Sharing of Genomic Information in the Beef Industry's Seedstock Sector

by

Adilbek Sultanov

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Agricultural and Resource Economics

Department of Resource Economics and Environmental Sociology
University of Alberta

© Adilbek Sultanov, 2021

Abstract

Recent advances in genomics have revolutionized selective breeding in many agriculturally important animal species such as dairy cattle, poultry, and pork. However, the adoption of genomic selection in the beef industry has been slower for reasons such as the existence of multiple breeds, the poor extent of phenotyping, lack of use of artificial insemination, and lower profit margins. Sharing genomic data between Breed Associations has been considered a solution and some information sharing platforms have emerged. Information sharing can also support the declining beef demand by improving meat quality. However, information sharing also makes BAs compete on product consistency, correlates their strategies, and increases signal variability.

Our aim in this thesis is to analyze the viability of IS to individual BAs in the seedstock sector. We use the study of oligopolistic competition under uncertainty to develop a game-theoretic model. Two models are proposed. In the basic specification, we consider the cases of full information sharing versus no information sharing. In the next specification, we allow for coalitional information sharing where breed associations can share in a coalition.

In the basic model, we found that in general, information sharing is less likely to occur when information is valuable to the breed association - either for improving the quality of the trait or the production decision. Second, we found that BAs with a large market size will not share information in a market with close substitutes. Finally, we found that information sharing increases profitability in expectation but makes the profit less predictable. One implication of this result is that Full Information Sharing may happen eventually as BAs have a better prior.

In the extension of the model, we found that coalitional information sharing will prevail for most values of market differentiation, except when BAs sell very strong or weak substitutes. This implies that once coalitions are established, it is hard to achieve full information sharing. Another implication is that the presence of IS coalitions introduces beef product differentiation in terms of quality.

To my parents

Acknowledgements

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Recent advances in genomics have revolutionized selective breeding in many agriculturally important animal species. In 2001, Meuwissen, Hayes, and Goddard published a landmark article that outlined the approach, subsequently called genomic selection, as well as the statistical methods of estimating the genetic merit of individuals based solely on the phenotypic and genotypic data of their ancestors (Van Eenennaam et al. 2014). Within the last two decades, genomic selection transformed the dairy industry (Meuwissen, Hayes, and Goddard 2013), and led to advances in other animal and plant industries (Van Eenennaam et al. 2014).

However, the adoption of genomic selection in the beef industry has been slower for several reasons (Hayes, Lewin, and Goddard 2013). Compared to dairy cattle, developing accurate genomic evaluations in beef populations is more difficult due to factors such as the existence of multiple breeds, poor extent of phenotyping (assessment of expressed traits), lack of artificial insemination (AI), and lower profit margins throughout the value chain (Berry, Garcia, and Garrick 2016).

Conducting multi-breed genomic evaluations to exploit information from multiple breeds has been considered one of the solutions for improving the accuracy of genomic predictions (Berry, Garcia, and Garrick 2016) as this method compensates for the small number of reference animals within a breed (Meuwissen, Hayes, and Goddard 2016). However, multi-breed evaluations are more complex than single breed

evaluations and, thus, require more data to be able to make accurate predictions (Berry, Garcia, and Garrick 2016).

Despite these difficulties, information sharing (IS) platforms to facilitate multi-breed evaluations are emerging. The most prominent one to date is the American Simmental Association's (ASA) International Genetic Solutions (IGS), which pools information from 17 industry partners and has around 20 million animal records (IGS 2021). In Australia, BREEDPLAN - a widely used genetic evaluation service for beef cattle internationally, recently announced plans to develop multi-breed evaluations (Meat and Livestock Australia 2020). In Canada, the Canadian Beef Improvement Network (CBIN), a third-party IS institution, is under development and has recently been funded by the four major breed associations (BAs) (Briere 2021). BAs are non-profit organizations representing their respective breeds in a country or region.

The benefits of IS to the beef industry have often been emphasized by its advocates. For example, IGS positions themselves as an "unprecedented collaboration among progressive breed associations to enhance beef industry profitability" (IGS 2021). Dave Sibbald, Interim Chair of the CBIN, highlighted the role of IS in strengthening the position of the Canadian beef industry against competing industries when he said:

> "CBIN is not about maintaining market share, it's about growing it. The global market is highly competitive, so we need new initiatives to confirm our leadership and support the great beef message we already have."

Indeed, IS can be beneficial to the beef industry. Beef consumption has been declining for about three decades in Canada for financial, health and food safety and environmental reasons (StatCan 2020). The beef industry's lack of vertical integration and poor information flow have also played a role in the deterioration of product quality and the subsequent decline of the demand for beef (Canovas et al. 2017). IS can help regain some of the demand for beef by increasing the accuracy of genomic selection and improving meat quality. Additionally, IS can increase breeders'

profitability by reducing production cost amid falling beef prices (Kemp and Atkins 2019).

Pooling information also helps BAs to meet the changing needs of commercial buyers. There has been a perceived shift among breeders from a breed-based approach to a performance-based approach to breeding. As more data become available and genomic tests get cheaper, increasingly more commercial producers choose animals based on their individual performance data regardless of the breed (Lynch-Staunton 2020). Responding to the demand, more breeders rely more on individual genomic performance for their selection decisions, and many adopt an open herd book approach (Lynch-Staunton 2020). IS brings the animals' performance metrics to a common base enabling commercial buyers to directly compare the animals on their performance. Thus, the BAs who engage in IS and open their animals to be directly comparable put themselves in a better position to satisfy the new demand from commercial buyers and may even stand a better chance of survival in the long-term (Kemp and Atkins 2019). There is evidence that bulls that perform well in multiple traits are priced higher, suggesting that the commercial market is ready to pay more for animals with well-rounded traits, rather than animals with a narrow trait focus (Industry Voice by Neogen 2021).

However, there are also costs associated with sharing information, which may be reflected in some BAs' reluctance to share. One of the main worries is the potential fall in the relative ranking of one's animals when compared against all others in the industry (Lynch-Staunton 2020). Currently, BAs report EPDs on different EPD bases that are set arbitrarily by each BA. This makes EPDs harder to compare across breeds, and reduces competition of BAs on performance metrics. Thus, currently BAs mainly compete by differentiation, i.e. by promoting signature traits to their customers - commercial producers. Sharing information essentially brings the BAs' performance metrics to one base. This makes animals directly comparable across breeds and, thus, increases the competition between BAs by having them compete on

performance. Since information revelation is largely irreversible, BAs will find it hard to come back to performance evaluations with arbitrary bases that they had before. Knowing this, some BAs who feel their animals will rank lower after information revelation may be reluctant to engage in IS and choose to promote signature traits to their loyal customer base. However, as the commercial market becomes more savvy with performance metrics, they may choose to move towards BAs that provide standardized performance metrics.

In a context where demand is uncertain, another cost of IS to the individual BAs is the correlation of their production strategies under IS. It has been shown in the IS literature that increasing the correlation of production strategies erodes profitability in markets with substitutes (Vives 1984). IS tends to change production decisions making production levels more homogeneous across BAs. Moreover, in the longer run, as BAs pursue the animal that maximizes profit and shift away from differentiation - driven breeding to performance - driven breeding, their breeds may become increasingly genetically similar. If we assume a certain number of winner genes and traits that maximize profit, it is possible that eventually breeds will aim to incorporate those genes in their animals. A parallel can be drawn with the dairy industry, which is almost exclusively dominated by the Holstein breed. In Canada, 94% of the million dairy cows are Holstein, while the next biggest one, Jersey, takes up about 4% of the dairy cattle population (Holstein Canada 2021). The Holstein breed became very popular due to high volume of milk production, which led to more dairy farmers switching to breeding Holstein (Kral 2017). The industry's convergence to breeding Holsteins has also been facilitated by the broad use of AI, which allowed for an even more narrow targeting of specific genes (Kral 2017). It has been calculated that in the US 99% of the male Holsteins were descendants of just two bulls born in the 1960s (O'Hagan 2019), which was enabled by the use of AI. In the beef industry, as profit-maximizing traits or combinations of traits are identified more precisely, AI is used more widely and the costs of genotyping decrease, we may see a similar con-

vergence of the beef industry to a particular genetic makeup. The breeders of these more homogenized animals will earn less, as they will compete in a market with close substitutes.

Additionally, as production strategies become more correlated, breeds may get more vulnerable to the variability of the pooled signal. First, the loss of genetic diversity that may result from IS will increase the chances of genetic mutations. The severity of negative mutations will be greater if the genes are ubiquitous in the cattle population. For instance, as of now high reliance on Angus genetics may have exposed the cattle herd to a higher chance of congestive heart failure (Thoren 2020). The focus on profit maximization also means that other potentially important traits may be selected out (Kral 2017). For example, as the climate is changing, traits like heat and drought tolerance will become more important, but may have disappeared by the time they become critical. The same can be said about epidemics and other extreme or unpredictable events. There is evidence showing that the US cattle populations are losing environmental adaptations due to the use of semen of bulls from different environments, which decreases the frequency of beneficial environmental adaptations (BEEF 2021). For example, cows in Colorado could be losing adaptations that ease cardiovascular stress at high altitudes as a result of this (BEEF 2021).

As far as we know, no formal analysis has been done to understand the economic viability of IS to individual BAs. BAs differ in the number of registered animals, market size, signature traits, and levels of adoption of genomic selection. These factors may affect the profitability of sharing information. The industry is dominated by the four largest breeds and their respective BAs: Angus, Simmental, Hereford, and Charolais. These big BAs may have different incentives to share information compared to smaller BAs (Lynch-Staunton 2020). BAs with established brands and customer base stand to lose more if sharing information results in a fall in the ranking of their animals. Furthermore, BAs also differ in their levels of adoption of genomics: while some routinely use high-density tests, others are just starting to introduce

genomic tests (Lynch-Staunton 2020). BAs that have more data and understanding of genomics may feel reluctant to share with others, especially if BAs are in strong competition.

This thesis aims to add to our understanding of individual incentives of BAs by using models of industrial organization and Bayesian games. The overall goal of this thesis is to formally analyze the viability of IS to individual BAs in the seedstock sector. To do so we ask the following questions:

1. In a market without IS, would a BA be better off if all BAs pooled their information? Under which circumstances?

2. Are BAs better off sharing information in groups, or coalitions, of BAs? Will BAs prefer to exclude others from the group?

3. Can the BA be better off sharing part, rather than all, of its information?

We answer the first question in Chapter 3 where we establish the economic models of the BA's profit under private information and full information disclosure. We introduce and define the model parameters that include the market size, trait development, and a BA's selection decision. Our aim in this chapter is to compare the two extreme cases: no IS and complete information disclosure to see under what circumstances IS becomes viable. For our basic model we draw inspiration from a seminal work by Vives (1984) while also adding a component that deals with breed improvement through selection. In general, we find that BAs will tend to withhold information when it is valuable, i.e. when the BA has much to learn from new information and its signature trait has substantial room for improvement. Furthermore, we find that as BAs with a large market size have more at stake, they are more reluctant to share information when close substitutes are sold. However, we also find that BAs with a large market size will be willing to share information once they develop sufficient knowledge. Importantly, we find that, although on average IS increases profitability,

it also makes the BAs more susceptible to the fluctuations of the pooled signal. This suggests that IS is inherently risky, which may explain the reluctance of some BAs to share information.

We answer the second question in Chapter 4 where we model a group of two BAs who share information with each other with one BA staying on its own. We show under which conditions the individual BA is willing to join the IS coalition if the opportunity is provided. Then we look at whether the coalition BAs are better off sharing within a coalition or leaving, as well as inviting the third BA to join. We find that full IS prevails when BAs sell weak substitutes. In contrast, when breeds are strong substitutes, the coalition BAs prefer to decouple their strategies and no IS happens. Finally, we find that when breeds are moderate substitutes, IS in coalitions prevails. We thus conclude that if BAs can make their own decision about sharing information, as opposed to when all BAs commit to sharing information, a new incentive structure emerges that prevents full IS to occur. This is similar to Marcoul's (2020) result in that we find that sharing in coalitions is likely to occur when BAs sell moderately strong substitutes.

Chapter 5 deals with the third question regarding the equilibria that emerge when BAs decide to share part, as opposed to all, of their data pool. Our preliminary findings suggest that established BAs in a market with strong substitutes will only engage in IS if a significant proportion of the data pool is shared, and will otherwise decline IS altogether.

The rest of the thesis is organized as follows. Chapter 2 provides a background of the beef industry, an overview of genomics and its use in the beef industry, and a brief literature review. Chapter 3 establishes the basic economic model and explores full sharing by commitment. Chapter 4 explores asymmetric equilibria with coalitions. Chapter 5 goes into the BA's decision on partial IS. Finally, Chapter 6 concludes our findings. The appendix to the thesis provides detailed proofs of lemmas and propositions made throughout the thesis.

Chapter 2

Background and Literature Review

2.1 Background

2.1.1 The beef cattle industry in Canada

The beef industry is important to Canada as it contributes around $18 billion to Canada's GDP annually, and consists of nearly 60 thousand farms and ranches and 11 million total cattle and calves. Around 3 million animals are finished annually, with around $8 million of cattle and calves received (Witte 2018). Alberta is the largest producer across all provinces in Canada: it accounts for 69% of Canada's fed cattle production with 1.6 million head annual output, and comprises 149 feedlots of 1000 head or more. The province also has the most cattle, accounting for 41.6% of the national herd. More than 40% of Canada's beef breeding stock (beef cows and heifers for beef herd replacement) and around 60% of the national feeder cattle come from Alberta (Alberta Cattle Feeders Association 2019). Of the $62.2 billion total farm cash receipts in 2018, $25 billion were receipts from livestock, with cattle contributing $8.1 billion. In 2018, total farm cash receipts of Alberta were $13.6 billion, representing 21.8% of Canada's total. Total provincial receipts from livestock and livestock products were $6.5 billion, of which cattle contributed around $4.8 billion (Statistics Canada 2019).

The Canadian beef industry is a complex system that involves numerous different participants (Goddard et al. 2016). The supply chain consists of the production

segments and the marketing segments. The production segments are seed-stock, cow/calf, backgrounding, and feedlot or finishing operations, while the marketing segments are beef processing (packing) and retail/sales. The beef value chain begins with the breeding or seedstock operations that produce genetic resources such as animals for breeding, semen and embryos which are used in the operations down the value chain (Field 2017). The breeding segment is essential to the success of the Canadian beef sector as the collective genetic decisions made by the breeding sector need to match the goals of commercial cow-calf producers as well as market specifications down the value chain (Witte 2018).

Cow-calf operations select cows to produce calves based on desirable traits using genetic resources from the breeding segment. Cow-calf operations raise beef cattle by keeping a permanent herd of cows to produce calves that are later weaned from mother cows and sold. Calves weigh between 220 and 250 kg at weaning, which usually occurs at six to eight months of age. After this, the calves are placed on a forage-based diet on open pastures for grazing and nursing. This happens on most farms until weaning in the fall (CCA and BIC 2010). Cow-calf can be purebred or commercial operations. A purebred operation typically raises one breed of cattle where all the cattle is sold through purebred sales, while a commercial operation raises crossbred cattle or purebred cattle that has not been registered (Barkley 2012).

After weaning, some calves are sent to a backgrounding operation. Backgrounding is an intermediate stage in cattle production which relies more heavily on forage and grains to increase a calf's weight until between 350 and 450 kg before placing it in a feedlot (Womach 2005; Fan 2018). Depending on the breed, some calves are sent directly to feedlots (Athwal 2002; Twine 2014). The cattle are then placed in a feedlot where they are fed to weights of 550-600 kg at about 18 to 24 months of age (Twine 2014), at which point they are ready for slaughter. The calves are placed on high-energy diets such as barley, corn and to some extent, on wheat and oats (Maclachlan and Stringham 2016).

The next stage is the processing segment, which currently consists of 6 processing plants in Alberta (Agriculture and Agri-Food Canada 2019). They are operated by the following meat companies: Lacombe Meat Research Centre, JBS Food Canada Inc., Cargill Limited, Bouvry Export Calgary Ltd., Canadian Premium Meats Inc., and Harmony Beef Company Ltd. The carcasses are graded by quality, and yield grades indicate the percentage of red meat present in the carcass (Witte 2018).

The beef industry has been extensively using technology, and one innovation that has become widely adopted is genomic selection (Van Eenennaam and Drake 2012; Rutherford 2016), which is now widely used for breeding decisions by both the seed-stock producers and cow-calf ranchers (Rutherford 2016). Genomic selection is the use of statistical methods, such as prediction equations, to estimate the genetic merit of a genotyped animal based on genotypes and phenotypes of its ancestors (Van Ee-nennaam et al. 2014). In the US, Canada, and other countries Breed Associations (BAs) use genomic data to increase the accuracy of Expected Progeny Differences (EPDs) which are expected differences in the performance of the progeny of a partic-ular animal for a certain characteristic or trait, and the performance of a calf of an average animal (Schmid 2013). EPDs and individual performance data are routinely reported in sale catalogs of bulls, which are used by cow-calf producers in purchasing decisions (Rutherford 2016; Vestal et al. 2013).

2.1.2 Genetics 101

A quick recap of basic genetics may be useful to the reader. DNA is the hereditary material in most living organisms. It is stored as a code that determines an animal's growth, appearance, behavior, and other traits. DNA is found in all living organisms and gets passed from one generation to the next, improving their ability to survive and thrive. The DNA molecule consists of four chemical "bases" – adenine (A), thymine (T), guanine (G), and cytosine (C). These bases are paired with hydrogen bonds, forming an ordered string which is an instruction for creating or regulating a

hormone, enzyme or other product. Each cell has a mechanism that reads the DNA code three letters at a time. These three letter codes tell the mechanism to start reading at a certain point. From there the combinations of base pairs form specific amino acids, after which the mechanism is instructed to stop reading. The order of these bases forms a genetic sequence, and the chain of amino acids the gene codes for is referred to as protein. Many genes have specific functions in the body; some code for specific enzymes, proteins, and hormones, while others determine whether certain genes are turned on or off (Beef Cattle Research Council 2017).

Mutations are slights changes in the DNA sequence. The effects of mutations can be diverse: they can change the length of the sequence, insert amino acids, stop proteins from being produced, or make other adaptations. Some mutations result in the changes in physical traits; for example, the polled gene that removes horns in cattle is a result of changes in chromosome one. Mutations to the leptin gene in chromosome four affect the depth of back fat, lean yield and days to market. Environmental factors can affect the expression of genes, even though the animal's genetics are determined at conception. For example, if an animal has genetic potential for weight gain, it will fail to gain weight under any type of deprivation (Beef Cattle Research Council 2017).

Single Nucleotide Polymorphisms (SNPs) are the most common type of genetic variation (Stoneking 2001) and are essentially spots in the DNA sequence where mutations happen. They determine the variety of physical attributes we see in animals such as liveweight, height, coat color, horns, marbling, growth rate and many other traits (Beef Cattle Research Council 2017). Thus, identifying information about SNPs is important in breeding because it allows breeders to target economically important traits in the breed. Most genomic tests today rely on SNP technology. There are about 35 million SNPs in a single animal, and identifying them all would be expensive. Therefore, genomic labs develop smaller SNP "panels" which are different in SNP density. The more dense panels capture a larger number of SNPs and are typically

more expensive. The panels can be broadly divided into five categories:

- Small panels: analyze up to around 2,000 SNPs and are used mainly to test for parentage, genetic abnormality, coat color and other simple traits. Simple traits are traits that only involve a few SNPs.

