value of the penalty coefficient (λ) in a lasso model. We will measure the impact of all PTAs on the price of the animal's semen.

In this case, only three PTAs have non-zero coefficients (fat percent, type, and rear udder width). As expected, since most traits are highly correlated, the bulk of the variation can be attributed to a reduced number of variables.

Those traits are defined as:

- *Type* is the difference in score classification points relative to the herd average; it predicts the animal's genetic ability to transmit to the offspring's final classification or appraisal score in the breed's program as evaluated by the Holstein Association.
- *Fat Percent* is the expected second-lactation fat production in pounds compared to the breed average. This trait predicts the animal's genetic ability to transmit concentration of milk fat; expressed as percentage points compared to the breed base.
- *Rear Udder Width* is the width of the rear udder where the udder attaches to the body, measured in inches.

Thus, our identification strategy depends on the assumption that conditional on a set of PTAs, the price difference between the two groups must be random. Consequently, on average, the runner-up's relatives are a good counterfactual for the winner's relatives. Interestingly, in terms of genetic traits, there is a significant difference for many traits, but with the opposite sign to what one would expect, that is, runner-up's relatives tend to have higher levels than winner's relatives.

We are interested in the time impact of winning the World Dairy Expo on the average price across the lines of winners and runner-ups; as an event study, we stack all three years (2001-2003) into one single time index ranging from three periods before to two periods after. Since the WDE takes place in October, the effect must be noticeable on the immediate subsequent evaluation (November); we then define the base period as the August evaluation.

Our empirical specifications are:

$$price_{it} = \beta_0 + \beta_1 treat_i + \beta_2 post_t + \beta_3 treat_i \times post_t + \varepsilon_{it} \quad (4.6)$$

$$price_{ikt} = \sum_{t=-3}^3 \beta_t treat_{ikt} + \gamma_k + \tau_t + \varepsilon_{ikt} \quad (4.7)$$

$$price_{ikt} = \sum_{t=-3}^3 \beta_t treat_{ikt} + \mathbf{X}_{it}\Gamma + \gamma_k + \tau_t + \varepsilon_{ikt} \quad (4.8)$$

Where $price_{ikt}$ is the price of a unit of animal genetics from bull i that belongs to family k in period t . The variable $treat_{ikt}$ is the treatment dummy, equal to 1 if k are the winner's relatives and 0 otherwise. $\mathbf{X}_{it}$ is the matrix of PTAs, γ_k and τ_t are family and period fixed effects.

4.4 Results

We run three distinct models, one for each year. We define the pre-intervention period as all evaluations released between February and August and the post-intervention period as the evaluations from November to

May of the following year.

	semen_price		
	2001	2002	2003
	(1)	(2)	(3)
winner $\times$ post = 1	1.379** (0.060)	0.209 (1.573)	-1.130** (0.023)
Observations	754	1,357	1,553
R ²	0.22872	0.24063	0.34103
winner fixed effects	✓	✓	✓
post fixed effects	✓	✓	✓
sire fixed effects	✓	✓	✓
breed fixed effects	✓	✓	✓
status fixed effects	✓	✓	✓

Table 4.2. Price equations

Table 4.2 shows that winning the WDE leads to an additional dollar of price for the 2001 winner but to a 30-cent lower price in 2002 and a further 1.2 dollars less in 2003.

	semen_price		
	2001	2002	2003
	(1)	(2)	(3)
winner $\times$ time = -2	-0.561 (1.160)	-0.153 (2.342)	-0.310 (1.594)
winner $\times$ time = 0	0.198 (1.183)	1.243 (1.919)	-0.053 (1.424)
winner $\times$ time = 1	0.198 (1.183)	0.691 (2.707)	-1.105 (1.786)
winner $\times$ time = 2	2.319 (2.058)	0.830 (2.693)	-1.604 (1.825)
winner $\times$ time = 3	1.563 (1.814)	2.309 (2.148)	-1.097 (1.894)
Observations	754	1,357	1,553
R ²	0.23019	0.24177	0.34119
sire fixed effects	✓	✓	✓
period fixed effects	✓	✓	✓
breed fixed effects	✓	✓	✓
status fixed effects	✓	✓	✓

Table 4.3. Price equations (multiple periods)

Table 4.3 shows the results of running three separate models per year, using the period before the intervention as the base, we have that none of the coefficients are significant, and that the sign of the effect changes across years, for 2001 and 2002, all coefficients for the post-WDE period are positive, but these signs reverse for 2003.

4.4.1 Multiple-period differences-in-differences

We now turn to the model where we decompose the variations by quarter, as detailed in equation (6). All specifications include fixed effects at the group level (winner or runner-up relative), sire, breed, and status. There are two possible status classes: “Active” for dairy bulls with 10 or more milking daughters in the USA and “Foreign” for bulls with 10 or more milking daughters outside of the USA and less than 10 in the USA, with a CDCB genetic evaluation and has semen available for sale.

Table 4.3 shows that post-intervention coefficients are negative (but not significant) for 2001, but positive for 2002 and 2003. Winning the WDE increases the premium by around 8 dollars on average for the winner’s relatives.

4.4.2 Aggregate differences-in-differences

Finally, we analyze the full sample on a single model by aggregating all pre and post-intervention periods into one model. We test three different specifications for equation (7): first one without covariates, except for the interaction of the treatment dummy and time period, then one including all traits, and the third specification includes only those traits selected by Lasso, as detailed earlier. Finally, we replace genetic traits with three selection indices, which are weighted averages of a specific subset of traits designed to rank bulls:

1. **Total Performance Index (TPI)**: a summary index that weights production, health, and conformation PTAs to provide a ranking of Holsteins based on genetic merit.
2. **Net Merit (NM)**: index of several PTAs that are of economic importance, intended to rank animals with a higher dollar amount that will make more money during a lifetime for a producer on a purely genetic basis.
3. **Cheese Merit (CM)**: factors in the same traits as Net Merit, but Cheese Merit assigns a negative economic weight on PTA Milk and places more emphasis on Protein Pounds.