- Low-density panels: contain 5,000-30,000 SNPs, are used for genomic prediction and selection when there are higher density genotypes available for reference.

- Medium density panels: have between 50,000 and 150,000 SNPs and are popular in beef genomic selection programs. These panels are usually applied to the most influential animals of the breed.

- High Density panels: analyze 500,000 to 1 million SNPs and are not common in the beef industry as they are expensive.

- Whole Genome Sequencing (WGS), which is typically used for research purposes.

(Beef Cattle Research Council 2017).

The information collected from SNP panels on a specific breed constitutes a BA's genomic data. When a BA has collected a large genomic database from conducting the SNP panel testing, it has a deeper knowledge of the genetic makeup of its breed and can therefore conduct selection more precisely and more quickly. Specifically, superior genomic data enables a BA to increase its evaluation and selection accuracy. Falconer (1960) provides a classic equation of the rate of genetic change toward a given breeding objective, which is expressed as:

$$\Delta G = irAL$$

where i is the selection intensity (the share of animals who are selected to pass on genes), r is the accuracy of selection (correlation between the estimated breeding value and the true breeding value), A is genetic variation in the population, and L is

the generation interval (average age of parents when their offspring is born) (Falconer 1960). Thus, obtaining more genomic data improves the accuracy of selection and accelerates the speed of genetic change, enabling the BA to produce an animal with superior traits.

The data from SNP panels can potentially also be shared between BAs. Effectively collecting and processing genomic data is a lengthy and expensive process and BAs are not equal in the amount and quality of data they have. Some BAs have done SNP testing for decades and use superior testing technology than others. More data is also needed to improve the accuracy of selection, especially for more complex traits. This creates the possibility of sharing genomic data between the BAs. Therefore, the data from SNP panels is what we mean when talk about information sharing (IS) in section 2.2 and the subsequent chapters.

An animal's performance with regards to a specific trait is measured by a metric called Estimated Breeding Value (EBV). EBV is a within-breed evaluation of an animal's genetic merit, which equals twice the expected progeny difference (Van Eenennaam et al. 2014). In the calculation of EBVs, the performance of individual animals within a contemporary group is directly compared to the average of the animals in that contemporary group. A contemporary group consists of animals of the same sex and age within a herd. EBVs are expressed in the units of measurement for each particular trait. They are shown as positive or negative differences between an individual's genetic difference and that of the comparison group. For example, a bull with an EBV of +50 kg for 600 Day Weight, has a genetic merit of 50kg above the breed base of 0 kg. The breed base is a historical benchmark, so the average EBV each year will be different. Therefore, the absolute value of the EBV is not important, but rather the differences in EBVs between animals (Agricultural Business Research Institute 2019). The accuracy of the EBV is a value between 0 and 1 that reflects how close the estimated value is to the true genetic value of the animal. It is based on the amount of performance information available on the animal and particularly

the number of progeny analyzed. A higher accuracy indicates a lower likelihood that the EBV will change as more information is collected on the animal or its progeny. EBVs are just as likely to decrease as they are to increase as more information becomes available (Ryan 2016). There is essentially no difference between EBV and EPD in terms of the information content. While an EBV predicts the genetic merit of an animal itself, an EPD is a predictor of the genetic merit of an animal's progeny. Calculation-wise, EBV equals twice the value of EPD. EPDs are used by the US beef industry, but most of the rest of the world uses EBVs (Notter et al. 2016).

2.1.3 Genomic selection in beef cattle

The adoption of GS in the beef industry has been slow compared with the dairy industry (Bolormaa et al. 2013). The accuracy of predictions in the beef cattle has been quite low so far compared to the dairy cattle possibly due to breed heterogeneity, less advanced structures and breeding programme, preference for natural service, cross-breeding, as well as effective population size (Johnston, Tier, and Graser 2012). Van Eenennaam et al. (2014) reported accuracies in the range of 0.3 to 0.7, which, in their view, is low and due to the lower quality of the reference population in beef cattle. Compared to the dairy cattle, beef cattle have fewer reference animals, and the reference population are not as closely related to the target population in the beef cattle industry as in the dairy industry (Meuwissen, Hayes, and Goddard 2016). Additionally, the beef industry does not use artificial insemination (AI) very heavily. As a result, fewer high-accuracy sires are available to provide a within-breed reference population.

Due to the low use of AI in the beef industry so far, developing phenotyped and genotyped training populations requires a much bigger reference population and is thus more expensive than in other meat industries. Consequently, the development of training populations in the beef cattle industry has been stalled. Some companies, such as GeneSeek and Zoetis, developed their own training populations or paid for

genotyping of semen collections from AI sires put together by individual researchers (McClure et al. 2010). This involvement of private genomic firms introduced a proprietary component into the way animals are ranked based on genetic merit (Garrick 2011), and has made it difficult to obtain the validation data of the resulting genomic predictions (Van Eenennaam et al. 2007).

Another big impediment to the widespread adoption of genomic selection in beef cattle is economic viability. Genomic selection is particularly good for traits that are hard to select for traditionally. In beef cattle some of these traits – such as feed conversion efficiency and beef quality – are also expensive to record. It is costly to set up large training populations, especially for one company alone.

One possible solution to the small number of reference animals within a breed has been the use of multi-breed reference populations. Harris, Johnson, and Spelman (2009) showed that prediction equations developed for one breed did not perform well in another breed. Bolormaa et al. (2013) report that using a common training population increased accuracy slightly, but not as much as if the same number of animals were from the same breed. Using a reference population of pure breeds in the US by Akanno et al. (2014) to predict within a crossbred population in Canada yielded low accuracy as well, since the target breed was not included in the population. Since different breeds are selected for different purposes and different intensity, the allele frequencies and linkage disequilibrium are different. This makes the prediction of EPDs for a breed not included in the reference set challenging (Lourenco et al. 2017). Some researchers report low across-breed accuracy (Kachman et al. 2013), while others report reasonable accuracies when the breeds are only a few generations apart (De Roos, Hayes, and Goddard 2009). Meuwissen, Hayes, and Goddard (2016) suggested that this situation could be improved by using Bayesian methods that allow some SNPs to have a larger effect than others, since the prediction can make better use of SNPs in high LD with the QTL, helping the information transfer across breeds (Bolormaa et al. 2013; Khansefid et al. 2014). It has been shown that some

of the accuracy has to do with the relationship between individuals in the training population and those in the selection candidate population (Habier et al. 2010).

To date, there hasn't been much adoption of GS in the beef industry because it is not as economically justified to improve the accuracy of EPDs on young beef sires used on natural mating pastures where they might sire 100 offspring compared to seedstock animals that will have thousands of genetic descendants (Van Eenennaam, Van Der Werf, and Goddard 2011). Many of the traits associated with the profitability of beef production are expensive or difficult to measure (e.g., feed efficiency or fertility), occur late in life (e.g., stayability - the probability a cow will remain in the herd until 6 years of age), or are experienced by a downstream segment of the cattle industry, such that the relevant phenotypes (e.g., disease susceptibility in the feedlot) are never relayed back to the breeder (Van Eenennaam et al. 2014).

There is no preexisting database of phenotypes for these traits that can form a training population. Some countries use electronic animal identification to form a database of phenotypes from the entire production chain (Wickham and Dürr 2011). There are also several publicly funded initiatives to create datasets of genotypes and phenotypes of complex traits (Van Eenennaam et al. 2014).

Pooling data across countries is an attractive option for enlarging the training database for traits that are expensive or difficult to measure (Dürr and Philipsson 2012). This approach will improve the accuracy of prediction only if the populations share a common genetic base. The development of phenotyped populations with the thousands of individuals that will be needed to obtain high precision is a significant hurdle to the adoption of GS in beef cattle. Ironically, these are the traits that will most likely benefit from GS because there are currently no selection criteria for them (Van Eenennaam et al. 2014).

Meuwissen et al. (2016) predict that two methods will be used in the future to calculate genomic estimated breeding values (GEBVs): within-breed (wbGS) and across breed (abGS) evaluation. GEBVs are breeding values derived from information in an

animal's DNA (*Genomic Breeding Values*). On the other hand, the training population may be comprised of multiple breeds or cross breeds. In this case, increasing the accuracy of EBVs could be achieved with dense SNPs and nonlinear methods. For wbGS, there isn't much potential to improve the accuracy aside from increasing the reference population. In contrast, there are many avenues for improvement in abGS.

If wbGS produces sufficiently accurate EBVs (e.g. ¿ 0.9), then there will be no need for abGS. However, animals will be measured on a larger number of traits, some of which will need to be recorded on a large scale, for example methane emissions. This reduces the opportunities for large within-breed reference populations for wbGS. Thus, Meuwissen et al. (2016) believe that abGS will produce more accurate EBV predictions that will be stable across populations that are spatially and temporally diverse. If only wbGS is used in the future, only the largest breeds and lines will remain since only they will have large enough training populations to generate accurate EBVs (Meuwissen, Hayes, and Goddard 2016). This suggests a need for sharing genomic information across different breeds to obtain a more accurate selection. We discuss the implications of that in Chapter 3.

2.1.4 The economics of genomics in the beef cattle industry

Some research has been done on the economics of genomic selection. Ballenger et al. (2016) considered which sectors of the beef cattle industry – seed stock sector, cow-calf operations, stockers and feedlots, processors, retailers, and final consumers – would be willing to consider paying for genomic testing. They reasoned that each cattle producer or buyer has to decide if the current cost of buying a genomic tool is worth the benefit down the road in the form of more valuable or less costly-to-raise offspring. This calculation, in their view, depends on the accuracy of predicting the traits of interest, as well as how much value there is in enhancing those traits in the offspring (Ballenger et al. 2016).

The vertical integration of the industry has implications for how the benefits of

using genomic tools are likely to be captured or shared in the beef supply chain. Compared to poultry, hogs and dairy, beef production has a longer biological cycle spanning vast geographical areas. Furthermore, the ownership of the animals changes multiple times across the production chain. As a result, there is less vertical integration and control in the beef industry than in other livestock sectors (Ballenger et al. 2016). A study by Van Eenennaam and Drake (2012) concludes that although each sector of the beef industry may benefit from genomic selection, they will profit less than in poultry, hogs, and dairy because of the relative lack of vertical control in the beef sector.

Cow-calf ranch operators may be benefiting from genomic information already. However, there is little evidence about use of genomic tools on the ranch itself. Ballenger et al. (2016) describe a meeting with some Wyoming stock growers where most said they were "not very knowledgeable" about today's genomic tools, most have never used them, and most identified high costs as an impediment to the adoption. Yet, most also expressed their interest in using DNA tests to better predict a number of economically relevant traits, such as reproductive performance, calving ease and feed conversion efficiency. They were also interested in how DNA testing can help them to decide which heifers to keep for breeding. It makes sense because cow-calf operators make more money if they can grow calves more efficiently, using less feed, more quickly, or with fewer losses (Ballenger et al. 2016).

The benefits of genomic testing for the feedlot sector has not been thoroughly studied as well. Potentially, feedlot operators would be interested in cooperating with the seed stock sector to ensure that the animals gain the required weight more quickly in the feedlot environment, are less prone to bovine diseases, or exhibit superior carcass traits. DeVuyst et al. (2011) assessed the relationships of growth and carcass traits and Igenity panel scores for average daily gain (ADG), marbling, rib-eye area, tenderness, fat thickness, and USDA yield grade. Igenity panels predict genetic variation in heifers and steers and are used to score commercial bulls being used as sires. The

panel provides 1-10 scores for 16 traits and 3 indexes, in order to rank cattle by merit (Neogen 2019). They found low but significant correlations between carcass measurements and corresponding Igenity panel scores (DeVuyst et al. 2011). Thompson et al. (2014) measured the value of feedlot operators of using comprehensive tests with information on seven economically relevant traits to manage cattle in the feedlot. They found that marbling and ADG panel scores would generate the biggest benefits, but the costs of the tests would outsize the returns (Thompson et al. 2014).

In the downstream sectors, genomic tools may attract the interests of restaurants and retail chains who may be willing to work with commercial beef producers, but this is not the industry norm. DNA testing may be attractive by allowing sellers to guarantee a particular quality of end product or for buyers to find it.

Ballenger et al. (2016) have noted the need for more economic analyses to be done in regard to the adoption of genomics research. Economists need to work in collaboration with genomic experts to evaluate which genomic advancements offer the most promise in terms of net return and successful development. Economic analyses could also inform regulatory questions coming in the future, such as the benefits of public or third-party validation service for beef genomic tools (Van Eenennaam et al. 2007).

Imperfect information seems to be one of the big problems in the beef industry. Adversity has stemmed from the ownership structure of the industry, distrust, and imperfect information. Low vertical integration has led to unfamiliarity of participants at different stages with motives, margins, and business strategies at other stages, creating strain among industry participants. In Canada, tensions escalated following BSE (mad cow disease) as fed cattle price dropped precipitously (Schroeder 2003). There has been little to no market signal from the consumer to seedstock, thus no feedback for desirable changes to be made to the product. The high degree of segmentation of the beef industry has been detrimental to the end product and consumer satisfaction. Poor information flow has resulted in the decline in high-yielding car-

casses in the past decade – from 66% to 41%. Over the past thirty years, consumers have been substituting beef for chicken, partly as a result of inconsistent product, primarily in the trait of tenderness (Canovas et al. 2017).

2.2 Information Sharing in the Beef Industry

2.2.1 What are the benefits and risks for a BA to share genomic information?

BAs are not-for-profit organizations that promote respective cattle breeds. Their principal purpose is the registration and identification of animals and the keeping of animal pedigrees (*Animal Pedigree Act* 1985). BAs collect data from their member base to promote breed development and are the sole representatives of their respective breeds, and make strategic decisions on their behalf. Given the increasing need for more data and the growing complexity of targeted traits, data pooling for multi-breed evaluations seems to be a viable option. Let us consider some of the potential benefits and costs associated with sharing information.

The main benefit of sharing information is increased accuracy of genomic evaluations. Compared to dairy cattle, beef cattle breeds typically have a large effective population, making it critical to have a higher reference population to make accurate predictions. This creates a need for using information across different breeds, especially for predicting more complex traits (Meuwissen, Hayes, and Goddard 2016). However, this presupposes that information from one breed can improve predictions for another breed. It has been suggested that combining data from different breeds is only useful if they share a common genetic base (Van Eenennaam et al. 2014). In general, sharing genomic data between two unrelated breeds doesn't seem to be very useful for improving predictions on purebred animals. That being said, research has shown that using genomic information of animals from multiple breeds consistently increased the accuracy of prediction within a particular breed compared to just using that single breed's genomic information (IGS 2018), meaning that other breed's

information can be successfully used to improve one's genetic predictions.

Let us now consider some of the costs of sharing information between BAs. Data is the BA's lifeblood; it determines the quality of its genetic evaluations for the members. Some BAs have made considerable investments into setting up infrastructures and procedures for performing evaluations and amassed large data sets over the years, while many smaller BAs are just starting to introduce genetic evaluations. The former may feel that they bring more to the table by sharing information, and breeds who have not invested much capital into developing their data infrastructure would disproportionately benefit from a sharing agreement. It is not clear to what extent this is a concern to larger BAs.

Regardless of their attitude, BAs may not be able to directly control the sharing of their data. When commercial buyers purchase purebred animals or breeding material, they have the right to perform genetic evaluation on the progeny where they want (Kemp and Atkins 2019), usually in the form of "composite cattle"[1]. They may choose to do their evaluation with another BA's evaluation lab if it performs composite EPD evaluations. Thus, the second BA gains access to the first BA's data through the composite cattle. This situation is illustrated with the example of International Genetics Solutions (IGS). IGS is a result of a collaboration between the American Simmental Association and the Red Angus Association of America, which started in 2010 and today is a collaboration of 15 BAs (IGS 2021). Technically, IGS is the American Simmental Association's evaluation lab, but it has developed into a multi-breed genetic evaluation system, possibly the largest in the world. They have collected data on more than 19 million animals (Brink 2020) with 400,000 new animals added annually (IGS 2018). Roughly 40% of the IGS database is Angus, the bulk of which has come in the form of composite cattle from commercial producers (Kemp

---

[1]Composite cattle are a result of crossing two or more existing breeds to improve hybrid vigour (Future Beef 2011). Hybrid vigour is the improved function of any biological quality in a hybrid offspring compared to its parents. For example, crossing Angus and Brahman creates advantages in fertility, longevity, adaptability and mothering ability due to heterosis, resulting in the Brahman composite breed (American Cattlemen 2014).

and Atkins 2019). Thus, IGS does not need the explicit consent of other BAs to collect their data. Two of the other largest breeds - Hereford and Charolais - do not participate in IGS, yet IGS collects their data through members who have their genomics in their herds. Thus, some data sharing is enabled just by the nature of the beef cattle business.

It is possible that the sharing of purebred data is still relevant and is not made redundant by the data coming from composite cattle. First, the volume and quality of data collection from composite cattle are lower than that obtained from BAs directly. Second, genetic testing is not as ubiquitous among commercial producers, as the majority of them is comprised of small producers (Jelinski and Waldner 2018) who typically sell animals after weaning, so are less interested in using genetic tools. Also, unlike breeders of purebred cattle who are often required to keep good records of animals such as via Whole Herd Reporting, commercial producers are not bound by such obligations. Therefore, it appears that the availability of composite data has not made the direct sharing agreements by BAs redundant at this point in time.

Furthermore, the decision of sharing information may depend on perceived competition. If cow-calf producers tend to cross two BAs' animals together, then it would make sense for the two BAs to share information between each other as their interests are aligned. On the contrary, if cow-calf producers usually choose between the two BAs' animals, the BAs may not want to share their information because they would be sharing with a competitor. There appear to be degrees of compatibility between different breeds based on the types of their economically important traits. British breeds, such as Angus (Black and Red), Hereford (Horned and Polled), Shorthorn, Devon, Welsh and Galloway, have similar traits, such as smaller mature size which is reached at an earlier age, less growth potential, excellent fertility and calving ease, higher carcass quality grades, but carcasses with a lower percentage of saleable product (Greiner 2005a). Continental breeds, which include Charolais, Chianina, Gelbvieh, Limousin, Maine Anjou, Salers, and Simmental, are generally larger

in mature size, reach mature size at an older age, produce carcasses with less fat and a higher percentage of saleable product, have lower quality grades, and produce more calving difficulty when mated to cows of the British breeds (Greiner 2005a). Thus, it may be important use a combination of British and Continental genetics that complement each other in a breeding program to produce an end product that has both acceptable carcass quality and retail product yield (Greiner 2005a).

At the same time, the distinction between a partner and a competitor may not be as clear-cut. Cow-calf producers are encouraged to choose breeds to produce calves that are appropriate for their end use (Greiner 2005a). For example, the choice of a sire will differ for a cow-calf producer who sells crossbred replacement heifers than for a producer who sells all calves at weaning. The former will be primarily interested in maternal and reproductive performance, with growth rate and calving ease a secondary concern. For the latter, growth rate and calving ease should be a priority. A producer practicing retained ownership (keeping the animal until slaughter weight) will be primarily interested in carcass traits. Additionally, heterogeneity in cow-calf producer preferences can affect their valuation of bull traits (Boaitey 2017). These producers may value the same animal differently based on the same information, due to differences in their understanding of the information, production goals, and farm practices. Further, farmers' risk perceptions may play a role due to a lack of familiarity with newer breeding tools such as genomics (Boaitey 2017). Importantly, a cow-calf producer's decision to purchase bulls may be contingent on personal relationships. Bull buyers typically have their preferred suppliers and require a reason to change (Speer 2011).