	semen_price			
	No covariates	All traits	Selected traits	Indexes
	(1)	(2)	(3)	(4)
winner $\times$ time = -2	-0.520* (0.049)	-0.510 (0.091)	-0.507* (0.056)	-0.571** (0.042)
winner $\times$ time = 0	-0.516*** (0.006)	0.000 (0.000)	-0.381** (0.008)	-0.316*** (0.005)
winner $\times$ time = 1	-1.282*** (0.012)	-1.170* (0.165)	-1.386*** (0.005)	-1.361*** (0.019)
winner $\times$ time = 2	-1.539** (0.059)	-1.634* (0.217)	-1.649** (0.042)	-1.576** (0.060)
winner $\times$ time = 3	-0.945** (0.035)	-1.429* (0.165)	-1.110** (0.029)	-1.082** (0.045)
Observations	3,664	1,942	3,620	3,611
R ²	0.89102	0.93923	0.90622	0.90879
winner fixed effects	✓	✓	✓	✓
time fixed effects	✓	✓	✓	✓
bull fixed effects	✓	✓	✓	✓
breed fixed effects	✓	✓	✓	✓
status fixed effects	✓	✓	✓	✓

Table 4.4. Price equations for the aggregate model

Results from the aggregated model can be found in Table 4.4, coefficient signs are negative, in accordance with the separate models by year, we see that winning the WDE does not lead to higher prices in posterior periods. These results are robust across specifications.

Our results suggest that there is no clear relation between winning the WDE and an increase in the price of semen for award-winning bulls and their close relatives. This can be explained by the fact that winning bulls are already “popular”, for a sire to earn points in the Premier Sire category, he must have either fathered a large enough number of daughters that participated in the contest (quantity), or one of his daughters must also be a superstar that earned high scores in several categories (quality). Then, it is possible that winning the WDE does not have any sizeable impact on both outcomes, since for this effect to exist, both winners and runner-ups should be (relatively) unknown sires, but if that were the case, it would be unlikely that they had a sizeable progeny that would allow them to have a large number of participating daughters.

4.5 Conclusion

Dairy farmers make breeding decisions to alter the gene pool of their herd by choosing a bull based on his traits, but those traits are expressed only in females, for that reason, those characteristics must either be evaluated on the animal’s daughters or through genomic testing.

Before discovering such traits, farmers chose a bull based on his ancestry or physical characteristics (“type”.) This article argues that because these methods are still in use by the dairy industry, there may be a positive effect on the price of a bull after winning the Premier Sire category in the World Dairy Expo that also “trickles down” to all his relatives whose genetics are supplied in the market at the same time.

Previous articles focused only on genetic traits (PTAs) and observable traits such as an animal’s coat color, weight, or breed but did not consider pedigree for their analyses. One likely reason is that pedigree cannot be easily converted into a continuous or discrete variable that can be used as an independent variable in a hedonic regression model. Also, until recently, obtaining pedigree data in a form that could be used for economic analysis was challenging.

We used the WDE winners and runner-ups for three years (2001-2003) when there was a significant turnover of bulls at the top of the ranking. The point difference between the first and the second position was relatively small, thus providing us with a natural experiment to test the impact of pedigree on the price of bull genetics. The total points accrued by any contestant are a function of his daughter’s performance; hence, whatever matters for the WDE judges matters for the whole market. Consequently, cattle shows play a crucial role in influencing how farmers think of genetics by explicitly showing the link between traits and observed characteristics of an animal.

We show that the difference in prices between the winner’s and runner-up’s lines does not match the expected behavior. We assumed that winning the WDE would lead to a price premium for the winner’s line. However, the results contradict our expectations. The observed difference in prices was negative, and we interpret this as evidence that winning the WDE does not lead to additional gains in information since all participants are already “celebrity” sires, i.e., well-known animals that have already fathered thousands of daughters.

Our empirical application includes a set of three stacked treatments, which may lead to negative weights for the coefficient estimates (Callaway and Sant’Anna, 2021; Goodman-Bacon, 2021). When there are more than two periods stacked and a non-treated group, some treated groups may have a negative weight for some periods; a future version of this article will deal with this issue in greater detail.

APPENDIX A: APPENDIX TO CHAPTER 2

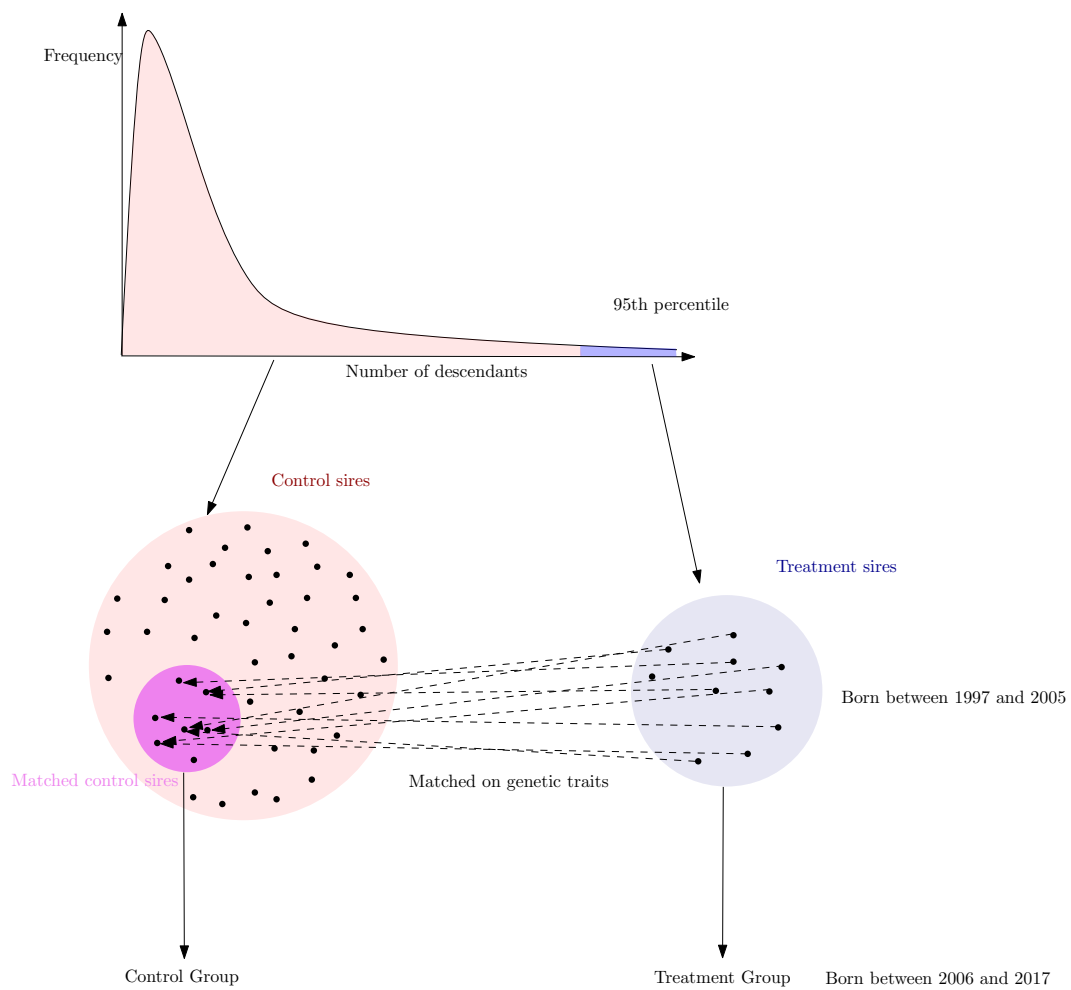


Figure A1. Sample selection mechanism

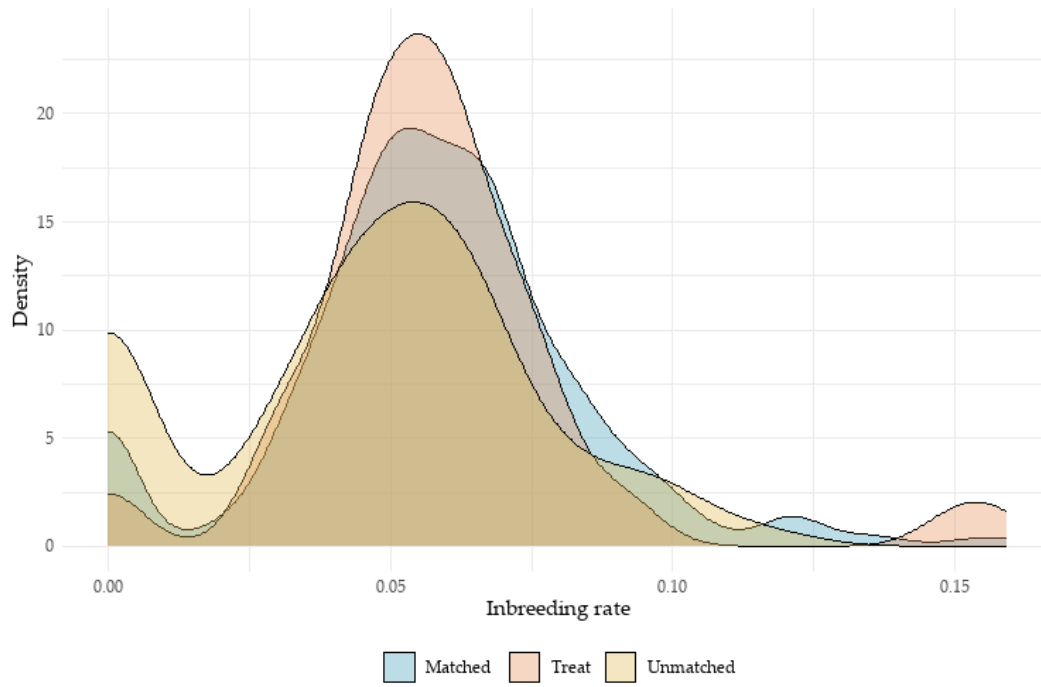


Figure A2. Distribution of Inbreeding rates across pre-treatment groups

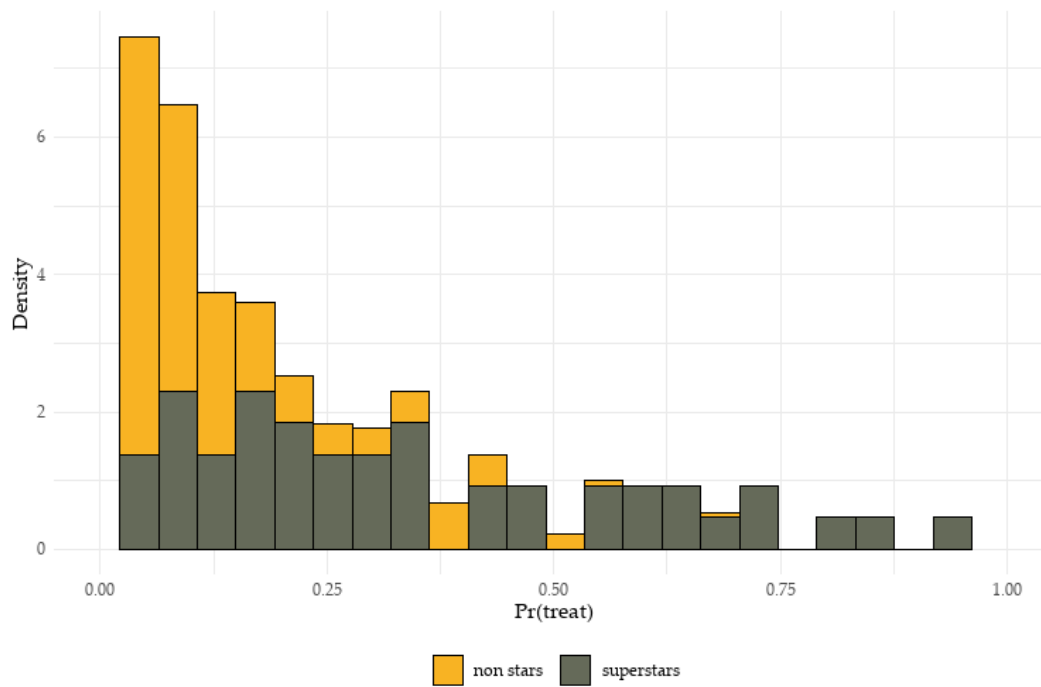


Figure A3. Propensity score by class of bull

Trait	Units	Value (\$/PTA unit)
Milk	Pounds	0.002
Fat	Pounds	4.18
Protein	Pounds	4.67
Productive Life	Months	34
Somatic Cell Score	Log	-74
Body Weight Composite	Composite	-45
Udder Composite	Composite	19
Foot/Leg Composite	Composite	3
Daughter Pregnancy Rate	Percent	11
Calving Trait Subindex	USD	1
Heifer Conception Rate	Percent	1.1
Cow Conception Rate	Percent	2.2
Livability	Percent	9.8
Health Trait Subindex	USD	1.0
Residual Feed Intake	Pounds	46.2
Effective Calving	Days	2.1
Heifer Livability	Percent	5.0