Nonetheless, BAs seem to be picking up signals from the commercial sector. For example, the Canadian Charolais Association noted that there are more and more Angus-based herds using Charolais bulls, suggesting that commercial cattlemen will be increasingly using Charolais bulls in the future (CCA 2019). If this practice becomes more widespread, it would suggest that breed complementarity is becoming

a more important factor for commercial producers.

A BA may worry that IS may reveal genetic imperfections in their animals, which may carry a reputational risk for the BA, or which could be used against them by other BAs if they found birth defects or other traits where it underperforms. For example, president and CEO of Five Rivers Cattle Mike Thoren mentioned that Angus may be responsible for increasing rates of congestive heart failure in US cow herd, possibly due to Angus's excessive pursuit of maximizing quality grade as a single trait focus at the expense of live performance (Thoren 2020).

Another risk for the BA is that sharing information opens the possibility for comparison between the animals of different breeds. A BA may worry that their animals that rank high in their within-breed evaluation will rank lower when compared to animals of other breeds. This would create a possibility for customers to compare the BA's animals directly to other breeds, which is not as simple because currently within-breed EPDs are only comparable within breeds due to differences in the genetic bases (Kuehn and Thallman 2018). The problem is that EPDs are not measured the same way across breeds. The EPD number depends on the base set by a BA. Currently, each BA chooses their base arbitrarily - it could be set to be a point in time, or a reference of high-accuracy sires (Spangler 2019). This makes EPDs from different breeds incomparable, and to compare them, the EPDs need to be brought to a common base. It has also been suggested that the way BAs set an EPD base has some marketing potential. For example, if the EPD base is set so that half the EPDs are negative and it is generally considered that higher EPDs are better, then a BA can simply add a set number to every animal's EPD, so that virtually all the animals have positive genetic predictions for the trait. In the words of Dr. Matt Spangler, beef genetics specialist and professor at the University of Nebraska-Lincoln: "If your breed's EPDs for growth are so much larger than mine simply because your base is bigger, you can advertise better than I can. That's a large part about the base; how does it make me look compared to my competition?" (Spangler 2019).

The introduction of IGS in 2010 established a single base for all the breeds' EPDs. Dr Bob Hough, retired executive vice president of the Red Angus Association of America, wrote about this process: "In addition to being difficult to do, this was a political football because it required cooperation between BAs and every breed wanted to be seen in the best possible light"(Hough and Halladay 2019). It was noted that with a few recent exceptions, BAs don't typically focus on the ease of across-breed comparisons when they come up with their EPD bases - the breeds working with IGS being the notable exception (Kuehn 2019). The EPDs of the 15 BAs participating in IGS are thus directly comparable (CLA 2015).

Currently, there are ways to bring individual EPD scores to one scale, such as the across-breed EPD adjustment factors produced annually by the U.S. Meat Animal Research Center (USMARC). The USMARC table can be used to adjust the EPD scores of animals from different breeds to make their merits comparable across breeds (Kuehn and Thallman 2019). However, using adjustment tables introduces more inaccuracy and has never seen widespread adoption among commercial cattlemen (Bullock 2015). The best way to compare bulls of different breeds is to pool their data, including crossbred data from known breed percentages, into one genetic evaluation, which is what has been achieved in IGS (Bullock 2015). However, USMARC EPDs are apparently still important for comparing breeds with Angus (Black) or others that are not in the IGS analysis (Hough and Halladay 2019).

Finally, sharing information also carries a risk of homogenizing the BAs' breeding strategies. If we assume that all BAs pursue the optimal animal, or the perfect bull, that is common across all breeds, then sharing information between breeds will bring them closer to the optimal animal thus making their breeds more similar. In a theoretic contribution, Vives (1984) showed that increases in the correlation of the signals erodes profitability under Cournot competition. Essentially, when the data is pooled, individual BAs have the same information as every other BA to make selection decisions and so the decisions will become more similar.

2.2.2 What types of information sharing exist in the industry?

Sharing genomic data between BAs may take several different forms. One way is for two BAs to share data between each other, which already happens to a certain extent. For example, Angus data is merged with smaller BAs who promote their crossbreeds to help make better breeding decisions.

Another way BAs could share information is by having a third-party institution collect the data, perform evaluations and report the EPDs back to each respective BA. Such a third-party institution could be a government initiative like the Canadian Beef Improvement Network (CBIN) being developed by the Canadian Beef Breeds Council (CBBC). This network plans to include breed associations, academia, beef industry groups, service providers and government and is planned to "link data networks and provide analysis tools for commercial and seedstock breeders" (CBBC 2018). CBIN will pool the information on all breeding animals in Canada, rank them on several different indices, and make the information available on a platform where customers would be able to bid on bulls. The initiative is supposed to improve the overall quality of the Canadian herd, while individually there may be winners and losers (Lynch-Staunton 2020). The idea seems to have been around for several years, but the work is still underway despite the support of industry organizations. Money needs to be invested in this project, but it is not clear whether the government or the industry players are ready to pay for it yet.

2.3 Literature review on information sharing

2.3.1 Overview of information sharing

In order to understand what may motivate competing BAs to share their genomic information with each other, it will be useful to review the literature on IS in oligopolies. This literature focuses on understanding the firm's optimal decisions and the resulting

market equilibria in the context of uncertainty. It belongs to the study of Bayesian games – a branch of game theory where players have incomplete information about other players.

There are very few general principles that explain the incentives to share information among oligopolists (Vives 1990). In its simplest form, the following two-stage game between the firms is considered. In the first stage, and before receiving any private information, a firm commits to share/reveal information or to keep it private. The information shared is assumed verifiable in the sense of Tirole (1986). After the firms have received their private signals and reports have been sent, the information agency makes public the information to the participants in the agreement (exclusionary disclosure) or to everyone in the industry (nonexclusionary disclosure). At the second stage, firms compete in a Bertrand or Cournot fashion and a Bayesian game of incomplete information is played. Firms may share information about the current or past behavior of customers, orders and prices, as well as cost and demand conditions.

Information sharing agreements are usually moderated by trade associations that typically distribute an aggregate statistic of firms' private signals (Vives 1990). In monopolistic competition, information can be shared under non-exclusionary or exclusionary disclosure. In the former case, information is shared with every market participant. In the latter case, it is provided to members only. It is clear that under non-exclusionary disclosure it is costly to share information as it is possible to free ride. Under exclusionary disclosure rule, IS may occur if the membership fee is not too high.

The evidence of IS among firms happening in real life is limited. Vives (2006) cites several papers studying IS experimentally prior to 2006. Genesove and Mullin (1999) study information exchange in the Sugar Institute and find that firms do not generally misreport but withhold information. Doyle and Snyder (1999) find study production plan announcements in the trade press in the automobile industry and find that a firm's announcements affect a competitor's responses: announcements

of increased production are met with increased production, which the researchers find to be consistent with announcements signaling a common demand parameter. Christensen and Caves (1997) find that in the pulp and paper industry, unexpected capacity announcements by the rivals promote project abandonment in sub-industries with low concentration levels. Armantier and Richard (2003) simulate a hypothetical agreement to share cost information by American Airlines and United Airlines in Chicago O'Hare airport. They find that IS would improve profitability and moderately harm consumers. Cason (1994) finds that pricing behavior is influenced by IS decisions. Ackert, Church, and Sankar (2000) find that in a Cournot game with cost uncertainty, where it cannot be verified if the firm received information, when a firm receives information about industry-wide costs, unfavorable information is disclosed but favorable information is withheld. Contrary to theory, when information is about a cost-specific shock, disclosure is not affected by the favorableness of information. Finally, Lemarié and Marcoul (2018) report that pesticide producers create IS clubs where they share information regarding pesticide resistance with competitors.

One of the main reasons the evidence of IS is scant is due to antitrust regulations. Scherer and Ross (1990) note that the law on trade associations is one of the most nuanced branches of antitrust doctrine. The present position in US courts seems to be that IS is not illegal *per se*, and that it should be challenged only if it helps to reach agreements on prices or to restrict competition. The attempts during 1920s and 1930s to form cartels using trade associations to monitor the agreements ended with consent decrees that would regulate trade associations. Nowadays, information exchanges about prices in oligopolistic markets are strictly monitored. In general, antitrust authorities look with suspicion at information exchanges of individuals firms' data, prices and quantities, because they can give away collusion attempts (Vives 2006).

2.3.2   Models of information sharing in oligopolies

The models we build in this thesis are based on some of the classic works in the field of IS economics. One of the earliest papers in the IS literature goes back to Novshek and Sonnenschein (1982) who extended the Cournot duopoly model to a case in which firms may acquire, release, or agree to pool information about uncertain demand. They proposed a linear demand where the only uncertainty concerned the value of the quantity intercept term distributed over an interval. The goal was to see how a model of duopoly information equilibrium can give insights about the incentives to acquire, release, or pool information. They considered a market in which there are a number of sources that provide information regarding the value of the random quantity intercept term a, and each source provides an unbiased estimate of a. Each firm receives a signal that contains information about the value of the quantity intercept a of the demand function. This signal is the sum of the information received from its private sources and the pooled sources. Then, the firm chooses its level of output, given the value of the signal it received, which corresponds to the maximum expected profit.

Novshek and Sonnenschein's (1982) work was followed by Vives (1984) in a seminal paper where he considered a duopoly model with firms having private information about an uncertain linear demand. He demonstrated that in the case that the goods are substitutes (not) to share information is a dominant strategy for each firm in Bertrand (Cournot) competition, while the opposite is true for complements.

In another work, Gal-Or (1985) considered an oligopoly with an uncertain market demand and firms observing private signals where each firm decides on how completely it will reveal its signal. It was found that no information sharing is the unique Nash equilibrium of the game regardless of how correlated the private signals are.

Fried (1984) examined a duopoly with private-value uncertainty and he found that sharing information is beneficial to the duopolists, which is the opposite result from Vives (1984) and Gal-Or (1985). Shapiro (1986) analyzed the sharing of private

cost information and also complemented the results of Fried (1984). In general, if the information structure is of the common-value type, the results of Clarke (1983), Vives (1984), Gal-Or (1985), and Li (1985) suggest that no information sharing is the unique equilibrium. By contrast if the information structure is of the private-value type, the results of Fried (1984), Shapiro (1986), and Li (1985) point to the exactly opposite conclusion (Vives 1990).

In our analysis we draw from a recent paper by Marcoul (2020) that analyzed a commercial fishery where fishers compete for a pool of fish and each fisher has a private signal about fish abundance. At first, the case of full IS among all fishers was considered. It was found that fishers have an incentive to share their private signals despite congestion costs, and this increases aggregate harvests while making it harder to predict harvests. Marcoul (2020) then considered the existence of several IS groups without communicating and showed that it often dominates fishing with private information as well as with full IS.

We also rely extensively on two works in the field of IS economics that studied agricultural markets. Hueth and Marcoul (2006) studied incentives for IS among agricultural intermediaries in imperfectly competitive markets for farm output. They found that information may reduce the expected profit of intermediary firms and even when IS increases expected profits, intermediary firms face a prisoner's dilemma as they have an incentive to withhold information. In another IS paper on agriculture, Lemarié and Marcoul (2018) considered a pesticide market where resistance to pesticides can develop. They distinguished between two cases of demand for pesticides, one where users are able to coordinate their use and one where they are not. They found that when firms can obtain new scientific information, they have an incentive to share information with competitors.

This thesis uses many of the standard principles of the IS literature outlined in subsections 2.3.1 and 2.3.2. We model an oligopoly where, at the first stage, receive private signals and commit to sharing or not sharing information, and at the sec-

ond stage compete in a Cournot fashion. In Chapter 3 BAs each have one private signal that is communicated to a third party evaluation agency. The agency collects the signals and forms an aggregate statistic, which is then distributed under non-exclusionary disclosure to each BA. In Chapter 4 the aggregate signal is distributed under exclusionary disclosure to coalition members. Like in Novshek and Sonnenschein (1982) and Vives (1984), firms (BAs in our case) acquire and share information about an uncertain linear demand. As in Vives (1984), we first analyze the equilibrium when players keep their signals private and then compare the resulting profit with an equilibrium with full IS. We then modify Vives's (1984) model by adding a breeding component which represents a BA's attempt to improve its breed's quality with genomic selection. Like Marcoul (2020) we consider asymmetric equilibria with IS within IS coalitions. Specifically, we analyze a market with two players sharing inside a coalition while the third staying on its own. Finally, we follow Vives's (1984) method for analyzing a duopoly with partial IS where both players commit to sharing an equal number of signals, with the addition of a breeding component.

Chapter 3

A Basic Model of a Market with Genomic Information

3.1 The Underlying Economic Environment

As Breed Associations (BAs) adopt genomic selection[1], there is an increasing need for more data in order to have more accurate results. However, there are only so many animals of the same breed, often not enough to make meaningful predictions especially for the more valuable complex traits (Meuwissen, Hayes, and Goddard 2016). Pooling data *across* breeds has been considered the way forward for the beef industry (Meuwissen, Hayes, and Goddard 2016). However, early attempts failed to significantly increase prediction accuracies (Berry, Garcia, and Garrick 2016). Until recently it was considered that one BA's data was not helpful to another BA because breeds are not close enough genetically to usefully connect it. But with progress in genomic science, researchers are finding ways to successfully connect data from many breeds in multi-breed evaluations (Saatchi et al. 2014). Consequently, information sharing (IS) between breeds is becoming possible and it opens the way to new economic questions: should BAs share its information with others? What are the costs and benefits of IS? In this chapter we will consider an extreme IS arrangement where all BAs decide to pool their data together. We will examine whether the BA is better

---

[1]To recall, genomic selection is the use of statistical methods to estimate the genetic merit of an animal based on the genotypes and phenotypes of its ancestors (Van Eenennaam et al. 2014)

or worse off sharing information in such an arrangement.

We consider a set of BAs whose objective is to maximize profit by selling their product to their beef producer base. A BA, who uses genomic selection technology, strives to provide the best breeding product to its producers. By "breeding product" we mean a registered purebred animal to be used as a sire or a dam in a producer's herd. Alternatively, a breeding product may also be purebred semen purchased for artificial insemination. BAs operate in an oligopolistic market where each BA differentiates itself by defining its own set of animal traits that it wishes to promote through its marketing activity and wants to offer the best profitability to its producer base. Producers expect a breeding product that offers a high performance in economically valuable traits with the guarantee that these predicted yields are accurate and consistent. For instance, a large variance in feed efficiency can introduce substantial uncertainty to the (net) price received for each unit of product and, as such, may be unacceptable for a producer. Similarly, a large variance in an animal's birth weight will introduce a risk of losing calves, which is costly to the producer and may be considered unacceptable.

In a beef market with n (where $n \geq 2$) BAs, the price commanded by a unit of breeding product by BA i will first depend on its total quantity and quality. Second, competition (cooperation) from the other $n - 1$ BAs who promote different traits will also decrease (increase) the price received by BA i because all beef products are ultimately offered on the same (albeit differentiated) market. Therefore, the price (or inverse demand) for BA i's product is

$$P_i\left(\theta, \theta_d, T_i\right) = \underbrace{\overbrace{A_i\left(1 + \theta\right)}^{precision\ component} - \overbrace{T_i \cdot \mathbb{E}\left[\left(\theta - \theta_d\right)^2\right]}^{selection\ component}}_{BI_i\ =\ \text{Breeding Impact}} - q_i - \gamma \sum_{j \neq i} q_j \qquad (3.1)$$

The inverse demand in (3.1) is interpreted as the price net of marginal cost which is constant.[2] The parameter θ (that we later discuss) is an unknown target genetic

---

[2]See for example Novshek and Sonnenschein (1982)

parameter that the BA tries to better reach, while θ_d represents the selection decision ultimately made by the BA. The parameter A_i is the market size for the BA's product.

The parameter T_i relates to the phenotypic trait (or set of traits) promoted by the BA. More precisely, it measures the share of the trait value that can be improved through genomic selection, or the 'improvability' of the trait. A BA with a high T_i promotes a trait that has significant room for improvement through genomic selection, so improving this trait will be very valuable to the BA; conversely a BA with a low T_i promotes a trait which is already well developed and cannot be improved much further. There may be several reasons why a BA would promote one trait and not another. First, it is reasonable to think that BAs are constrained by their breed's gene pool accumulated over many years of selection. Thus, BAs are often bound to promote the traits they already have. Another reason is that cattle buyers have heterogeneous preferences for traits. For example, some producers may value the animal's docility to minimize the risk of injury even if it means foregoing yield.[3]

We assume that A_i is much greater than T_i:

$$A >> T \tag{3.2}$$

which guarantees that a BA receives a strictly positive price in (3.1). Since BAs have already selected for their traits for a long time, T_i contains potential marginal improvement of the trait while the rest of the trait value is contained in parameter A_i. This is true in the real world since all beef breeds sell at a positive price, even if a breed's quality can be significantly improved through selection.

In the basic model, we assume as a first step that BAs are symmetric and identical and thus $A_i = A$ and $T_i = T$ for $i = 1, 2, ..., n$. Therefore, although each trait promoted by each BA is different, the improvability of each trait is identical across all BAs. This assumption, which can be relaxed, simplifies the exposition. To maximize the value of the breeding impact in expression (3.1), the BA collects genomic

---

[3]Boaitey (2017) found heterogeneity in cow-calf producer preferences according to perceptions of risk, calf retention practices and familiarity with genomics.

information regarding trait i and seeks to select the best version of the breed possible given the environmental constraints of its producers. This best possible breed can only be determined by estimating the unknown parameter θ that characterizes the animal. The parameter θ_d represents the *selection decision* made by BA i given the genomic information it has collected. This decision represents an attempt to approach as closely as possible the unknown parameter θ, or the "target" decision that yields the optimal animal.

There are two distinct ways in which the information on the optimal animal drives the BA's net price - the *precision component* and the *selection component*, both featured in the breeding impact of equation (3.1). In the *precision component*, the BA uses better information to make a prediction about an uncertain market size. In other words, a better knowledge of θ helps the producer make better production decisions. In the *selection component*, the BA uses new information to improve its animal breeding to help produce animals with enhanced traits. Finally, in expression (3.1) BI stands for Breeding Impact, which refers to the combined effect of the BA's use of genomic selection via the improvement of the breed's traits and information about the demand on the market price for its breeding product.