Table A1. Economic values for **Net Merit**

Variable	Control mean (n = 126)	Treatment mean (n = 337)	t.statistic	p-value	df
Production traits					
Milk	1366.841	1326.074	1.996	0.047	194.194
Fat	1.977	1.131	11.393	0.000	240.576
Protein	1.287	1.580	-3.785	0.000	273.710
Longevity and Health traits					
Somatic cell score	37.820	43.231	-2.207	0.028	217.114
Productive life	1.999	0.813	11.212	0.000	174.613
Conformation traits					
Type	2.027	1.809	2.619	0.009	299.167
Stature	1422.469	1380.570	0.686	0.493	289.615
Strength	1.291	1.471	-2.100	0.037	244.177
Body depth	45.992	45.982	0.005	0.996	207.117
Dairy form	2.584	1.433	9.934	0.000	188.248
Rump angle	-0.539	-0.311	-2.555	0.011	261.210
Thurl width	2.569	2.331	3.072	0.002	286.063
Rear legs side view	2.215	2.045	2.466	0.014	304.506
Rear legs rear view	0.025	0.088	-0.564	0.573	264.047
Foot angle	3.186	3.286	-5.608	0.000	339.823
Fore udder	2.388	2.310	0.976	0.330	288.241
Rear udder height	1.820	0.982	9.429	0.000	223.345
Rear udder width	1.346	1.221	1.194	0.234	247.251
Udder cleft	0.297	0.004	3.112	0.002	348.119
Udder depth	2.567	1.261	15.043	0.000	330.459
Teat front place	2.369	1.963	9.647	0.000	281.936
Teat length	2.085	1.695	3.837	0.000	243.773
Other					
TPI	1.783	1.679	0.985	0.326	288.011
Semen price	24.961	21.134	4.034	0.000	185.181

Table A2. Balance table for all variables

	No covariates		inbreeding	
	(1)	(2)	Traits (3)	Traits and interactions (4)
treat $\times$ post = 1	1.453*** (0.4250)			
treat $\times$ yob = 2005		0.0944 (0.2975)	0.0489 (0.2989)	0.0252 (0.2829)
treat $\times$ yob = 2006		-0.2183 (0.3067)	-0.1182 (0.2933)	-0.2789 (0.2534)
treat $\times$ yob = 2007		0.2328 (0.3038)	0.2826 (0.2676)	0.2848 (0.2534)
treat $\times$ yob = 2008		0.4035 (0.3148)	0.3246 (0.2798)	0.3294 (0.2425)
treat $\times$ yob = 2010		-0.4589 (0.2827)	-0.4721* (0.2826)	-0.5583** (0.2667)
treat $\times$ yob = 2011		-0.0164 (0.3599)	-0.0455 (0.2596)	-0.1671 (0.2511)
treat $\times$ yob = 2012		1.224** (0.4906)	1.252** (0.5633)	1.151* (0.5981)
treat $\times$ yob = 2013		1.433*** (0.4124)	1.083** (0.4631)	0.9980** (0.4725)
treat $\times$ yob = 2014		1.780*** (0.3017)	1.600*** (0.3514)	1.501*** (0.3068)
treat $\times$ yob = 2015		1.637*** (0.3453)	1.363*** (0.4159)	1.255*** (0.4256)
treat $\times$ yob = 2016		2.414*** (0.2424)	1.766*** (0.2697)	1.710*** (0.2564)
treat $\times$ yob = 2017		2.849*** (0.4456)	2.770*** (0.4429)	2.734*** (0.4507)
treat $\times$ yob = 2018		3.417*** (0.4530)	3.576*** (0.4566)	3.515*** (0.4847)
treat $\times$ yob = 2019		2.005** (0.8611)	2.071** (0.8827)	2.022** (0.8785)
p-value for nonzero pre-effect		0.416	0.498	0.107
Observations	15,491	15,491	15,491	15,491
R ²	0.16988	0.47369	0.49942	0.50471
Within R ²	0.00702	0.01235	0.04150	0.05163
treat fixed effects	✓	✓	✓	✓
post fixed effects	✓			
yob fixed effects		✓	✓	✓
parent_company fixed effects			✓	✓

Table A3. Difference-in-Differences estimates

	No covariates		inbreeding	
	(1)	(2)	Traits (3)	Traits and interactions (4)
treat $\times$ post = 1	1.217*** (0.4671)			
treat $\times$ yob = 2005		-0.0189 (0.2991)	-0.0863 (0.2966)	-0.1019 (0.2842)
treat $\times$ yob = 2006		-0.1702 (0.3088)	-0.1060 (0.2951)	-0.2813 (0.2520)
treat $\times$ yob = 2007		0.1376 (0.3131)	0.1949 (0.2826)	0.1698 (0.2699)
treat $\times$ yob = 2008		0.3104 (0.3286)	0.2154 (0.2896)	0.2295 (0.2508)
treat $\times$ yob = 2010		-0.5930** (0.2916)	-0.6049** (0.2935)	-0.6837** (0.2705)
treat $\times$ yob = 2011		0.0069 (0.3567)	-0.0263 (0.2577)	-0.1519 (0.2509)
treat $\times$ yob = 2012		1.115** (0.4877)	1.134** (0.5545)	1.052* (0.5685)
treat $\times$ yob = 2013		1.131** (0.4479)	0.9319** (0.4026)	0.8830** (0.3758)
treat $\times$ yob = 2014		0.8006 (0.6513)	0.6818 (0.6000)	0.6248 (0.5656)
treat $\times$ yob = 2015		0.5968 (0.8029)	0.3803 (0.8051)	0.3005 (0.7810)
treat $\times$ yob = 2016		0.9684** (0.3983)	0.6753** (0.3389)	0.5804* (0.3175)
treat $\times$ yob = 2017		2.558*** (0.5683)	2.484*** (0.5576)	2.437*** (0.5518)
treat $\times$ yob = 2018		3.407*** (0.4524)	3.567*** (0.4620)	3.493*** (0.4911)
treat $\times$ yob = 2019		1.995** (0.8613)	2.083** (0.8848)	2.035** (0.8804)
p-value for nonzero pre-effect		0.772	0.799	0.414
Observations	15,352	15,352	15,352	15,352
R ²	0.16918	0.47434	0.50014	0.50506
Within R ²	0.00575	0.01144	0.04123	0.05067
treat fixed effects	✓	✓	✓	✓
post fixed effects	✓			
yob fixed effects		✓	✓	✓
parent_company fixed effects			✓	✓

Table A4. Difference-in-Differences estimates from alternative scenario

APPENDIX B: APPENDIX TO CHAPTER 3

Table B1 shows the correlation coefficients for the bulls' traits variables. The correlation matrix shows only two correlations above 0.5: i) median weaning weight EPD and median final weight EPD, and ii) weight and age.