We thus assume that the BA's knowledge of the target parameter for the optimal animal θ has a *dual* consequence on the net price of the BA's product. The first relates to the *selection component* in equation (3.1). As the BA obtains better knowledge of θ, it is able to improve selection by developing its best version of the breed, θ_d, that approximates the true optimal animal, θ, more closely. Thus, the quality of the BA's breed is improved, reflecting positively on the price. The knowledge of θ also affects the net price through the parameter A featured in the *precision component* of equation (3.1). Parameter A encompasses information about the market size and production cost for the BA's product. From the perspective of the market size, producers who purchase the BA's animals care about the consistency in the quality of the breed's traits. If the BA's understanding of the optimal animal is imprecise, then it cannot

guarantee a consistent quality of the breed with respect to the trait it promotes. On the other hand, uncertainty around θ introduces uncertainty to the production cost. Better knowledge of the optimal animal brings information on how to take care of the animal so as to have the phenotypic traits realized in the best manner. Thus, uncertainty about θ increases the risk of mismatching the genomics of the animal to its environment, thereby raising the cost of production. For example, reaching optimal results on raising cattle with certain highly nutritious forages like alfalfa often depends on proper management techniques so as to avoid pasture bloating, which can sometimes be fatal. Consequently, a BA will try to minimize the market size and production cost uncertainties by trying to obtain the best knowledge about the optimal animal.

The parameter $\gamma \in [-1, 1]$ refers to market differentiation between different breeds. It describes to what extent animals of different breeds are substitutes or complements in the purchasing decision of the producer. In this work we assume that all breeds have the same degree of differentiation, or $\gamma_{ij} = \gamma$ for $i, j = 1, 2, ..., n$, for simplicity of calculation. Two breeds with $\gamma > 0$ are substitutes and producers will tend to choose one of them. When $\gamma < 0$, breeds are complements and the animals are usually bought together. Thus, we assume that either all breeds are substitutes or all are complements depending on the value of γ.[4] This has to do with trait differentiation and breed complementarity. For the producer, combining some traits is more profitable than combining other traits. Since BAs offer different traits, it is more profitable for the producer to use some breed combinations over others. For example, producers raising terminal cattle tend to combine British breeds - which are known for their strong maternal traits - with Continental breeds - which are known for their fast growth - to obtain a terminal animal that easily calves and quickly gains weight. Thus, we

---

[4]In reality when $n \geq 2$ we will have γ_{ij} meaning that every BA will have a different value of breed complementarity and market differentiation. For simplicity of calculation we assume that all BAs have the same value of γ, which is a limitation of our model. This assumption can be relaxed in the future.

assume that British-Continental breeds have a negative value of γ, meaning that they are complements on the market. On the other hand, pairs of breeds whose traits are not complementary are assumed to have a positive parameter γ, meaning they are substitutes in the market. For example, Angus and Hereford are both British breeds and have similar traits. They are usually crossed to obtain even stronger maternal traits, but crossing them to produce terminal animals is unlikely to be as profitable as crossing Angus and Simmental, which is a combination of British and Continental breeds. Therefore, we expect purchases of British-Continental pairs of breeds to be more common in the production of terminal cattle as opposed to British-British or Continental-Continental breeds.

3.2 Information collection, learning, and the BA's production decision

Each BA collects genomic and phenotypic information from its producer base. This information collected by its producers relates to their breeding stock and is channeled to the BA who, in turn, uses this data to perform the best selection possible. Producing the optimal animal with respect to the genetic trait T_i is (statistically) very unlikely so that, even with the best prediction equations, errors in the breeder's choice will occur, and the BA's selection decision will never fully reach the parameter for the optimal animal, that is $\theta_d \neq \theta$. The notation $\mathbb{E}\left[(\theta - \theta_d)^2\right]$ in expression (3.1) thus refers to the cost of making a decision that deviates from the optimum.

The data collection process is usually done by the BA using Whole-Herd Reporting (WHR). WHR is a system in which seedstock breeders, either voluntarily or mandatorily, collect performance data on each animal and report it to their BA (Hough 2019). This includes keeping a record of the annual production of cows, as well as the performance of every weaned calf. Breeders are often required by their BA to comply with WHR for their registered cattle, a practice which reduces selective bias in performance data. The collected and reported data is used to perform genomic

evaluations and update existing Expected Progeny Differences (EPDs), which are estimates of an animal's genetic value as a parent (Greiner 2005b).

For simplicity, we assume that the outcome of the data collection process is a signal, s_i. This signal allows the BA to, first, strictly improve the breed's performance with regard to the economically valuable trait it proposes to its producers. The BA is able to perform a better selection decision θ_d, by selecting its best version of the breed, as featured in the *selection component* of expression (3.1), and increase the speed of genetic change based on more accurate evaluations of the EPDs. Second, the signal also provides information regarding the unknown (random) parameter θ featured in the *precision component* of expression (3.1).

Initially, the BA treats the unknown parameter as a random variable for which it holds a prior belief. More precisely, it assumes that $\theta \sim N\left(\bar{\theta}_{pr} \equiv 0, \sigma_\theta^2 \equiv 1/\tau_\theta\right)$ where τ_θ denotes the precision (inverse of variance) of the prior held by the BA. Our *precision component* reflects the idea that a BA has a general expectation that the market size and production cost for its breed are encompassed by the net parameter A_i. Yet, both of those things are uncertain, and by introducing the random variable about the optimal animal, θ, into $A_i (1 + \theta)$ we aim to represent variations around A_i due to this uncertainty.

Formally, we assume that the signal s_i, which imperfectly reveals θ, is drawn from a normal distribution S_i defined by $S_i = \theta + \varepsilon_i$ where θ is the unknown target genomic parameter to be evaluated and $\varepsilon_i \sim N(0, \tau_\varepsilon)$ is a noise whose precision is $\tau_\varepsilon > 0$. We assume that θ and ε are jointly independent which implies that S_i is an unbiased estimator of θ; that is $\mathbb{E}[S_i|\theta] = \theta$. In addition, we also assume that, for $i = 1, 2, ..., n$, the $\varepsilon_i's$ are iid.

Thus, a BA who receives a signal s_i can calculate a posterior estimate of the mean of θ equal to $\theta_p = (1 - \rho)\bar{\theta}_{pr} + \rho s_i = \rho s_i$ where $\rho \equiv \tau_\varepsilon/(\tau_\varepsilon + \tau_\theta)$ is the correlation parameter between the true θ and the signal s_i (see DeGroot 2005, p. 167). Therefore, ρ is naturally interpreted as the quality of information provided by the genomic

information collection process.

In making its selection decision, the BA behaves as a Bayesian forecaster who calculates the unknown mean θ of a distribution with precision τ. The optimal decision is to pick $\theta_d = \mathbb{E}(\theta)$ which minimizes the value of the risk $\mathbb{E}\left[(\theta - \theta_d)^2\right]$ given by $Var[\theta]$.[5] In this context, given the observation of the genomic information signal $S_i = s_i$, the Bayes decision is $\delta^*(s_i) = \mathbb{E}[\theta|S_i = s_i] = \rho s_i$ where this last expression represents the expectation of the posterior distribution. The Bayes risk is the variance of the posterior and is equal to $\eta^* = \mathbb{E}[Var(\theta|s_i)] = 1/(\tau_\varepsilon + \tau_\theta)$.[6]

Remark 3.1 *It can be shown that the expected (squared) error* $\mathbb{E}\left[(\theta - \theta_d)^2\right]$ *equals the Bayes risk when the random variable θ is replaced by $BA'_i s$ estimate ρs_i :*

$$
\begin{aligned}
\mathbb{E}\left[(\theta - \rho_i s_i)^2\right] &= \mathbb{E}\left[(\theta - \rho_i(\theta + \varepsilon_i))^2\right] \\
&= \mathbb{E}\left[\theta^2 \rho_i^2 - 2\theta^2 \rho_i + \theta^2 + 2\theta\varepsilon_i \rho_i^2 - 2\theta\varepsilon_i \rho_i + \varepsilon_i^2 \rho_i^2\right] \\
&= \rho_i^2 \mathbb{E}\left[\theta^2\right] - 2\rho_i \mathbb{E}\left[\theta^2\right] + \mathbb{E}\left[\theta^2\right] + 2\rho_i^2 \mathbb{E}[\theta\varepsilon_i] - 2\rho_i \mathbb{E}[\theta\varepsilon_i] + \rho_i^2 \mathbb{E}\left[\varepsilon_i^2\right] \\
&= (1 - \rho_i)^2 \frac{1}{\tau_\theta} + \rho_i^2 \frac{1}{\tau_\varepsilon} \\
&= \frac{1}{\tau_\theta + \tau_\varepsilon}
\end{aligned}
$$

Therefore, conditional on receiving the signal s_i and selecting its best version of the breed (by choosing $\theta_d = \rho s_i$), the new private estimate of the BI is

$$
BI_i = A(1 + \rho s_i) - \frac{T}{\tau_\varepsilon + \tau_\theta} \tag{3.3}
$$

Note that in expression (3.3), the unknown parameter θ is replaced by its best estimate ρs_i. From an *ex ante* standpoint, the breeding impact is a random variable which depends on the s_i brought by the information collection. The expected breeding

---

[5]Rewrite $E\left[(\theta - \theta_d)^2\right]$ as $E(\theta)^2 - 2\theta_d E[\theta] + \theta_d^2$ which is a convex second degree polynomial in θ_d. It is minimized for $\theta_d = E[\theta]$ so that $E\left[(\theta - E[\theta])^2\right] = E[\theta]^2 - E[\theta]^2 = Var[\theta]$

[6]Note that unlike the posterior mean, the posterior variance of a normal conjugate is independent of the signal s_i.

impact

$$\mathbb{E}\left[BI_i\right] = A - T/(\tau_\varepsilon + \tau_\theta) \tag{3.4}$$

is strictly increasing in the precision of the signal, τ_ε, whereas its variance

$$Var\left[BI_i\right] = (A)^2 \rho^2 Var\left[s_i\right] \tag{3.5}$$
$$= \frac{A^2 \tau_\varepsilon}{\tau_\theta\left(\tau_\varepsilon + \tau_\theta\right)}$$

increases with the precision of the signal, τ_ε.[7] Therefore, in terms of value, an increase in the expected yield value also comes with an increase in the variability of this value. The implication is that as the BA invests in a more precise signal by improving its genomic evaluation capacity, it actually increases the variability of its breeding impact. Interestingly though, the increase in variance happens at a decreasing rate[8] until the variance converges to its maximum value at A^2/τ_θ. This means that, keeping market size A constant, the BA with a high prior precision τ_θ can learn new information without introducing significant uncertainty about the demand intercept. In contrast to that, if a BA has weak prior knowledge, it can be very risky to learn new information. It seems that if not much is known about the BA's animal, new information can either sharply increase or decrease the intercept of the market demand for this animal. For example, as higher density tests are introduced by a BA, if the BA discovers that a breed has a genetic defect that is hard to eliminate, this new piece of information will sharply decrease the market's valuation of the animal. As a counter example, a discovery that the BA's animal has a rare and valuable trait will significantly increase the market demand intercept. On ther other hand, if much is already known about the BA's animal, then any new signal is unlikely to change the market's valuation of the animal.

We can make another observation about the BI and the value of information.

---

[7] $\frac{\partial Var[BI_i]}{\partial \tau_\varepsilon} = (A)^2 / \left(\tau_\varepsilon + \tau_\theta\right)^2 > 0$

[8] $\frac{\partial}{\partial \tau_\varepsilon} Var\left(BI_i\right) = -2\frac{A^2}{(\tau_\theta + \tau_\varepsilon)^3}$

Lemma 3.1 *In expectation, the marginal value of new information is positive and decreasing.*

Proof. All proofs are in the appendix. ∎

Thus, at first, new information is very valuable to the BA as it uses it to enhance the breed, but as the BA learns more, subsequent information adds less to breed improvement.

Now that we have described the production function of the BAs, we embed it in our market framework.

3.3 Market equilibrium with private genomic information

Each BA i collects its genomic information resulting in a signal s_i which represents the information held by its members. The BA considers this information private. In essence, the BA first receives its genomic information and then makes the best breeding decision possible. Given this decision, it then chooses the quantity of breeding product it markets given the competition by other BAs.

Conditional on receiving signal s_i, the objective function of the representative breeder of BA i is

$$\underset{q_i}{Max}\left\{\pi_i\left(q_i,q_j|s_i\right) = \left[A\left(1+\rho s_i\right) - \frac{T}{\tau_\varepsilon+\tau_\theta} - q_i - \gamma\sum_{j\neq i}q_j\right]q_i\right\}. \tag{3.6}$$

Each BA uses a strategy that is *affine* in its own signal s_i that we denote by $q_i = B_0 + B_1 s_i$ where B_0 and B_1 are two real numbers to be determined. In a symmetric equilibrium, this implies that BA i expects BA j with $j \neq i$ to choose $q_j = B_0 + B_1 s_j^e$ where s_j^e is the rational expectation of BA i regarding BA $j's$ signal s_j. However, producer information is kept private by each BA, so BA i does not observe s_j and can only estimate it using its own piece of information s_i; it is easily shown that $\mathbb{E}\left[s_j|s_i\right] = \rho s_i.$[9] In our context, a Perfect Bayesian Equilibrium is a quantity

---

[9]$\mathbb{E}\left[s_j|s_i\right] = \left(1-\rho\right)\bar\theta + \rho s_i = \rho s_i$ since $\bar\theta = 0$. We could relax the assumption that the ε_i are

schedule and a pair of beliefs regarding θ held by each BA i with $i = 1, 2, ..., n$. Each BA holds its own belief (estimate) of θ and also holds an estimate of its competitor's belief (second order belief). Given this pair of beliefs, the quantity decided by each BA is then the best response to the quantity decided by its competitors.

In a Cournot setting, the first order condition of expression (3.6) is

$$\left(A - \frac{T}{\tau_\varepsilon + \tau_\theta} \right) + A\rho s_i = 2q_i + \gamma \sum_{j \neq i} q_j \text{ for } i \neq j \text{ and } i = 1, 2, ..., n.$$

Replacing q_i and q_j by $q_i = B_0 + B_1 s_i$ and $q_j = B_0 + B_1 \mathbb{E}[s_j | s_i]$, respectively, we get

$$A - \frac{T}{\tau_\varepsilon + \tau_\theta} + A\rho s_i = [2 + \gamma(n-1)] B_0 + [2 + \gamma(n-1)\rho] B_1 s_i,$$

which yields by identification

$$B_0 = \frac{A - \frac{T}{\tau_\varepsilon + \tau_\theta}}{2 + \gamma(n-1)} \text{ and } B_1 = \frac{A\rho}{2 + \gamma(n-1)}.$$

Therefore, the quantity offered by BA i, as a function of genomic information s_i, in equilibrium[10] is

$$q_i = \frac{A - \frac{T}{\tau_\varepsilon + \tau_\theta}}{2 + \gamma(n-1)} + \frac{A\rho}{2 + \gamma(n-1)} s_i \text{ for } i = 1, 2, ..., n \tag{3.7}$$

Note that, in expectation, each BA will induce the production of

$$\mathbb{E}[q_i] = \left(A - \frac{T}{\tau_\varepsilon + \tau_\theta} \right) / [2 + \gamma(n-1)]. \tag{3.8}$$

This level corresponds to a world where selection decisions strictly improve the breed's performance in traits without introducing uncertainty. The signal s_i can be interpreted as the latest results obtained from regularly conducted EPD evaluations. When

---

independent. If one assumes that $Cov(\varepsilon_i, \varepsilon_j) = E[\varepsilon_i \varepsilon_j] = \sigma > 0$ for $i \neq j$ then we can show that $E[s_j | s_i] = \rho(1 + \sigma\tau_\theta) s_i$. In this case, the estimate of s_j comes not only from the true parameter θ but also from the correlation of the signal's noise.

[10]There is a restriction on the values of γ in the denominator since the BA's best production strategy given in (3.7) must be positive, and otherwise the BA must stop production. The limitation is $\gamma > -2/(n-1)$ which binds for $n > 3$. It thus allows for a market with three BAs who sell perfect complements. However, as n increases we cannot have full complements anymore. This is realistic because it is hard to imagine a market with BAs selling perfect complements. This limitation on γ, therefore, does not drastically restrict the application of this model.

the BA receives a signal about EPD improvements, $s_i > 0$, this represents encouraging information and producers will increase their production (the implicit assumption here being that, at the time when producers make their *actual* breeding decision, it is hard for them to observe the signal obtained by other BAs). The expected profit is given by

$$\bar{\pi}_i = \underbrace{\frac{\left(A - \frac{T}{\tau_\varepsilon + \tau_\theta}\right)^2}{[2 + \gamma(n-1)]^2}}_{selection\ component} + \underbrace{\frac{\tau_\varepsilon(\tau_\theta + \tau_\varepsilon)A^2}{\tau_\theta[2(\tau_\theta + \tau_\varepsilon) + (n-1)\gamma\tau_\varepsilon]^2}}_{precision\ component}. \tag{3.9}$$

As seen from expression (3.9), the BA's expected profit is strictly positive and entails the *selection component* and the *precision component*. The *selection component* corresponds to the portion of the profit that comes from improving the quality of the animal. It contains the loss function with parameter T, which is minimized given the BA's best available genomic knowledge. This refers to the BA's best attempt at improving breed quality through genomic selection. In this component the increase in signal precision, τ_ε, is transformed into an increase in the accuracy of genomic selection and a strictly better quality of the animal. The *precision component* varies with signal precision via two factors. In the first factor, an increase in signal precision improves the BA's estimation of the market demand and increases the BA's profit. However, in the second factor increasing signal precision also correlates the strategies of all BAs and decreases profit.

The first derivative of the *selection component* with respect to signal precision is given by

$$\frac{\partial \bar{\pi}_i^{selection}}{\partial \tau_\varepsilon} = \frac{2\left(A - \frac{T}{(\tau_\theta + \tau_\varepsilon)}\right)T}{(\tau_\theta + \tau_\varepsilon)^2[(n-1)\gamma + 2]^2} > 0, \tag{3.10}$$

where $\bar{\pi}_i^{selection}$ refers to the portion of BA$_i's$ profit from the *selection component* in (3.9). Note that the numerator is positive under our assumption in (3.2) that A is much greater than T. In expression (3.10) as the BA improves signal precision, it improves its breed's trait and receives more profit as a result. Thus, it always makes sense for a BA to invest in improving signal strength from the perspective of

43

improving the breed.

The first derivative of the *precision component* with respect to signal precision is given by

$$\frac{\partial \bar{\pi}_i^{precision}}{\partial \tau_\varepsilon} = A^2 \frac{2\left(\tau_\theta + \tau_\varepsilon\right) - \gamma \tau_\varepsilon \left(n - 1\right)}{\left(2\left(\tau_\theta + \tau_\varepsilon\right) + \gamma \tau_\varepsilon \left(n - 1\right)\right)^3} \tag{3.11}$$

where $\bar{\pi}_i^{precision}$ refers to the portion of BA$'_i s$ profit from the *precision component* in (3.9). The expression in (3.11) has an ambiguous sign and is positive for $\gamma \in \left(\frac{-2(\tau_\theta + \tau_\varepsilon)}{(n-1)\tau_\varepsilon}, \frac{2(\tau_\theta + \tau_\varepsilon)}{(n-1)\tau_\varepsilon}\right)$ which will bind (be between -1 and 1) for $\mu > 2/\left(n - 3\right)$, where $\mu = \tau_\varepsilon / \tau_\theta$ is the relative quality of genomic information acquired. A high value of μ means that the BA's learning of new information has more precision than the precision of its prior information, or what it already knows. This would refer to a case when the BA is beginning to adopt genomic selection where previously it relied on less precise conventional selection. Thus, in the case when the BA has a strong signal precision relative to prior precision (μ is high) improving signal precision will reduce the *precision component* when BAs are strong substitutes or strong complements. On the other hand, a low value of μ means that the BA already developed a strong genomic knowledge base and its learning precision is lower, to the precision of its prior knowledge. This would be a case of a BA that already adopted genomics selection on a wide scale and collects genomic data from its breeders. Thus, when a BA has a weak signal precision relative to prior precision (μ is low), the *precision component* will be increasing with signal precision as long as there are not too many BAs in the industry. To put the model in perspective, consider the case of Canada where there 26 registered BAs ($n = 26$). In this case, under the assumption that all BAs have the same signal precision τ_ε and the same γ, this would mean that when the BAs' prior precision τ_θ is at least $11\frac{1}{2}$ times greater than signal precision τ_ε (or $\mu \lesssim 0.08$), improving signal precision will increase the *precision component* and the whole profit.

3.4   Sharing genomic information

Let us consider that after receiving signal s_i from the n BAs, a third-party IS group with genetic evaluation capacity is created. It forms the statistic $\bar{s}_i = \Sigma_1^n s_i / n$ which is used to compute new EPDs.[11]   Thus, with IS, each BA benefits from the *enhanced* equation of genomic selection that uses the entire pool of information.[12] Given $\bar{s}$ BA i selects a bull whose Bayes risk (i.e., the variance of the posterior) is equal to $\eta_s^* = \mathbb{E}\left[Var\left(\theta|\bar{s}\right)\right] = 1/\left(n\tau_\varepsilon + \tau_\theta\right)$. The information about the demand also comes with a different (and better) estimate of the variable θ. Given the prior $\theta \sim N\left(\bar{\theta}_{pr} \equiv 0, \sigma_\theta^2 \equiv 1/\tau_\theta\right)$, the posterior estimate of θ in the IS case is given by $\delta_s^*\left(\bar{s}\right) = \left(1 - \rho_n\right)\bar{\theta} + \rho_n\bar{s} = \rho_n\bar{s}$ where $\rho_n \equiv n\tau_\varepsilon/\left(n\tau_\varepsilon + \tau_\theta\right)$ is a correlation parameter between the true θ and the signal $\bar{s}$. Thus, with IS the price for the breeding product is

$$P_i\left(\theta, \bar{s}\right) = A\left(1 + \rho_n\bar{s}\right) - \frac{T}{n\tau_\varepsilon + \tau_\theta} - q_i - \gamma\sum_{j\neq i} q_j. \tag{3.12}$$

Note that with IS, the Bayes risk $\eta_s^* = \mathbb{E}\left[Var\left(\theta|\bar{s}\right)\right] = 1/\left(n\tau_\varepsilon + \tau_\theta\right)$ is strictly lower. Therefore, in expectation, using all the data available from all the BAs results in a breeding impact

$$\mathbb{E}\left[BI_s\right] = A - T/\left(n\tau_\varepsilon + \tau_\theta\right), \tag{3.13}$$

which is strictly greater compared to when only private information is used (see (3.3)).

---

[11]With normally distributed noise there is no loss of informativeness in transmitting the mean of the signals rather than the whole vector of signals (Hölmstrom 1979; Shavell 1979)

[12]the pooled signal $\bar{s}$ has a strictly higher precision compared to BA$_i'$s private signal s_i. To see this, we compare

$$p_{\bar{s}} = \frac{1}{Var\left(\theta + \sum_1^n \varepsilon_i/n\right)} = \frac{n\tau_\theta\tau_\varepsilon}{\tau_\theta + n\tau_\varepsilon}$$

and

$$p_{s_i} = \frac{1}{Var\left(\theta + \varepsilon_i\right)} = \frac{\tau_\theta\tau_\varepsilon}{\tau_\theta + \tau_\varepsilon}$$

to get

$$\frac{p_{\bar{s}}}{p_{s_i}} = 1 + \frac{(n-1)\tau_\theta}{\tau_\theta + n\tau_\varepsilon} > 1$$

At the same time, the variance of the BI when information is pooled is

$$Var\left[BI_s\right] = nA^2\tau_\varepsilon/\tau_\theta\left(\tau_\theta + n\tau_\varepsilon\right) \tag{3.14}$$

which is strictly greater than the variance of the BI when operating on private information.[13] This suggests that IS increases the variance of the intercept, thus making the market demand less predictable. Although in expectation the value of the breeding impact increases, BAs are now subject to a higher variability of the pooled signal. While we do not formally introduce risk aversion in this model, it seems that IS is inherently risky and from this perspective, our model can explain the reluctance of some BAs to enter such pooling agreements. In what follows, we emphasize the change in market structure following IS as a reason for not joining the pool.

Note that IS changes the objective function of the BA because it alters both the stochastic part of the expected profit in (3.9) related to the prediction of the demand and the non-stochastic part related to the improvement in breeding. The latter is our addition to the model of Vives (1984), which introduces nuance to the result, as will be shown later in this work.

The representative breeder of BA_i maximizes the new objective function

$$\underset{q_i}{Max}\left\{\pi_i^s\left(q_i, q_j|\bar{s}\right) = \left[A\left(1 + \rho_n\bar{s}\right) - \frac{T}{n\tau_\varepsilon + \tau_\theta} - q_i - \gamma\sum_{j\neq i}q_j\right]q_i\right\}. \tag{3.15}$$

In the IS case, all BAs have the same information and hence the same strategy $q_i^s = B_0^s + B_1^s\bar{s}$. The first order condition is written as

$$\left(A - \frac{T}{n\tau_\varepsilon + \tau_\theta}\right) + A\rho_n\bar{s} = 2q_i + \gamma\sum_{j\neq i}q_j.$$

Note that $q_i = q_j = B_0^s + B_1^s\bar{s}$ so that

$$\left(A - \frac{T}{n\tau_\varepsilon + \tau_\theta}\right) + A\rho_n\bar{s} = \left[2 + \gamma\left(n - 1\right)\right]B_0^s + \left[2 + \gamma\left(n - 1\right)\right]B_1^s\bar{s}.$$

---

[13]
$$Var\left[BI_s\right] - Var\left[BI_i\right] = A^2\tau_\varepsilon\frac{n - 1}{\left(\tau_\theta + \tau_\varepsilon\right)\left(\tau_\theta + n\tau_\varepsilon\right)} > 0$$

which yields by identification

$$B_0^s = \frac{A - \frac{T}{n\tau_\varepsilon + \tau_\theta}}{2 + \gamma(n-1)} \text{ and } B_1^s = \frac{A\rho_n}{2 + \gamma(n-1)}$$

or

$$q_i(\bar{s}) = \frac{A - \frac{T}{n\tau_\varepsilon + \tau_\theta}}{2 + \gamma(n-1)} + \frac{A\rho_n}{2 + \gamma(n-1)}\bar{s} \text{ for } i = 1, 2, ..., n. \qquad (3.16)$$

Compared to (3.7), the expected yield obtained with each breeding unit is strictly higher since $\mathbb{E}[q_i(\bar{s})] > \mathbb{E}[q(s_i)]$. The expected profit with sharing is given by

$$\bar{\pi}_i^s = \underbrace{\frac{\left(A - \frac{T}{n\tau_\varepsilon + \tau_\theta}\right)^2}{[2 + \gamma(n-1)]^2}}_{selection\ component\ (+)} + \underbrace{\frac{A^2 n\tau_\varepsilon}{\tau_\theta(\tau_\theta + n\tau_\varepsilon)[2 + \gamma(n-1)]^2}}_{precision\ component\ (+)}. \qquad (3.17)$$

We can perform comparative statics of $\bar{\pi}_i^s$ with respect to τ_ε to obtain

$$\frac{\partial \bar{\pi}_i^s}{\partial \tau_\varepsilon} = \underbrace{\frac{2nT\left(A - \frac{T}{\tau_\theta + n\tau_\varepsilon}\right)}{(\tau_\theta + n\tau_\varepsilon)^2(n\gamma + 2 - \gamma)^2}}_{\Delta selection\ component\ (+)} + \underbrace{\frac{A^2 n}{(\tau_\theta + n\tau_\varepsilon)^2(n\gamma + 2 - \gamma)^2}}_{\Delta precision\ component\ (+)} > 0 \qquad (3.18)$$

which is strictly positive. Note that unlike under private information in (3.9), under IS the *precision component* in (3.18) strictly increases with signal precision. Thus, it seems that under IS, learning new information does not make the market more competitive unlike when the BAs keep their information private. Since all the individual signals are shared under IS, there is no threat from the competitors improving their signals and turning them against each other. Thus, under IS, the BA has an incentive to invest in the development of a more precise signal by introducing higher density testing, investing in genomic evaluation infrastructure, hiring geneticists, and other measures that increase signal precision.

We can compare BA$_i's$ profit with private information in (3.9) and profit with IS in (3.17):

$$\Delta\bar{\pi} \equiv \bar{\pi}_i^s - \bar{\pi}_i = \underbrace{\frac{K_s^2 - K^2}{(2 + (n-1)\gamma)^2}}_{selection\ component\ (+)} + \underbrace{\frac{A^2 \tau_\varepsilon(n-1)\left[(2-\gamma)^2(\tau_\theta + \tau_\varepsilon) - n\gamma^2(\tau_\theta + n\tau_\varepsilon)\right]}{(n\gamma - \gamma + 2)^2(\tau_\theta + n\tau_\varepsilon)(2\tau_\theta + 2\tau_\varepsilon - \gamma\tau_\varepsilon + n\gamma\tau_\varepsilon)^2}}_{precision\ component\ (+/-)},$$

$$(3.19)$$

47

where $K_s = A - \frac{T}{n\tau_\varepsilon + \tau_\theta}$ and $K = A - \frac{T}{\tau_\varepsilon + \tau_\theta}$. Whether BAs want to commit to sharing information depends on the two distinct effects of IS. The *selection effect* is always positive and implies that by pooling more genomic information, IS BAs are able to obtain breeding products with strictly higher yield. The *precision component* has an ambiguous sign which depends on the existing product differentiation in breeding.

We can factorize (3.19) to get

$$\Delta\bar{\pi} = \underbrace{\left[\frac{\tau_\varepsilon\,(n-1)}{(\tau_\theta + \tau_\varepsilon)^2\,(\tau_\theta + n\tau_\varepsilon)^2\,[\gamma\,(n-1)+2]^2\,[\tau_\varepsilon\,(n-1)\,\gamma + 2\,(\tau_\theta + \tau_\varepsilon)]^2}\right]}_{(+)}\underbrace{P_A}_{(+/-)}$$

where

$$P_A = \left[(\tau_\theta + \tau_\varepsilon)^2\,(\tau_\theta + n\tau_\varepsilon)\left(-\,(n-1)\,(\tau_\theta + \tau_\varepsilon + n\tau_\varepsilon)\,\gamma^2 + 4\,(1-\gamma)\,(\tau_\theta + \tau_\varepsilon)\right)\right]A^2$$
$$(3.20)$$
$$+ \left[2\,(\tau_\theta + \tau_\varepsilon)\,(\tau_\theta + n\tau_\varepsilon)\,(\tau_\varepsilon\,(n-1)\,\gamma + 2\,(\tau_\theta + \tau_\varepsilon))^2\right]AT$$
$$- \left[(2\tau_\theta + \tau_\varepsilon\,(n+1))\,(\tau_\varepsilon\,(n-1)\,\gamma + 2\,(\tau_\theta + \tau_\varepsilon))^2\right]T^2$$

is a second-degree polynomial in A or T. Inspecting the coefficient in front of the trait improvability parameter T leads to the following proposition.

Proposition 3.1 *IS is always undesirable for high levels of trait improvability T.*

Proof. All proofs are in the appendix. ∎

It seems that when the BA's trait can be significantly improved with genomic selection, the signal is too valuable to be shared with all BAs in the market, and so it will be kept private.

Analyzing (3.20) with respect to market size A yields another proposition.

Proposition 3.2 *Depending on the degree of product differentiation,*

1. *if BAs sell close substitutes ($\gamma_i < \gamma < 1$), IS will be declined (accepted) if BAs have a big (small) market size.*

2. *if BAs sell differentiated products* ($\gamma_{ii} < \gamma < \gamma_i$), *IS will be declined (accepted) if BAs have a small (big) market size.*

Proof. All proofs are in the appendix. ∎

Note that the result in Proposition 3.2 converges to that of Vives (1984) for $T = 0$. To recall, Vives (1984) found that IS will be undesirable for any $\gamma > \gamma_i$. Once we introduce a breeding component where a trait can be improved by T into the model, the result still holds for big BAs, but the reverse is true for moderately sized BAs.

The polynomial in (3.20) is concave or convex with respect to parameter A depending on the sign of a concave polynomial in γ

$$- (n - 1) (\tau_\theta + \tau_\varepsilon + n\tau_\varepsilon) \gamma^2 + 4 (1 - \gamma) (\tau_\theta + \tau_\varepsilon) \tag{3.21}$$

with a positive root

$$\gamma_i = \frac{2}{\sqrt{\rho n^2 + (1 - \rho) n + 1}} \in (0, 1) \tag{3.22}$$

Inspecting (3.22) we find that as ρ approaches 1, the threshold γ_i above which big BAs are unwilling to share information, decreases, suggesting that big BAs will start declining IS at lower values of γ_i (higher differentiation). We make the following proposition.

Proposition 3.3 *Big BAs (BAs with big market size) will be less willing to share information when the signal is informative.*

Proof. All proofs are in the appendix. ∎

Finally we make a proposition about the effect of the number of BAs in the market on the desirability of IS.

Proposition 3.4 *As the number of BAs in the market increases, IS becomes less (more) attractive to big (small) BAs.*

Proof. All proofs are in the appendix. ∎

This result is due to the cost associated with the *correlation of strategies*, which is a major concern for BAs. In simple terms, it is not advantageous to do what others are doing, especially when the others sell similar products. By the same token, it is even less advantageous to share when there are many other BAs who all sell similar products. Our results regarding how γ_i changes with ρ and n are illustrated in Figure 3.1 below.
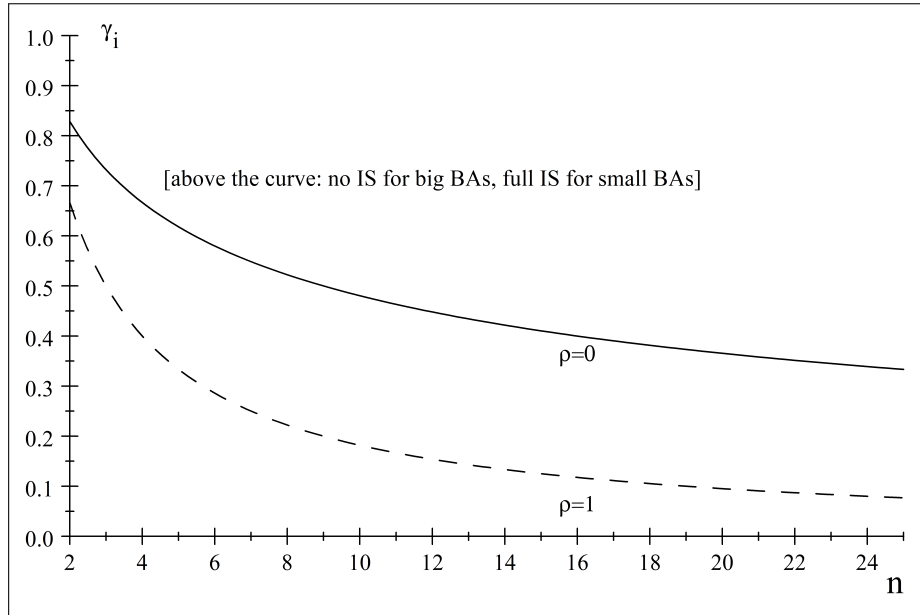
Figure 3.1: Threshold γ_i at which big BAs decline IS falls more precipitously with the number of BAs in the market, n, when the signal quality is high.

3.5 Discussion

First, we have shown in (3.4) that learning new information about the optimal animal increases a BA's expected Breeding Impact by allowing the BA to conduct better genomic selection for its trait. However, learning new information about the breed is risky, especially when the BA is just beginning to introduce genomic selection. Specifically, we have shown that the variance of the breeding impact BI in (3.5) converges to A^2/τ_θ. It increases with signal precision τ_ε, and at a faster rate when the BA's prior information τ_θ is low or market size A is high. It seems that if not much

is known about a breed's genomic content, revealing new information about it has the potential to significantly increase or decrease the market demand for that breed, where the effect of the new information revelation is magnified by the market size A. Thus, a BA with a large market size A will be more susceptible to the positive or negative effect of new information revelation. The model predicts, therefore, that a BA with a large market size but low genomic technology adoption will be the most reluctant to share information. For example, the Canadian Hereford Association is just adopting genomic tools, and is doing it cautiously as it may be more exposed.

We have also shown in (3.1) that the value of new information increases at a decreasing rate. This suggests that the early stages of genomic technology adoption have the highest return on investment, especially when prior knowledge τ_θ is low or trait improvability T is high. While the effect of prior knowledge is intuitive, the effect of trait improvability needs explaining. Recall that a high T indicates that a large portion of the promoted trait can be improved by genomic selection. Therefore, as the BA learns new information when T is high, it derives more value out of the learning process early on because there is much to glean about the trait from new data.

One intuitive finding of this chapter in (3.10) is that the *selection component* of the profit always increases with improved information. This shows that the BA will use the new information to improve the quality of its animal, and from this perspective it is always a good idea to invest in improving the quality of the signal.

However, we have shown that learning new information also makes the market more competitive. We saw in (3.11) that when BAs are just beginning to adopt genomic selection (i.e. when τ_θ is low), then increasing signal precision τ_ε may reduce the BAs' profits via the *precision component* when BAs are strong substitutes or strong complements. This suggests that when new information is valuable, learning new information by all BAs reduces their profit for strong substitutes (and complements). On the other hand, when BAs sell moderately differentiated products, the profit due

to the *precision component* is increasing with signal precision. In other words, when information is valuable to all BAs in the market, then learning new information increases the competition in the market such that improving the signal is only profitable under moderate product differentiation. Furthermore, if we increase the number of BAs n, the market becomes even more competitive and the *precision component* will be reduced with new information even under moderate product differentiation.

With regards to IS, first we found in (3.13) that IS allows a BA to strictly improve the quality of its animal. This finding is intuitive because the main reason for IS is improving the precision of genomic evaluations and speeding up genomic selection. We established that this is indeed happening with IS. However, we found in (3.14) that IS also increases the variance of the intercept, thus making it harder for a BA to predict the market demand. This suggests that IS raises the value of the animal but also makes it less predictable. If BAs vary in their risk aversion, the more risk averse will be reluctant to share.

Also we found in (3.18) that, unlike with private information, under IS learning new information does not make the market more competitive. Since the BAs pool their signals together, there is no cost involved in the improvement of the signal precision of other BAs. That being said, the magnitude of the increase in the *precision component* under IS is inversely related to the market differentiation γ, such that the *precision component* increases the least for perfect substitutes.

In Proposition 3.1 we found that no IS will take place when the BAs' traits are highly improvable through genomics, but full IS will result when the traits are moderately improvable. This suggests that if all BAs have such traits that can be significantly improved by applying genomic selection, BAs will stick to their private information at first, but as the traits get improved, T will decrease and IS will become more attractive to the BAs. So, if the BA's traits are already well developed, IS becomes more likely, since new information ceases to be as valuable as before.