Variables	Weight	SC	Age	Red coat	Median BW EPD	Median WW EPD	Median FW EPD	Median SC EPD	Median Maternal EPD
Weight	1.000								
SC	0.396	1.000							
Age	0.624	0.265	1.000						
Red coat	-0.036	-0.042	-0.083	1.000					
Median BW EPD	0.175	0.130	0.010	0.107	1.000				
Median WW EPD	0.169	0.217	-0.037	-0.037	0.397	1.000			
Median FW EPD	0.211	0.205	-0.053	-0.089	0.367	0.638	1.000		
Median SC EPD	0.024	0.413	-0.145	0.002	0.101	0.250	0.303	1.000	
Median Maternal EPD	0.002	-0.004	-0.003	0.122	0.018	-0.104	-0.018	0.048	1.000

Table B1. Correlations coefficients between explanatory variables

We checked for multicollinearity using the Variance Inflation Factor (VIF). Table A2 shows variance inflation factors for all variables in all five models. Goldberger (1991) shows that the standard error of any given coefficient j can be written as:

$$\hat{\beta}_j = \frac{1}{1 - R_j^2} \frac{s^2}{(n - 1)Var_j}$$

Where Var_j is the sample variance of variable j and R_j^2 is the r-squared coefficient from a regression of j against all other regressors included in the model, hence $1/(1 - R_j^2)$ is the variance inflation factor of regressor j . The higher the variance of j explained by all other variables, the higher the VIF; these are the values reported in columns (1) to (5). For fixed effects, this value is the Generalized Variance Inflation factor (GVIF), equal to the average VIF for all levels corrected by the degrees of freedom:

$$GVIF = VIF^{1/2 \times df}$$

The rule of thumb is generally a VIF above 10, which could indicate strong multicollinearity. The average of all our estimates ranges between 1 and 3.7.

Variance Inflation Factors	(1)	(2)	(3)	(4)	(5)
Weight	1.971	1.974	1.974	1.475	1.692
Scrotal circumference	1.510	1.510	1.510	1.243	1.412
Age	1.828	1.837	1.837	1.375	1.639
Color	1.057	1.119	1.131	1.111	1.920
Birth Weight EPD	1.264	1.264	1.265	1.135	1.373
Weaning weight EPD	1.835	1.838	1.847	1.378	1.598
Final weight EPD	1.890	1.902	1.926	1.407	1.634
Scrotal circ. EPD	1.401	1.413	1.425	1.212	1.408
Maternal EPD	1.037	1.090	1.110	1.098	1.409
Stud FE		1.148	1.177	1.297	3.642
PP FE			1.142	1.077	1.819
Year FE				1.053	1.397
Sire FE					1.045