Next, in Proposition 3.2 we found that when products are undifferentiated (γ is

close to 1), it only makes sense for a BA to share information if the market size A is moderate. When products are highly differentiated (γ close to 0), it makes sense to share information only when the market size is large. The key takeaway is that if BAs sell close substitutes, big BAs will decline IS. We see this happening with some large BAs that are cautious about sharing information, such as the top two BAs in Canada. The Canadian Charolais Association and the Canadian Hereford Association.

One important takeaway from Proposition 3.3 is that when the information from the signal is very valuable - i.e., the BAs are learning significantly from new information - big BAs will be wary of sharing their information with other BAs. But as the BAs learn more information, the signal becomes less valuable and the big BAs will be willing to share information if products are reasonably differentiated. This result is reassuring for the beef industry because it suggests that if big BAs possess strong knowledge they will be more willing to share information.

Finally, in Proposition 3.4 we found that big BAs will prefer not to share information when the number of BAs in the market increases. This suggests that sharing information with all BAs in the market when the market size is big for all BAs is unsustainable. To recall, IS correlates strategies, so it becomes less beneficial to correlate strategies across a larger number of BAs.

In the next chapter we will consider the existence of IS coalitions and assess their viability in the context of seedstock beef industry.

Chapter 4

Information Sharing in Coalitions

4.1 The underlying economic environment

In the previous chapter we considered private information equilibria versus full information sharing (IS) equilibria. Full IS can be obtained for a large set of parameters whenever breed associations (BAs) have to make a 'once and for all' decision to share information or not. However, the reality of IS is often different. Typically, a third party invites all BAs to join an IS agreement without imposing a commitment to a simultaneous and irrevocable decision. In practice, a BA may consider who else is joining before deciding whether to enter the agreement. In other cases, all BAs may initially decide to join the IS agreement but one of them (or a subset of them) may decide it would rather be on its own. All this makes it difficult to predict whether IS, or what type of IS, will take place. One possible IS arrangement in between the two extremes of full IS and no IS is sharing within groups, or coalitions. In this chapter we will consider a simple case of coalitional IS- a market where two BAs form a coalition and one BA stays on its own. We wonder if such an arrangement will be stable in the long run; i.e., all coalition members are better off staying in the coalition and not allowing the outsider BA to join. We also want to see for which values of parameters a coalition will either dissolve, resulting in no IS, or invite the outsider BAs to join, resulting in full IS. Would a single BA prefer to stay out of the coalition and forgo

the benefit of improved genomic selection[1]? Would the coalition BAs be better off without a third member? The results of this chapter suggest that IS in coalitions is quite likely. In fact, we find that leaving it up to each BA to decide whether to share information or not is unlikely to result in full sharing, rather, we find that sharing in coalitions is the most common IS arrangement.

There are examples of IS happening within groups. For instance, International Genetic Solutions (IGS) - an initiative created in collaboration between American Simmental Association (ASA) and Red Angus Association of America (RAAA) - has been performing genetic evaluations for several BAs who now send their data to IGS. The sixteen BA members of IGS (IGS 2021) benefit from increased evaluation precision by having access to other members' data for an annual fee. While IGS can reportedly only use the data for evaluations, BAs are collaborating more in other areas such as marketing, advertising, education and youth development. IGS is a current example of a BA that creates an IS group. On the other hand, Canadian Beef Breeds Council's Canadian Beef Improvement Network (CBIN) is an example of a third party organization. CBIN is designed to be an industry-wide IS collaboration between BAs, academia and commercial industry (CBBC 2020). Some BAs, for example the Canadian Hereford Association, have expressed interest in joining once CBIN is launched (CHA 2021). However, CBIN is not yet operational, despite a consensus about the need for such an organization (Duckworth 2018). The stated reason is inadequate funding, but it is not clear if the problem is a misalignment of incentives between the participating sides.

Let us start our analysis by identifying reasons IS in coalitions may be beneficial for individual BAs. Arguably, the main benefit of sharing in coalitions is that member BAs can improve their breeding product based on enhanced information precision, while the non-member BAs have to rely solely on their individual signals. Hence, as

---

[1]To recall, genomic selection is the use of statistical methods to estimate the genetic merit of an animal based on the genotypes and phenotypes of its ancestors (Van Eenennaam et al. 2014)

opposed to the two symmetric cases of full IS or no IS, where no BA has an advantage or disadvantage, in asymmetric IS arrangements, such as coalition sharing, member BAs can offer products more favored by the consumer, and possibly expand their market share at the expense of the non-member BAs.

There are also potential costs involved in joining an IS coalition. Information sharing *correlates* BAs' production strategies, thereby reducing the BAs' expected profits (Vives 1984)[2]. This effect is greatest under full IS where BAs must share with all other BAs. Under coalition IS the effect may be less severe because members only share information between each other and can choose which BAs to accept into the coalition. Thus, if IS in coalitions is possible, we should expect BAs with complementary breeds to form coalitions more often than BAs with substitute breeds. There are also other costs of IS in coalitions beyond our model which are more dynamic in nature. BAs already perform their own genomic evaluations using proprietary measures of how well an animal is performing. After joining an IS group, the performance metrics may be brought to a common basis and become comparable across breeds inside the IS group, which may pose new risks as the standardization of performance metrics can either benefit or hurt a BA. An animal that performs well within its own breed's metrics may turn out to be average or inferior when compared against all others in the IS group. It can be argued that BAs who think their animals will perform poorly relative to others may consider abstaining from joining the group. On the other hand, they could potentially benefit the most from new information. Another potential cost of IS in groups has to do with the ownership of information. If IS is organized by a third party, such as a governmental organization, BAs may be more willing to join. On the other hand, if pooling of information is done by a rival BA, it may make more sense to join only if a BA is not a very strong competitor.

---

[2]In Cournot competition, when products are substitutes it is usually a good strategy for a firm to produce a smaller quantity when all competitors are producing large quantities. Conversely, when competitors are producing small quantities, a firm should produce a large quantity to differentiate from its competition (Bulow, Geanakoplos, and Klemperer 1985).

In section 4.2 we lay out the model of a market with a coalition of two BAs and one non-sharing BA. In section 4.3 we find the conditions under which a non-sharing BA will want to join the coalition, if it is given the option. In section 4.4 we examine the coalition BA's choice of staying in the coalition or leaving. In section 4.5 we investigate whether a non-sharing BA will be allowed to join the coalition. Finally, in section 4.6 we bring together the individual decisions of the BAs to find the resulting market equilibria.

4.2 The equilibrium with a coalition of two BAs and one non-sharing BA

In this section we want to investigate the possible existence of a market equilibrium which features genomic IS within a coalition of BAs but not across all BAs. In particular, we consider a market in which two BAs are in an IS coalition and one BA is on its own. As in the basic model in chapter 3, we shall assume that all three BAs are identical. That being said, the two sharing BAs' production strategies are different from the production strategy of the non-sharing BA. This is because the coalition BAs operate with more and better information that comes from a pooled signal. We will consider BA_1 and BA_c's strategies, where BA_1 is the non-sharing BA while BA_c is a member of the coalition of BAs. In our model the BAs belonging to a coalition will make identical production decisions and we only have to consider the actions of a representative member of the coalition. So for simplicity BA_1 will refer to the non-sharing BA, and BA_c will refer to a member of the coalition.

Each BA collects genomic and phenotypic information from its producer base. The prices for BA_1 and $BA_c's$ breeding products are given by P_1 and P_c, which are respectively:

$$P_1\left(\theta, \theta_d, T\right) = A\left(1 + \theta\right) - T \cdot \mathbb{E}\left[\left(\theta - \theta_{d_1}\right)^2\right] - q_1 - \gamma\left(q_c + q_3\right) \qquad (4.1)$$

$$P_c\left(\theta, \theta_d, T\right) = A\left(1 + \theta\right) - T \cdot \mathbb{E}\left[\left(\theta - \theta_{d_2}\right)^2\right] - q_c - \gamma\left(q_1 + q_3\right) \qquad (4.2)$$

BA$_1$ receives its private signal, s_1, and, given the prior $\theta \sim N\left(\bar{\theta}_{pr} \equiv 0, \sigma_\theta^2 \equiv 1/\tau_\theta\right)$, forms its posterior estimate of θ that equals $\delta_1^*(s_1) = (1 - \rho_1)\bar{\theta} + \rho_1 s_1 = \rho_1 s_1$, where $\rho_1 \equiv \tau_\varepsilon/(\tau_\varepsilon + \tau_\theta)$ is the correlation parameter between the true θ and BA$_1'$s signal s_1. BA$_1$ seeks to minimize its loss function $\mathbb{E}\left[(\theta - \theta_{d_1})^2\right]$ by making the selection decision θ_{d_1} with the lowest Bayes risk of $\eta_1^* = \mathbb{E}\left[Var\left(\theta|s_1\right)\right] = 1/(\tau_\varepsilon + \tau_\theta)$. Therefore, conditional on receiving its own signal, s_i, and selecting its best version of the breed (by choosing $\theta_{d_1} = \rho_1 s_1$), BA$_1$ faces the following price (inverse demand) for its product:

$$P_1(\theta, \theta_d, T) = \underbrace{A(1 + \rho_1 s_1) - \frac{T}{\tau_\varepsilon + \tau_\theta}}_{BI_1} - q_1 - \gamma(q_c + q_3) \tag{4.3}$$

where BI stands for Breeding Impact, which refers to the combined effect of BA's use of genomic selection, via improvement of the breed's traits and information about the demand, on the market price for its breeding product. Note that the breeding impact in (4.3) is the same as the breeding impact in the basic model with private information expressed in equation (3.4).

Meanwhile, the two BAs inside the coalition pool their signals, s_2 and s_3, together to form a statistic $s_c = (s_2 + s_3)/2$. Thus, each BA belonging to the coalition benefits from a more precise signal which leads to a better precision of genomic evaluation[3]. Given s_c, BA$_c$ selects for an animal with the lowest Bayes risk equal to $\eta_s^* = \mathbb{E}\left[Var\left(\theta|s_c\right)\right] = 1/(2\tau_\varepsilon + \tau_\theta)$. Furthermore, BA$_c$ has a better estimate of the random

---

[3]To show this, we will compare the precision of BA$_1'$s individual signal s_1 to the precision of BA$_c'$s pooled signal $\bar{s}$. The precision of the signal is the reciprocal of its variance. Thus, BA$_1'$s variance of signal is given by

$$Var(s_1) = Var(\theta + \varepsilon_1) = \underbrace{Var(\theta)}_{\frac{1}{\tau_\theta}} + \underbrace{Var(\varepsilon_1)}_{\frac{1}{\tau_\varepsilon}} + \underbrace{2Cov(\theta, \varepsilon_1)}_{0} = \frac{\tau_\theta + \tau_\varepsilon}{\tau_\theta \tau_\varepsilon}$$

and BA$_c'$s variance of signal is

$$Var\left(\frac{s_2 + s_3}{2}\right) = Var\left(\theta + \frac{\varepsilon_2 + \varepsilon_3}{2}\right) = \frac{1}{\tau_\theta} + \frac{1}{4}\left(\frac{2}{\tau_\varepsilon}\right) = \frac{\tau_\theta + 2\tau_\varepsilon}{2\tau_\theta \tau_\varepsilon}$$

and thus BA$_1'$s precision of signal is $p_1 = \frac{\tau_\varepsilon \tau_\theta}{\tau_\theta + \tau_\varepsilon}$ and BA$_c'$s precision of signal is $p_c = \frac{2\tau_\theta \tau_\varepsilon}{\tau_\theta + 2\tau_\varepsilon}$. Comparing them yields $\frac{p_c}{p_1} = 1 + \frac{\tau_\theta}{\tau_\theta + 2\tau_\varepsilon}$, which is greater than 1, meaning that BA$_c$ has a higher signal precision than BA$_1$.

variable in the demand intercept, θ. Given the prior $\theta \sim N\left(\bar{\theta}_{pr} \equiv 0, \sigma_\theta^2 \equiv 1/\tau_\theta\right)$, BA_c's posterior estimate of θ is $\delta_s^*\left(s_c\right) = \left(1 - \rho_c\right)\bar{\theta} + \rho_c s_c = \rho_c s_c$ where $\rho_c \equiv 2\tau_\varepsilon / \left(2\tau_\varepsilon + \tau_\theta\right)$ is the correlation parameter between the true θ and the coalition's pooled signal s_c. Therefore, a coalition BA faces a price for its product equal to

$$P_2\left(\theta, \theta_d, T\right) = \underbrace{A\left(1 + \rho_c s_c\right) - \frac{T}{2\tau_\varepsilon + \tau_\theta}}_{BI_c} - q_c - \gamma\left(q_1 + q_3\right) \qquad (4.4)$$

Immediately we see that BA_c's Bayes risk $\eta_s^* = 1/\left(2\tau_\varepsilon + \tau_\theta\right)$ is strictly lower than that of the non-sharing BA, which results in a higher demand intercept for BA_c and a better price received by the coalition's members. Note that, unlike in the symmetric model in the previous chapter, BA_1 and BA_c have different demand intercepts. In expectation, the breeding impact, $\mathbb{E}\left[BI_c\right] = A - T/\left(2\tau_\varepsilon + \tau_\theta\right)$ is greater than $\mathbb{E}\left[BI_1\right] = A - T/\left(\tau_\varepsilon + \tau_\theta\right)$. Exploring the variances of the breeding impacts, we can see that BA_c has a higher variance of the intercept[4]. This is in line with our finding in Chapter 3 that IS increases the variance of the intercept. It seems that while being inside of a coalition makes BA_c better-off in terms of increased expected BI, it also makes the return on breeding less predictable: there are states of the world where BA_c's breeding impact is actually lower than $\mathrm{BA}_1's$. Although we do not formally introduce risk aversion, one important consequence of the model is that IS is inherently risky. Once a BA joins an IS coalition and relies on the pooled signal, it has less control over where the breeding impact will go. This could also explain the reluctance of some BAs to join IS coalitions. We therefore expect more risk averse BAs to abstain from IS.

Given the information above, BA_1 and BA_c have the following objective functions,

---

[4]
$$Var\left[BI_1\right] = \frac{A^2 \tau_\varepsilon}{\tau_\theta} \frac{1}{\left(\tau_\theta + \tau_\varepsilon\right)} \text{ versus } Var\left[BI_c\right] = \frac{A^2 \tau_\varepsilon}{\tau_\theta} \frac{2}{\left(\tau_\theta + 2\tau_\varepsilon\right)}.$$

which are respectively:

$$Max_{q_1} \left\{ \pi_1 \left(q_1, q_c | s_1 \right) = \left[A \left(1 + \rho_1 s_1 \right) - \frac{T}{\tau_\varepsilon + \tau_\theta} - \mathbb{E} \left[q_1 | s_1 \right] - \gamma \left(q_c + q_3 \right) \right] \mathbb{E} \left[q_1 | s_1 \right] \right\}$$

(4.5)

and

$$Max_{q_c} \left\{ \pi_c \left(q_c, q_1 | s_c \right) = \left[A \left(1 + \rho_c s_c \right) - \frac{T}{2\tau_\varepsilon + \tau_\theta} - \mathbb{E} \left[q_c | s_c \right] - \gamma \left(q_1 + q_3 \right) \right] \mathbb{E} \left[q_c | s_c \right] \right\}.$$

(4.6)

Each BA uses a strategy that is *affine* in the signal it receives. $BA_1's$ strategy is based on its own signal s_1 that we denote $q_1 = B_0 + B_1 s_1$ where B_0 and B_1 are two real numbers to be determined. On the other hand, BA_c uses a strategy featuring the pooled signal s_c that we denote $q_c = C_0 + C_1 s_c$ where C_0 and C_1 are also two real numbers to be determined. Note that, unlike in the basic model in Chapter 3, the strategies of the BAs are asymmetric. BA_1 expects BA_c to choose $q_c = C_0 + C_1 s_c^e$ where s_c^e is the rational expectation of BA_1 regarding $BA_c's$ signal s_c. It can be shown that $BA_1's$ best estimate of s_c is $\mathbb{E} \left[s_c | s_1 \right] = \rho_1 s_1$. In turn, BA_c expects BA_1 to choose $q_1 = B_0 + B_1 s_1^e$ where s_1^e is $BA_c's$ rational expectation of $BA_1's$ signal s_1. Similarly, it can be shown that $\mathbb{E} \left[s_1 | s_c \right] = \rho_c s_c$.

In a Cournot setting, computing the first order condition of expression (4.5) for the non-sharing BA_1 yields

$$A - \frac{T}{\tau_\varepsilon + \tau_\theta} + A\rho_1 s_1 = 2q_1 + \gamma \left(q_c + q_3 \right).$$

Substituting q_1 and q_c by $q_1 = B_0 + B_1 s_1$ and $q_c = C_0 + C_1 E \left[s_c | s_1 \right]$ where $E \left[s_c | s_1 \right] = \rho_1 s_1$ we get

$$A - \frac{T}{\tau_\varepsilon + \tau_\theta} + A\rho_1 s_1 = \left(2B_0 + 2\gamma C_0 \right) + \left(2B_1 + 2\gamma \rho_1 C_1 \right) s_1,$$

which yields by identification

$$A - \frac{T}{\tau_\varepsilon + \tau_\theta} = 2B_0 + 2\gamma C_0$$

(4.7)

and

$$A\rho_1 = (2B_1 + 2\gamma\rho_1 C_1). \tag{4.8}$$

In an asymmetric equilibrium we need to derive the first order condition of the objective function of BA_c given in (4.6) and set it to zero, which yields:

$$A - \frac{T}{2\tau_\varepsilon + \tau_\theta} + A\rho_c s_c = 2q_c + \gamma(q_1 + q_3).$$

Substituting BA'_cs strategy $q_c = C_0 + C_1 s_c$ and the expectation of $BA'_1 s$ strategy $q_1 = B_0 + B_1 E[s_1|s_c]$ where $E[s_1|s_c] = \rho_c s_c$ we have

$$A - \frac{T}{2\tau_\varepsilon + \tau_\theta} + A\rho_c s_c = (2C_0 + \gamma(B_0 + C_0)) + (2C_1 + \gamma(C_1 + \rho_c B_1)) s_c,$$

which yields by identification:

$$A - \frac{T}{2\tau_\varepsilon + \tau_\theta} = 2C_0 + \gamma(B_0 + C_0) \tag{4.9}$$

and

$$A\rho_c = 2C_1 + \gamma(C_1 + \rho_c B_1). \tag{4.10}$$

Now we can find the constants in the BAs' strategies: B_0, C_0, B_1, C_1. Solving the system of equations comprising (4.7), (4.8), (4.9), and (4.10), we obtain

$$B_0 = \frac{A - \frac{T}{(\tau_\theta + \tau_\varepsilon)}\left[1 + \gamma\left(\frac{\rho_c}{2-\gamma}\right)\right]}{2(\gamma + 1)}, \tag{4.11}$$

$$B_1 = \frac{\rho_1(2 + \gamma(1 - 2\rho_c))}{2(2 + \gamma(1 - \gamma\rho_1\rho_c))}A, \tag{4.12}$$

$$C_0 = B_0 + \underbrace{\frac{\rho_c}{2(2-\gamma)(\tau_\theta + \tau_\varepsilon)}}_{(+)}T, \text{ and} \tag{4.13}$$

$$C_1 = \frac{\rho_c(2 - \gamma\rho_1)}{2(2 + \gamma(1 - \gamma\rho_1\rho_c))}A \tag{4.14}$$

and the equilibrium strategies are

$$q_1 = \frac{A - \frac{T}{(\tau_\theta + \tau_\varepsilon)}\left[1 + \gamma\left(\frac{\rho_c}{2-\gamma}\right)\right]}{2(\gamma + 1)} + \frac{\rho_1(2 + \gamma(1 - 2\rho_c))}{2(2 + \gamma(1 - \gamma\rho_1\rho_c))}As_1 \tag{4.15}$$

61

and

$$q_c = \frac{A - \frac{T}{(\tau_\theta + \tau_\varepsilon)}\left[1 - \frac{\rho_c}{2-\gamma}\right]}{2(\gamma + 1)} + \frac{\rho_c(2 - \gamma\rho_1)}{2(2 + \gamma(1 - \gamma\rho_1\rho_c))}As_c. \qquad (4.16)$$

We can see that, in expectation, BA_1 produces B_0 and BA_c produces C_0 and that $C_0 > B_0$ for all γ such that $-1 < \gamma \leq 1$. Therefore, in expectation, the BA inside the coalition will have a production with a higher value than the BA outside the coalition. This means that if selection decisions strictly improve the breed's performance without introducing uncertainty, the coalition BA produces a better quality animal. This is unlike the baseline model in Chapter 3 where all BAs produced animals of the same quality. In this model specification though, BA_c benefits from sharing information which improves its selection decision, while the BA outside the coalition operates using only its own genomic data. Later in this chapter, we will see that the non-sharing BA will not always want to join the coalition despite having an inferior product quality.

Furthermore, if we compare the components of the strategies q_1 and q_c that relate to the signal, which are respectively s_1 and s_c, we find that this part is strictly greater in q_c.[5] In this asymmetric equilibrium the sharing BAs rely on new information more than the non-sharing BA. This can be explained by the fact that the sharing BA has a higher signal precision, and so the information coming from the signal is more influential in the BA's production decision. In other words BA_1 has less trust in the new information and prefers to do what it always does, while BA_c is more willing to change its production strategy based on the new data.

By plugging in equilibrium strategies in (4.15) and (4.16) into BA_1 and BA_c's objective functions in (4.5) and (4.6) and taking expectations we obtain the first and second order moments of the signals:

$$\mathbb{E}\left[s_1^2\right] = \frac{\tau_\theta + \tau_\varepsilon}{\tau_\theta\tau_\varepsilon}; \ \mathbb{E}\left[s_1 s_c\right] = \frac{1}{\tau_\theta}; \ \mathbb{E}\left[s_1\right] = 0;$$

$$\mathbb{E}\left[s_c\right] = 0; \ \text{and} \ \mathbb{E}\left[s_c^2\right] = \frac{2\tau_\varepsilon + \tau_\theta}{2\tau_\varepsilon\tau_\theta}.$$

---

[5] $q_c - q_1 = \frac{A\tau_\theta\tau_\varepsilon(2-\gamma)}{2(2+\gamma(1-\gamma\rho_1\rho_s))(\tau_\theta+\tau_\varepsilon)(\tau_\theta+2\tau_\varepsilon)} > 0$, for $\gamma > -1$

These moments are necessary to compute the BA's equilibrium profits. For BA_1 we obtain

$$\pi_1 = \underbrace{\left(\frac{A - \frac{T}{(\tau_\theta + \tau_\varepsilon)} \left[1 + \gamma \left(\frac{\rho_c}{2-\gamma}\right)\right]}{2(\gamma+1)} \right)^2}_{\text{selection component}} + \underbrace{\frac{\tau_\varepsilon}{\tau_\theta} A^2 \frac{(\tau_\theta + \tau_\varepsilon) \left[2(2-\gamma)\tau_\varepsilon + (\gamma+2)\tau_\theta\right]^2}{\left[(-2\tau_\varepsilon^2)\gamma^2 + (\tau_\theta + \tau_\varepsilon)(\tau_\theta + 2\tau_\varepsilon)(\gamma+2)\right]^2}}_{\text{precision component}},$$

(4.17)

while for BA_c' it is

$$\pi_c = \underbrace{\left(\frac{A - \frac{T}{(\tau_\theta + \tau_\varepsilon)} \left[1 - \frac{\rho_c}{2-\gamma}\right]}{2(\gamma+1)} \right)^2}_{\text{selection component}} + \underbrace{\frac{\tau_\varepsilon}{2\tau_\theta} A^2 \frac{(\tau_\theta + 2\tau_\varepsilon) \left[(2-\gamma)\tau_\varepsilon + 2\tau_\theta\right]^2}{\left[(-2\tau_\varepsilon^2)\gamma^2 + (\tau_\theta + \tau_\varepsilon)(\tau_\theta + 2\tau_\varepsilon)(\gamma+2)\right]^2}}_{\text{precision component}}.$$

(4.18)

BA_c's selection component is greater than $BA_1's$ selection component, since the former equals $(C_0)^2$ and the latter $(B_0)^2$ and we showed earlier that $C_0 > B_0$ for all γ such that $-1 < \gamma \leq 1$. Therefore we see that BA_c earns a greater profit from breeding since it produces a product of better quality.

Comparing precision components, we find that for any $\sqrt{2} - 2 < \gamma \leq 1$, BA_c has a smaller precision component (see appendix). Thus, on the one hand, being in a coalition allows a BA to improve the quality of the breed and, on the other hand, a coalition BA receives a lower profit from the precision effect if goods are substitutes or even weak complements. From Vives (1984, p.72) we know that an increase in the precision of rivals' information and increases in the correlation of the signals will decrease a BA's expected profit, which explains our result: BAs in the coalition have their signals perfectly correlated because they receive the same pooled signal. However, our analysis departs from Vives (1984) in that we now explore the existing tensions within the coalitional equilibrium. There are other papers that look at this (e.g., Marcoul (2020)) but we are the first to do it in the context of the seedstock sector of the beef industry. In the next two sections we analyze whether BA_1 will want to join an IS coalition and whether BA_c will ever want to leave the IS coalition, respectively.

4.3   Should a non-sharing BA join an IS coalition?

We have seen that the non-sharing BA sells a product of lower quality and quantity than the coalition BA, yet at the same time its precision component is greater for a wide range of the market differentiation parameter γ. Naturally, the question is whether BA_1 would like to join the coalition to improve the quality of its breed, or would it be deterred by the competitive tensions inside the coalition. BA_1 faces a trade-off between improved quality of the breed and the cost of revealing its information to other BAs who may be its competitors. In order to find the best decision for the BA, we need to compare $BA_1's$ profit when it is alone to the profit it would get if it joined the coalition. We found the former in (4.17). With regards to the latter, BA_1 joining the coalition would result in the full IS equilibrium. We already derived the full IS profit in the previous chapter. Adapting it to our case of three BAs by setting $n = 3$ we get:

$$\bar{\pi}_s = \left(\frac{A - \frac{T}{3\tau_\varepsilon + \tau_\theta}}{2\left(\gamma + 1\right)} \right)^2 + \frac{A^2 \frac{3\tau_\varepsilon}{\tau_\theta\left(\tau_\theta + 3\tau_\varepsilon\right)}}{4\left(\gamma + 1\right)^2}. \tag{4.19}$$

Calculating the difference between (4.19) and (4.17) yields

$$\Delta\pi_1 = \underbrace{\frac{K_s^2 - K^2}{\left(2\left(\gamma + 1\right)\right)^2}}_{(+)\text{ selection component}} + \underbrace{\left(\frac{A^2\tau_\varepsilon}{4\left(\gamma + 1\right)^2\left(\tau_\theta + 3\tau_\varepsilon\right)\left[-2\tau_\varepsilon^2\gamma^2 + \left(\tau_\theta + \tau_\varepsilon\right)\left(\tau_\theta + 2\tau_\varepsilon\right)\left(\gamma + 2\right)\right]^2} \right)}_{(+)\text{ precision effect}} \underbrace{\frac{P\left(\gamma, \mu\right)}{(+/-)}}_{} \tag{4.20}$$

with $K_s = A - \frac{T}{3\tau_\varepsilon + \tau_\theta}$ and $K = A - \frac{T}{\left(\tau_\theta + \tau_\varepsilon\right)}\left[1 + \gamma\left(\frac{\rho_c}{2-\gamma}\right)\right]$

and

$$P\left(\gamma, \mu\right) = 4\left(\gamma + 1\right)\left(2 - \gamma\right)^3\mu^3 - \left(2 - \gamma\right)\left(-4\gamma + 28\gamma^2 + 9\gamma^3 - 32\right)\mu^2 \tag{4.21}$$

$$- 2\left(\gamma + 2\right)\left(7\gamma + 8\gamma^2 - 10\right)\mu - \left(2\gamma + \gamma^2 - 2\right)\left(\gamma + 2\right)^2$$

where $\mu = \tau_\varepsilon/\tau_\theta$ represents the relative quality of the genomic information acquired.

The first term in (4.20) is the selection component, and it is always positive (see appendix). In other words, the non-sharing BA will strictly improve its breeding

product by joining the coalition, due to the improved quality of the information it receives. Meanwhile, the second term (the precision component) has an ambiguous sign determined by the polynomial $P(\gamma, \mu)$ given in (4.21). This means that under certain competitive conditions, expression (4.20) is negative, making BA_1 worse-off after joining the IS coalition. In fact, we find that when BA_1 learns substantially from new information compared to what it already knows, it is eager to join an IS group. This happens because the value of learning new information outweighs the cost of sharing information with competitors. But as the BA gains more knowledge, its prior precision increases and the BA derives less value from new information, and at some point the BA starts to take into consideration market differentiation γ. Ultimately, it does not want to join the IS group when γ is close to 1, or in other words, when BAs are strong competitors. Because there is less to be gained from sharing information, the BA pays attention to the trade-off more.

Setting (4.21) to zero yields a threshold $\bar{\gamma}$ (see appendix) for which BA_1 is indifferent between joining IS or not. Figure 4.1 below plots $\bar{\gamma}$ against μ. In the graph, "join" ("not join") means that (4.20) is positive (negative), so BA_1 is better-off (worse-off) sharing its data with the other two BAs. We can see from the figure that on the interval of $\gamma < (\sqrt{3} - 1)$ BA_1 is always better-off joining the IS coalition. On the interval $(\sqrt{3} - 1) < \gamma < 1$, BA_1 will not want to join if $\gamma > \bar{\gamma}$ for a given level of μ. That being said, if $\mu \gtrapprox 2$, then BA_1 is always better-off joining the IS group[6].

Summarized in a table, we have:

	$\mu \lessapprox 2$	$\mu \gtrapprox 2$
$\gamma \in (\sqrt{3} - 1, 1)$	not join (-)	join (+)
$\gamma \in (0, \sqrt{3} - 1)$	join (+)	join (+)

Table 4.1: Non-coalition BA's best decision given relative precision and market differentiation

---

[6]The exact form is $\mu = \frac{1}{24}\left(a + \frac{721}{a} + 1\right) \approx 2.13$, where $a = \sqrt[3]{8857 + 864i\sqrt{397}}$
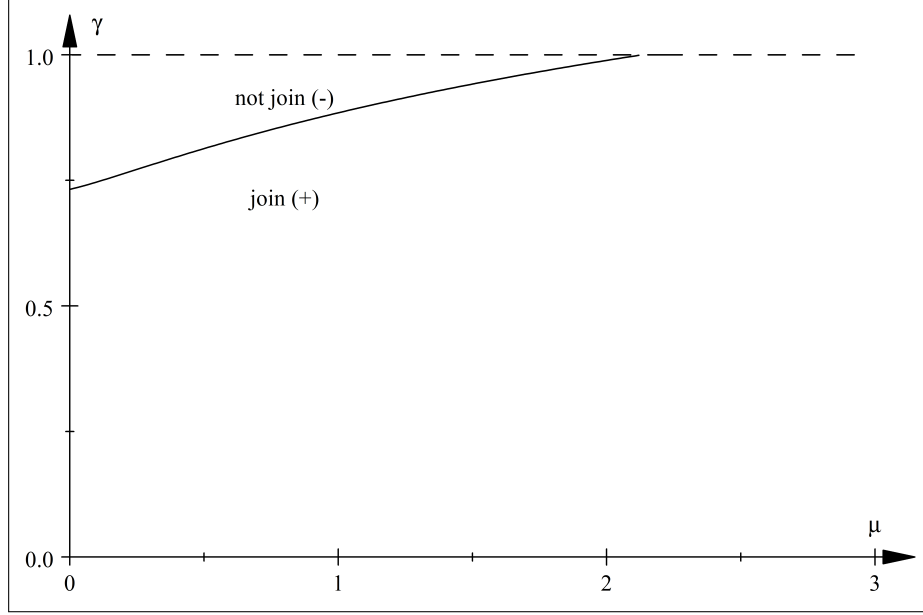
Figure 4.1: BA$'_1s$ decision to join or not to join IS coalition depending on market
differentiation γ and relative quality of information acquired μ

We have the following result:

Proposition 4.1 *In the IS game with coalitions, the relative precisions of the signal
determine the following regimes:*

1. *When BAs do not learn substantially from new information compared to what
   they already know (signal precision is less than twice the prior precision), there
   exists a threshold $\bar{\gamma} \in [0,1]$ such that a BA outside of the IS coalition will not
   want to join the coalition for any $\gamma > \bar{\gamma}$.*

2. *When BAs learn a lot from new information (signal precision is more than twice
   the prior precision), the BA outside the coalition will always want to join the
   IS coalition.*

Proof. All proofs are in the appendix ■

If the BA's prior information is strong relative to what it learns from the signal,
then it will not join IS if the goods are strong substitutes, $\gamma > \bar{\gamma}$. On the other hand,
BA$_1$ will want to join the IS group if it learns a lot from new information, $\tau_\varepsilon / \tau_\theta \gtrless 2$

In Proposition 4.1, we assume that T is small in magnitude because genomic selection represents marginal improvement to the traits that have been selected for a long time, and so the extent to which a trait can be improved with genomics will be small as well. This assumption about T is reasonable because T represents the share of the trait value that can be improved through better selection decisions, while the rest of the trait value is included in the parameter A.

The implications of point 2 of Proposition 4.1 is that BAs that are just starting to introduce genomic programs will want to join the IS group, regardless of market differentiation, γ. Formally we can say that such a BA learns more information from the signal than what it already knows, $\tau_\varepsilon/\tau_\theta \gtrapprox 2$. BAs that are new in the use of genomics likely have less data and lower accuracy, as they are yet to introduce widespread genomic testing for their breeders and/or high-density tests.

We can draw a parallel between our finding in Proposition 4.1 and the prominent example of IS in groups - IGS. American Simmental Association's IGS includes six Canadian BAs as members, two of which are the Canadian Angus Association and the Canadian Simmental Association, which are the two biggest (by number of annual registrations) BAs in Canada. The four others are the Canadian Gelbvieh Association (CGA), the Canadian Limousin Association (CLA), the Canadian Shorthorn Association (CSA), and the Canadian Salers Association (CSA) which are number five, six, seven and eleven in Canada by the number of registrations, respectively (Agriculture and Agri-Food Canada 2021). Noticeably, the third and fourth biggest Canadian BAs are not part of the IGS - Canadian Charolais Association (CCA) and Canadian Hereford Association (CHA). Charolais is a Continental breed like Simmental, and Hereford a British breed like Angus. So we make an argument that CCA and CHA do not wish to share their information with IGS partly because they are strong competitors to IGS's founding members (γ close to 1), and they do not learn as much from new information (signal precision relatively to prior precision, $\tau_\varepsilon/\tau_\theta$, is low), so both BAs would stand to lose from joining IGS, holding other factors constant. While it is

hard to measure precisely how advanced a BA is in its use of genomics, we know that all top four BAs - Angus, Simmental, Hereford and Charolais have adopted genomic testing and use high density tests. Comparatively, CGA, CLA and CSA have much to learn from new information (signal precision relatively to prior precision, $\tau_\varepsilon/\tau_\theta$, is high), and they are willing to join the IS group, holding other factors constant.

There may be other reasons why CCA and CHA are not joining IGS. Although we do not know whether CCA and CHA were ever explicitly invited to join IGS, the latter publicly claims that it is open to new members (Atkins 2021). Of course there are many other factors that can play into a BA's decision to join IGS. One other explanation is that both CCA and CHA have BA counterparts in different countries and they may prefer to share information within their own breeds internationally. However, even in that case they would still benefit from new information. Another explanation is that IGS is a relatively new entity and CHA and CCA may join after all. That being said, Canadian Hereford in their recent report indicated an interest in joining CBIN, rather than IGS. CBIN being a third-party initiative, may be seen as more impartial and thus a more desirable medium of IS.

4.4   When should a BA within the IS group leave?

Another way to consider whether IS in coalitions is viable is to analyze under which conditions a BA will consider leaving an IS agreement. It will be interesting to see whether one of the BAs in the sharing agreement have an incentive to leave. If it does so, there will be a situation when no IS takes place. We need to compare the profit of a BA inside an IS coalition with its profit when there is no IS. Let's take BA_c as an example. We found that BA'_cs equilibrium strategy in (4.16) and can find BA'_cs profit inside of a sharing agreement with another BA by substituting (4.16) into (4.6).

Taking expectations

$$\mathbb{E}\left[s_1^2\right] = \frac{\tau_\theta + \tau_\varepsilon}{\tau_\theta \tau_\varepsilon}; \ \mathbb{E}\left[s_1 s_c\right] = \frac{1}{\tau_\theta}; \ \mathbb{E}\left[s_1\right] = 0;$$

$$\mathbb{E}\left[s_c\right] = 0; \text{ and } \mathbb{E}\left[s_c^2\right] = \frac{2\tau_\varepsilon + \tau_\theta}{2\tau_\varepsilon \tau_\theta}$$

we obtain the coalition BA's expected equilibrium profit

$$\pi_c = \left(\frac{A - \frac{T}{(\tau_\theta + \tau_\varepsilon)}\left[1 - \frac{\rho_c}{2-\gamma}\right]}{2\left(\gamma + 1\right)}\right)^2 + \frac{1}{2}\frac{\tau_\varepsilon}{\tau_\theta}\frac{A^2\left(\tau_\theta + 2\tau_\varepsilon\right)\left[2\left(\tau_\theta + \tau_\varepsilon\right) - \tau_\varepsilon\gamma\right]^2}{\left[\left(-2\tau_\varepsilon^2\right)\gamma^2 + \left(\tau_\theta + \tau_\varepsilon\right)\left(\tau_\theta + 2\tau_\varepsilon\right)\left(\gamma + 2\right)\right]^2}. \tag{4.22}$$

Now we compare (4.22) with BA_c's profit if it stops sharing information, which corresponds to a special case of profit when nobody shares, only now with $n = 3$:

$$\bar{\pi}_i = \left(\frac{A - \frac{T}{\tau_\varepsilon + \tau_\theta}}{2\left(\gamma + 1\right)}\right)^2 + \frac{\tau_\varepsilon\left(\tau_\theta + \tau_\varepsilon\right)A^2}{4\tau_\theta\left(\tau_\theta + \tau_\varepsilon + \gamma\tau_\varepsilon\right)^2}. \tag{4.23}$$

Computing the difference with (4.22), we obtain

$$\Delta\pi = \underbrace{\frac{K_s^2 - K^2}{4\left(\gamma + 1\right)^2}}_{(+)\text{ selection component}} + \underbrace{\frac{A^2\tau_\varepsilon}{4\left(\tau_\theta + \tau_\varepsilon\left(\gamma + 1\right)\right)^2\left[\left(2\tau_\varepsilon^2\right)\gamma^2 - \left(\tau_\theta + \tau_\varepsilon\right)\left(\tau_\theta + 2\tau_\varepsilon\right)\left(\gamma + 2\right)\right]^2}}_{(+)\text{ precision component}}\underbrace{V\left(\gamma, \mu\right)}_{(+/-)} \tag{4.24}$$

with $K_s = A - \frac{T}{(\tau_\theta + \tau_\varepsilon)}\left[1 - \frac{\rho_c}{2-\gamma}\right]$ and $K = \left(A - \frac{T}{\tau_\varepsilon + \tau_\theta}\right)$. Note that $K_s > K$, since $0 < 1 - \frac{\rho_c}{2-\gamma} < 1$ indicating that BA_c earns more profit from trait improvement inside a sharing agreement. As expected, BA_c produces an animal of strictly better quality.