Table B2. Variance Inflation Factors from the estimated models

APPENDIX C: APPENDIX TO CHAPTER 4

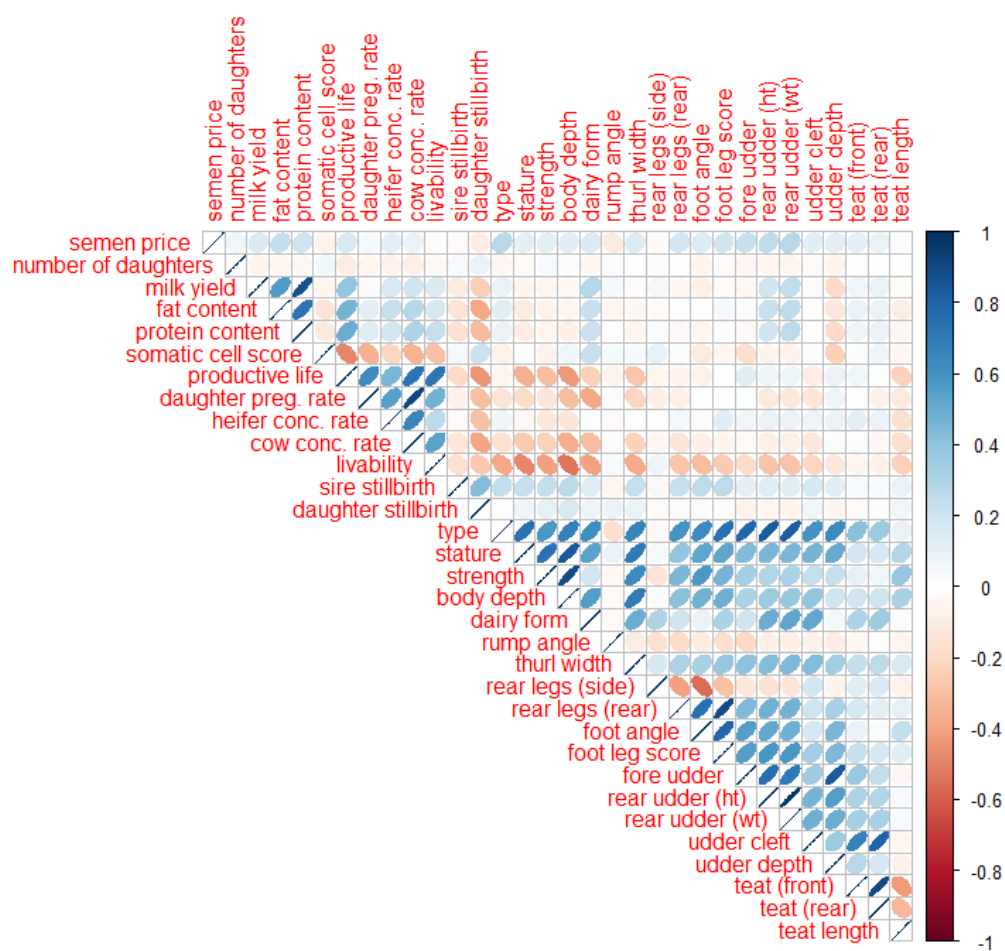


Figure C1. Correlation plot for the full set of genetic traits

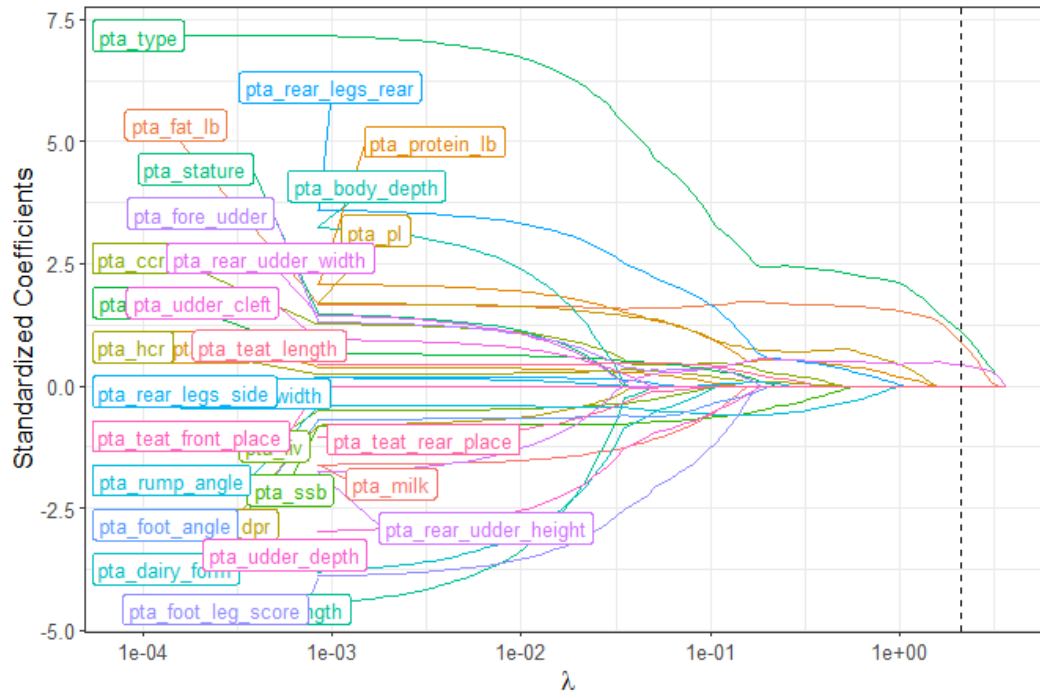


Figure C2. Coefficient plot for semen price

REFERENCES

- Adams, H. A., Sonstegard, T. S., VanRaden, P. M., Null, D. J., Van Tassell, C. P., Larkin, D. M., & Lewin, H. A. (2016). Identification of a nonsense mutation in APAF1 that is likely causal for a decrease in reproductive efficiency in holstein dairy cattle. *Journal of Dairy Science*, 99(8), 6693–6701. <https://doi.org/https://doi.org/10.3168/jds.2015-10517>
- Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, 119(1), 249–275.
- Bjelland, D., Weigel, K., Vukasinovic, N., & Nkrumah, J. (2013). Evaluation of inbreeding depression in Holstein cattle using whole-genome SNP markers and alternative measures of genomic inbreeding [Publisher: Elsevier]. *Journal of Dairy Science*, 96(7), 4697–4706. <https://doi.org/10.3168/jds.2012-6435>
doi: 10.3168/jds.2012-6435.
- Bourdon, R. (2000). *Understanding animal breeding*. Prentice Hall.
- Boyer, C. N., Burdine, K., & Laurent, K. (2020). Factors influencing bred beef heifer sale prices in a sequential auction. *Applied Animal Science*, 36(5), 754–759. <https://doi.org/https://doi.org/10.15232/aas.2020-02012>
- Boyer, C. N., Campbell, K., Griffith, A. P., DeLong, K. L., Rhinehart, J., & Kirkpatrick, D. (2019). Price determinants of performance-tested bulls over time. *Journal of Agricultural and Applied Economics*, 51(2), 304–314. <https://doi.org/10.1017/aae.2019.3>
- Callaway, B., & Sant’Anna, P. H. (2021). Difference-in-differences with multiple time periods [Themed Issue: Treatment Effect 1]. *Journal of Econometrics*, 225(2), 200–230. <https://doi.org/https://doi.org/10.1016/j.jeconom.2020.12.001>
- Cameron, A. C., & Trivedi, P. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press.
- Carvalho, G. R., Faria, W. R., Nardy, V. P. D. R., & Betarelli Junior, A. A. (2022). Hedonic pricing analysis for semen of dairy bulls in brazil. *PLOS ONE*, 17(4), 1–17. <https://doi.org/10.1371/journal.pone.0267109>
- Cole, J. (2019). The truth about inbreeding in dairy cattle. *Hoard’s Dairyman*. <https://hoards.com/print-article-26575-permanent.html>
- Cunningham, S. (2021). *Causal Inference: The Mixtape*. Yale University Press. <https://doi.org/doi:10.12987/9780300255881>

- Dhuyvetter, K. C., Jones, R., Turner, T., & Marsh, T. (2005). *Economic values associated with expected progeny differences (epd) for angus bulls at auction* (tech. rep.). Kansas Agricultural Experiment Station Research Reports.
- Dhuyvetter, K. C., & Schroeder, T. C. (2000). Price-weight relationships for feeder cattle. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, 48(3), 299–310. <https://doi.org/https://doi.org/10.1111/j.1744-7976.2000.tb00281.x>
- Dhuyvetter, K. C., Schroeder, T. C., Simms, D. D., Bolze, R. P., & Geske, J. (1996). Determinants of purebred beef bull price differentials. *Journal of Agricultural and Resource Economics*, 21(2), 396–410.
- Espasandín, A. C., & Ducamp, F. (2004, April). *El uso de cruzamientos vs. la utilización de razas puras para la producción de carne bovina* [Accessed: 2023-09-24]. https://www.produccion-animal.com.ar/genetica_seleccion_cruzamientos/bovinos_de_carne/64-cruzamientos_vs_razas_puras.pdf
- Etcheverry, M. (2009, July). Exportación de genética bovina: Diagnóstico de situación. Fortalezas y debilidades para competir en el mercado internacional (www.produccion-animal.com.ar, Ed.). https://www.produccion-animal.com.ar/informacion_tecnica/comercializacion/55-exportacion_genetica_bovina.pdf
- Farley, J. U. (1964). "brand loyalty" and the economics of information. *The Journal of Business*, 37(4), 370–381. Retrieved October 21, 2023, from <http://www.jstor.org/stable/2351231>
- Foundation, H. (2018). Interpreting and Utilizing Haplotype Information [Accessed: 2023-14-06].
- Foundation, H. (2021). Understanding Genetics and the Sire Summaries [Accessed: 2023-14-06].
- Freyaldenhoven, S., Hansen, C., & Shapiro, J. M. (2019). Pre-event trends in the panel event-study design. *American Economic Review*, 109(9), 3307–38. <https://doi.org/10.1257/aer.20180609>
- Goldberger, A. S. (1991). *A course in econometrics*. Harvard University Press.
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing [Themed Issue: Treatment Effect 1]. *Journal of Econometrics*, 225(2), 254–277. <https://doi.org/https://doi.org/10.1016/j.jeconom.2021.03.014>
- Greenwood, P. L. (2021). Review: An overview of beef production from pasture and feedlot globally, as demand for beef and the need for sustainable practices increase [Sustainable livestock systems for high-producing animals]. *Animal*, 15, 100295. <https://doi.org/https://doi.org/10.1016/j.animal.2021.100295>
- Guinan, F., Wiggans, G., Norman, H., Dürr, J., Cole, J., Van Tassell, C., Misztal, I., & Lourenco, D. (2023). Changes in genetic trends in us dairy cattle since the implementation of genomic selection. *Journal of Dairy Science*, 106(2), 1110–1129. <https://doi.org/https://doi.org/10.3168/jds.2022-22205>
- Hastie, T., Tibshirani, R., & Wainwright, M. (2015). *Statistical Learning with Sparsity: The Lasso and Generalizations*. Chapman & Hall/CRC.
- Hayami, Y., & Ruttan, V. (1985). *Agricultural development: An international perspective*. Johns Hopkins University Press.
- Hazel, L. N. (1943). The Genetic Basis for Constructing Selection Indexes. *Genetics*, 28(6), 476–490. <https://doi.org/10.1093/genetics/28.6.476>
- Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The Review of Economic Studies*, 64(4), 605–654.

- Hutchins, J., & Hueth, B. (2023). 100 years of data sovereignty: Cooperative data governance and innovation in us dairy. *Applied Economic Perspectives and Policy*, 1–26. <https://doi.org/https://doi.org/10.1002/aep.13339>
- Imbens, G. W., & Rubin, D. B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139025751>
- Jones, R. D., Turner, T., Dhuyvetter, K., & Marsh, T. (2008). Estimating the economic value of specific characteristics associated with angus bulls sold at auction. *Journal of Agricultural and Applied Economics*, 40(1), 455–28.
- Katz, M. L., & Shapiro, C. (1985). Network externalities, competition, and compatibility. *The American Economic Review*, 75(3), 424–440. Retrieved August 16, 2023, from <http://www.jstor.org/stable/1814809>
- Kerr, W. A. (1984). Selective breeding, heritable characteristics and genetic-based technological change in the canadian beef cattle industry. *Western Journal of Agricultural Economics*, 9(1), 14–28.
- Ladd, G. W., & Gibson, C. (1978). Microeconomics of technical change: What's a better animal worth? *American Journal of Agricultural Economics*, 60(2), 236–240.
- Lewbel, A. (2019). The identification zoo: Meanings of identification in econometrics. *Journal of Economic Literature*, 57(4), 835–903. <https://doi.org/10.1257/jel.20181361>
- Lush, J. L. (1937). *Animal breeding plans*. Collegiate Press, Incorporated. <https://books.google.com/books?id=CZbKoAEACAAJ>
- Madalena, F., Madureira, A. P., Maldini Penna, V., & Maldonado Turra, E. (1996). Fatores que afetam o preço do sêmen bovino. 1. raças nelore e gir leiteiro. *Revista Brasileira de Zootecnia*, 25, 428–436.
- Marcantonio, S. A. (2017, October). Argentine bovine genetics market (F. de Genética Bovina, Ed.) [Accessed: 2023-11-25]. <https://forodegeneticabovina.com.ar/wp-content/uploads/2017/10/Argentine-bovine-genetics-market-English-1.pdf>
- Martinez, C. C., Boyer, C. N., & Burdine, K. H. (2021). Price determinants for feeder cattle in tennessee. *Journal of Agricultural and Applied Economics*, 53(4), 552–562. <https://doi.org/10.1017/aae.2021.24>
- Melton, B. E., Colette, W. A., & Willham, R. L. (1994). Imputing input characteristic values from optimal commercial breed or variety choice decisions. *American Journal of Agricultural Economics*, 76(3), 478–491. <https://doi.org/https://doi.org/10.2307/1243659>
- Melton, B. E., Heady, E. O., & Willham, R. L. (1979). Estimation of economic values for selection indices. *Animal Production*, 28(3), 279–286. <https://doi.org/10.1017/S0003356100023199>
- Mitchell, J. L., Peel, D. S., & Brorsen, B. W. (2018). Price determinants of bred cows. *Journal of Agricultural and Applied Economics*, 50(1), 64–80. <https://doi.org/10.1017/aae.2017.20>
- Montiel-Olea, J. L., & Plagborg-Møller, M. (2019). Simultaneous confidence bands: Theory, implementation, and an application to SVARs. *Journal of Applied Econometrics*, 34(1), 1–17. <https://doi.org/https://doi.org/10.1002/jae.2656>
- Nilforooshan, M. A., & Saavedra-Jiménez, L. A. (2020). Ggroups: An R package for pedigree and genetic groups data. *Hereditas*, 157(1), 17. <https://doi.org/10.1186/s41065-020-00124-2>
- Prices and demands for input characteristics. (1976). *American Journal of Agricultural Economics*, 58(1), 21–30.
- Reca, L. G., Lema, D., & Flood, C. (2010). *El crecimiento de la agricultura argentina. medio siglo de logros y desafíos*. University of Buenos Aires Press.

- Richards, T. J., & Jeffrey, S. R. (1996). Establishing indices of genetic merit using hedonic pricing: An application to dairy bulls in alberta. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroéconomie*, 44(3), 251–264. <https://doi.org/https://doi.org/10.1111/j.1744-7976.1996.tb00149.x>
- Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1), 34–55.
- Rosen, S. (1981). The economics of superstars. *American Economic Review*, 71(5), 845–858.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55.
- Roth, J. (2022). Pretest with caution: Event-study estimates after testing for parallel trends. *American Economic Review: Insights*, 4(3), 305–22. <https://doi.org/10.1257/aeri.20210236>
- Roth, J., Sant'Anna, P. H., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *Journal of Econometrics*, 235(2), 2218–2244. <https://doi.org/10.1016/j.jeconom.2023.03.008>
- Schroeder, T. C., Espinosa, J. A., & Goodwin, B. K. (1992). The value of genetic traits in purebred dairy bull services. *Review of Agricultural Economics*, 14(2), 215–226.
- Schultz, T. (1983). *Transforming traditional agriculture*. University of Chicago Press.
- Smith, J. A., & Todd, P. E. (2005). Does matching overcome Lalonde's critique of nonexperimental estimators? [Experimental and non-experimental evaluation of economic policy and models]. *Journal of Econometrics*, 125(1), 305–353. <https://doi.org/https://doi.org/10.1016/j.jeconom.2004.04.011>
- Smith, K. V., DeLong, K., Griffith, A. P., Boyer, C. N., Martinez, C., & Jensen, K. L. (2023). Cow-Calf Producer Preferences for Bull Genomic Enhanced Expected Progeny Differences. *Journal of Agricultural and Resource Economics*, 48(3), 322848.
- Smith, L. A., Cassell, B. G., & Pearson, R. E. (1998). The effects of inbreeding on the lifetime performance of dairy cattle. *Journal of dairy science*, 81(10), 2729–2737. [https://doi.org/https://doi.org/10.3168/jds.S0022-0302\(98\)75830-8](https://doi.org/https://doi.org/10.3168/jds.S0022-0302(98)75830-8)
- Steyn, Y., Masuda, Y., Tsuruta, S., Lourenco, D., Misztal, I., & Lawlor, T. (2022). Identifying influential sires and distinct clusters of selection candidates based on genomic relationships to reduce inbreeding in the us holstein. *Journal of Dairy Science*, 105(12), 9810–9821. <https://doi.org/https://doi.org/10.3168/jds.2022-22143>
- Steyn, Y., Lawlor, T. J., Lourenco, D., & Misztal, I. (2023). The importance of historically popular sires on the accuracy of genomic predictions of young animals in the US Holstein population. *Journal of Dairy Science Communications*, 4(4), 260–264.
- Stigler, G. J. (1961). The economics of information. *Journal of Political Economy*, 69(3), 213–225. Retrieved October 21, 2023, from <http://www.jstor.org/stable/1829263>
- Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2), 175–199. <https://doi.org/10.1016/j.jeconom.2020.09.006>
- Sunding, D., & Zilberman, D. (2001). The agricultural innovation process: Research and technology adoption in a changing agricultural sector. In *Agricultural production* (pp. 207–261, Vol. 1). Elsevier. [https://doi.org/https://doi.org/10.1016/S1574-0072\(01\)10007-1](https://doi.org/https://doi.org/10.1016/S1574-0072(01)10007-1)
- Sy, H. A., Faminow, M. D., Johnson, G. V., & Crow, G. (1997). Estimating the values of cattle characteristics using an ordered probit model. *American Journal of Agricultural Economics*, 79(2), 463–476.

- Tang, M., Thompson, N., Boyer, C. N., Olynk Widmar, N. J., Stewart, T. S., Lofgren, D. L., & Minton, N. (2020). Temporal changes in angus bull attribute valuations in the midwest. *Journal of Agricultural and Applied Economics*, 45(3), 518–532. <https://doi.org/10.22004/ag.econ.302500>
- Thompson, T., Boyer, C. N., Martinez, C. C., Rowan, T. N., & Rhinehart, J. (2022). Valuation of genomic-enhanced expected progeny differences in bull purchasing. *Journal of Agricultural and Applied Economics*, 54(4), 713–722. <https://doi.org/10.1017/aae.2022.36>
- USDA. (2023). *Production, supply and distribution (psd) data* [Accessed: 2023-02-10]. <https://apps.fas.usda.gov/psdonline/app/index.html#/app/home>
- Van Raden, P., Neupane, M., Toghiani, S., Parker Gaddis, K., & Tempelman, R. (2021). *Net merit as a measure of lifetime profit: 2021 revision* (tech. rep.). USDA-AIP Research Report.
- Walburger, A. M. (2002). Estimating the implicit prices of beef cattle attributes: A case from alberta. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, 50(2), 135–149. <https://doi.org/https://doi.org/10.1111/j.1744-7976.2002.tb00424.x>
- Whitt, C. E., Tauer, L. W., & Huson, H. (2019). Bull efficiency using dairy genetic traits. *PLOS ONE*, 14(11), 1–14. <https://doi.org/10.1371/journal.pone.0223436>
- Wiggans, G. R., & Carrillo, J. A. (2022). Genomic selection in United States dairy cattle. *Frontiers in Genetics*, 13. <https://doi.org/10.3389/fgene.2022.994466>
- Wiggans, G., VanRaden, P., & Zuurbier, J. (1995). Calculation and use of inbreeding coefficients for genetic evaluation of united states dairy cattle. *Journal of Dairy Science*, 78(7), 1584–1590. [https://doi.org/https://doi.org/10.3168/jds.S0022-0302\(95\)76782-0](https://doi.org/https://doi.org/10.3168/jds.S0022-0302(95)76782-0)
- Yue, X.-P., Dechow, C., & Liu, W.-S. (2015). A limited number of y chromosome lineages is present in north american holsteins. *Journal of Dairy Science*, 98(4), 2738–2745. <https://doi.org/https://doi.org/10.3168/jds.2014-8601>
- Zimmerman, L. C., Schroeder, T. C., Dhuyvetter, K. C., Olson, K. C., Stokka, G. L., Seeger, J. T., & Grotelueschen, D. M. (2012). The effect of value-added management on calf prices at superior livestock auction video markets. *Journal of Agricultural and Resource Economics*, 37(1), 128–143.