The competitive breeding depends on $V\left(\gamma, \mu\right)$ given by

$$V\left(\gamma, \mu\right) = -\left(2\gamma^2\left(1 - \gamma\right)\left(3 - \gamma\right) - 8\left(1 - \gamma\right)\right)\mu^4 - \left(2 - \gamma\right)\left(7\gamma + 12\gamma^2 - 14\right)\mu^3$$
$$+ \left(-36\gamma - 17\gamma^2 + 4\gamma^3 + 36\right)\mu^2 - \left(20\gamma + 7\gamma^2 - 20\right)\mu + \left(4 - 4\gamma - \gamma^2\right) \tag{4.25}$$

where $\mu = \tau_\varepsilon/\tau_\theta$.

Proposition 4.2 *There exists a threshold $\hat{\gamma} \in (0, 1)$ such that a BA inside an IS coalition will want to leave the IS coalition for $\gamma > \hat{\gamma}$.*

Proof. $V\left(\gamma, \mu\right)$ can be expressed as a fourth-degree expression in γ. Using Descartes' rule, we can show that $V\left(\gamma\right)$ has only one (positive) root $\hat{\gamma}$ which occurs between 0 and 1, such that $V\left(\gamma\right) < 0$ for $\gamma > \hat{\gamma}$. A more detailed proof is in the appendix. ∎

As in Proposition 4.1, in Proposition 4.2 we assume that trait improvability T is small, making the difference in selection components in (4.24) small and positive. Then, for any $\gamma > \hat{\gamma}$, the precision effect becomes negative. Since the selection component is very small, the profit difference is negative in (4.24) and so BA_c is better off quitting the IS coalition when $\gamma > \hat{\gamma}$.
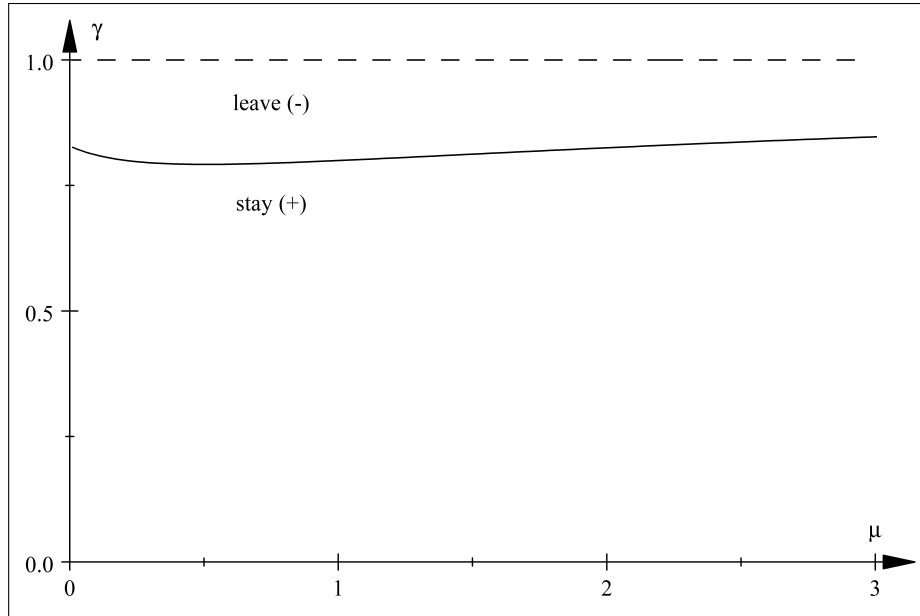

Figure 4.2: coalition BA's decision to stay in the coalition or leave the coalition depending on market differentiation γ and the relative quality of acquired information μ. BA_c will (not) want to leave the coalition above (below) the continuous line.

In Figure 4.2, which depicts the result of Proposition 4.2, "stay" ("leave") refers to an area where profit difference in (4.24) is positive (negative) and BA_c is better off (worse off) sharing information with BA_3. The intuition of this result is that an IS coalition of two BAs is stable as long as all three BAs are not very strong competitors ($\gamma \lessapprox 0.8$). Note also that as signal precision increases, BAs inside the IS coalition will want to stay in the coalition when goods are stronger substitutes. In other words, the threshold of differentiation $\hat{\gamma}$ for which BAs are indifferent between staying or leaving increases as signal precision increases relatively to prior precision. This result is expected because there is value in staying when a BA learns much from

70

new information despite sharing with stronger competitors. This result is in line with Proposition 4.1 where we found that BAs will always want to join the coalition if the BA learns a lot from new information (signal precision at least twice the prior precision). Thus, even for very high market differentiation levels (close, but not equal, to 1), IS in coalitions is possible if μ is very high or when BAs learn a lot of new information. If all BAs are just beginning to collect and use genomic data (μ is high), then IS in coalitions are stable even for very high γ.

4.5 Will BA_1 be allowed to join the coalition?

We can now ask whether the BA outside of the sharing coalition will be allowed in by the coalition members. This will only happen if the coalition members are better-off with an extra member on board. We thus need to compare BA_c's profit inside a coalition of two BAs, to the profit in full IS, which will be the same for all BAs.

Profit when BA_c is sharing with one BA is given in (4.22), while the profit of BA_c when all BAs are sharing is same as full sharing profit with three BAs, given in (4.19). Thus, BA_1 will be allowed inside the IS coalition if $\Delta\pi$ is positive, where $\Delta\pi$ is given by substituting (4.22) from (4.19):

$$\Delta\pi = \bar{\pi}_s \,(\text{full IS}) - \pi_c \,(\text{group IS}) \tag{4.26}$$

$$= \underbrace{\left\{ (K_{full})^2 - (K_{group})^2 \right\}}_{\text{selection component }(+)}$$

$$+ \underbrace{\frac{1}{4} \underbrace{\frac{A^2 \tau_\varepsilon}{(\gamma+1)^2 (\tau_\theta + 3\tau_\varepsilon) \left((2\tau_\varepsilon^2) \gamma^2 - (\tau_\theta + \tau_\varepsilon)(\tau_\theta + 2\tau_\varepsilon)(\gamma+2) \right)^2}}_{(+)} \underbrace{Z(\gamma,\mu)}_{(+/-)}}_{\text{precision component}}$$

where $K_{full} = \dfrac{A - \frac{T}{3\tau_\varepsilon + \tau_\theta}}{2(\gamma+1)}$, $K_{group} = \dfrac{A - T\frac{2(1-\gamma)\tau_\varepsilon + \tau_\theta(2-\gamma)}{(2-\gamma)(\tau_\theta + \tau_\varepsilon)(\tau_\theta + 2\tau_\varepsilon)}}{2(\gamma+1)}$, and

$$Z(\gamma,\mu) = -2\mu^2 (5\mu+1)\gamma^4 + 8\mu(2\mu+1)^2 \gamma^3 - (2\mu+1)\left(12\mu + 3\mu^2 + 5\right)\gamma^2 \tag{4.27}$$

$$- 4(2\mu+1)(5\mu+1)(\mu+1)\gamma + 4(2\mu+1)(\mu+1)^2$$

where $\mu = \tau_\varepsilon / \tau_\theta$.

The analysis of expression (4.26) yields the following result:

Proposition 4.3

1. *The coalition BA will allow the non-sharing BA to join for any $\gamma < 1/5$*

2. *There exists a threshold $\hat{\gamma} \in \left(1/5, 2/5 \left(\sqrt{6} - 1 \right) \right)$ such that a coalition BA will not allow an outsider BA to join the coalition for $\gamma > \hat{\gamma}$.*

3. *The coalition BA will not allow a BA to join for any $\gamma > 2/5 \left(\sqrt{6} - 1 \right)$*

Proof. All proofs are in the appendix. ■

Implicitly plotting $\hat{\gamma}$ yields Figure 4.3 where "allow" ("not allow") refers to the
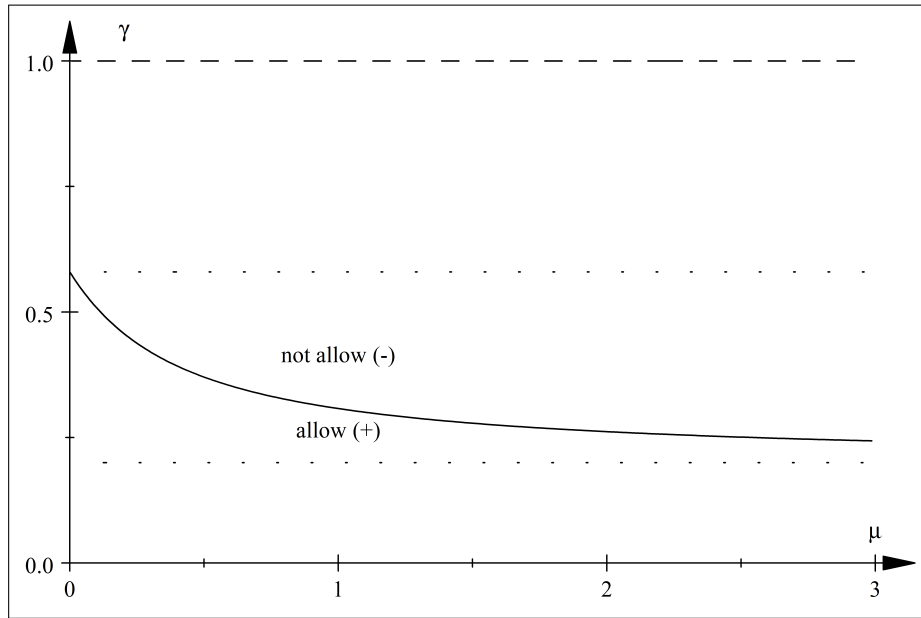

Figure 4.3: Coalition BA's decision to allow the outsider BA to join the IS group or not depending on market differentiation γ and relative quality of genomic information acquired μ. BA_c will (not) want to allow BA_1 to join the coalition below (above) the continuous line.

area where (4.26) is positive (negative) and the coalition BA is better off (worse off) letting the non-sharing BA share information in the coalition. In the figure we can see that for $\gamma \leq 1/5$ the outsider BA will always be allowed to join the coalition.

This result states that when signal precision is infinite, i.e. when BAs learn a lot from new information (μ goes to infinity), the two coalition BAs will not want to allow another BA to join beyond a certain threshold of differentiation $\gamma = 1/5$. At this point the information is too valuable to be shared even with a weak competitor. On the contrary, when BAs already possess information (μ is close to zero), a new BA will be allowed to join up to a point when $\gamma = \frac{2}{5}\left(\sqrt{6} - 1\right) \approx 0.6$. Here, the information is not as valuable and the coalition BAs are more comfortable sharing it.

In Figure 4.3, "allow" also refers to the area where full IS is strictly better than coalition IS, from the point of view of the coalition BA. We can see that a coalition BA will prefer full IS only for low to moderate values of market differentiation and will otherwise prefer coalition sharing. It seems that it is more advantageous for the two coalition BAs to share information against the loner BA for higher values of market differentiation.

4.6 When is IS in coalitions the equilibrium?

Figure 4.4 summarizes the main results of this chapter. The figure is divided into five areas for ease of explanation. The areas correspond to the equilibria that result given market differentiation γ and relative signal precision to prior precision μ. In area (1), the coalition BAs are better off staying in the coalition and allowing the outsider BA to join, while the outsider BA is better off joining. This results in a full IS equilibrium. In area (2), the coalition BAs are better off staying in the coalition, but not allowing the outsider BA to join, while the outsider BA prefers to join. Area (3) is similar except now the outsider BA is better off not joining. Both areas (2) and (3) result in IS coalitions. In area (4), the coalition BAs want to leave and the outsider BA does not want to join the coalition. In area (5) the coalition BAs are better off leaving the coalition while the outsider BA is better off joining the coalition. Thus, in areas (4) and (5) there is no IS.

One of the implications of this result is that IS in coalitions seems to be a com-
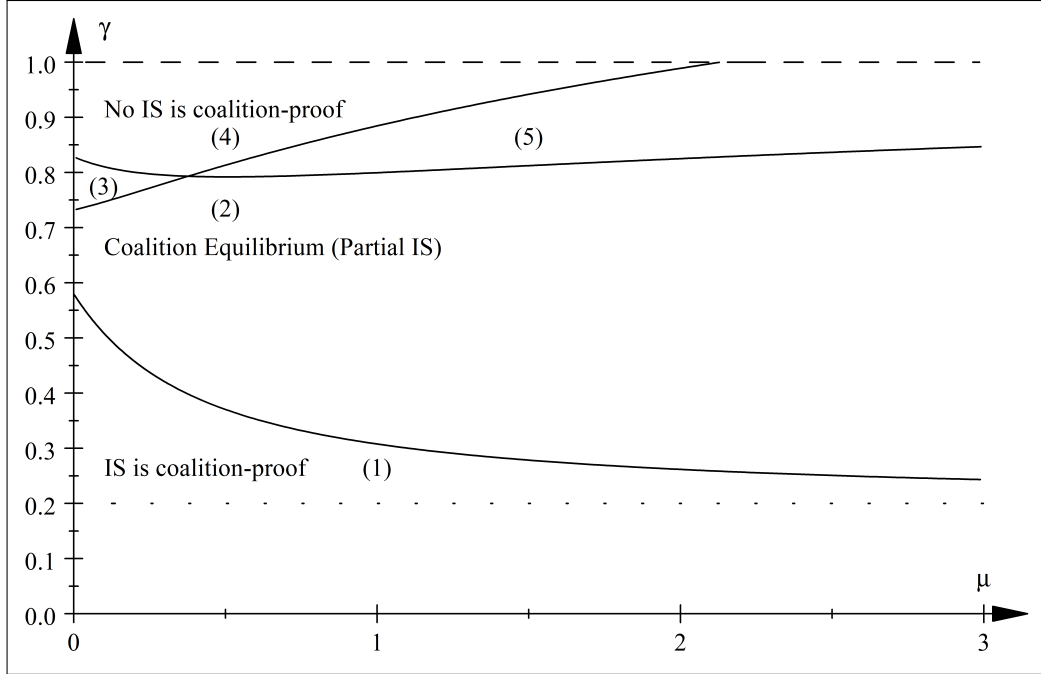
Figure 4.4: Equilibrium strategies of the BAs inside and outside of coalition depending on differentiation γ and relative quality of genomic information acquired μ. There is full IS in area (1), IS in coalitions in areas (2) and (3), and no IS in areas (4) and (5).

mon outcome. IS in coalitions is the dominant outcome on the interval of market differentiation γ between around 0.2 and 0.8. We may assume that most BAs are moderate competitors since they all sell bulls. Even if the breeds are complementary, the choice of a bull of a particular breed still excludes other breeds. In that case, we can see that IS will most likely be happening in the form of coalitions. That being said, on the interval $\gamma = 0.2$ and $\gamma = 0.6$, full IS equilibrium can be brought about if μ is lowered as BAs develop more knowledge while sharing in coalitions. However, on the interval between $\gamma = 0.6$ and $\gamma = 0.8$ there seems to be no way out of sharing in coalitions. In this case, full IS can be imposed by a regulator or a non-sharing BA can offer to compensate coalition members for their losses and join the coalition. For $\gamma > 0.8$ there is no IS unless μ is very high, in which case IS in coalitions becomes viable again. Finally, we can see that when BAs sell complementary or independent goods, full IS is always the preferred choice for all BAs.

74

Chapter 5

Conclusions and Future Research

Advances in genomics over the past two decades have transformed many agricultural industries. Yet, its adoption by the beef industry so far has been slower due to factors like big reference populations, lack of artificial insemination (AI), and fragmented ownership throughout the supply chain. Despite its difficulties, pooling information across different beef breeds has been considered a solution to these problems, which has led to the emergence of information sharing (IS) platforms. IS has the potential to support the declining demand for beef in North America, and to help individual purebred producers to meet the commercial producers' increasing demand for precise genomic evaluations. However, economic costs of IS to individual BAs include the risks related to comparing its animals with other breeds, correlation of production strategies eroding profitability, and the variability of the pooled signal. The benefits of IS to the industry and individual breed associations (BAs) are widely discussed, but no theoretical analysis has been conducted so far to understand the individual BAs' incentives to share information. The aim of this thesis is to formally analyze the viability of IS to individual BAs in the seedstock sector. The following sections will summarize our findings, draw implications and provide suggestions for future research.

5.1   Conclusions

In Chapter 3 we considered equilibria when all BAs are signed into a sharing agreement compared to when no IS occurs. In general, we found that IS increases profitability when the goods on the market are differentiated. However, when the goods are less differentiated, BAs with a big market size will decline IS if they can significantly benefit from privately acquired information. In other words, if BAs with a big market size can use their private signal to gain a deeper insight into the state of the market demand or to improve the signature trait they promote, they will withhold information when goods are undifferentiated. We also found that BAs with a big market size will be more risk averse and will only share when either the competitors sell differentiated products or their private signal alone is not informative enough. That being said, IS makes price more unpredictable as it is now subject to a greater variability of the pooled signal, thus explaining the reluctance of some BAs to share information.

In Chapter 4 we considered asymmetric equilibria where some BAs share information and others not. Specifically, we modeled a group of two BAs sharing information in a coalition, and one BA staying on its own. As in Chapter 3 we found that the degree of product differentiation is important in the BA's decision. Thus, full IS prevails for weak substitutes, and no IS prevails for strong substitutes, while sharing in coalitions prevails in the moderate levels of differentiation. We also found, similarly to Chapter 3, that when the relative value of newly acquired information to known information is high, full IS is less likely to occur and coalitional IS is more likely to prevail. The individual BA will strive to become a coalition member for most levels of product differentiation, especially if the signal is informative. The coalition BAs will allow the individual BA to join only if the products are highly differentiated, except when the signal is not informative - in which case the BA will be admitted even if products are moderate substitutes. The coalition members may themselves

reconsider their membership in the coalition when the goods are strong substitutes. Overall, this leads to the prevalence of coalitional IS for the moderate range of product differentiation.

5.2   Implications

One of the bigger implications of Chapter 3 is that BAs with a big market size will be more likely to participate in IS once they have established a strong genomic evaluation capability. In Chapter 3 we found that BAs with a big market size will decline IS in the situations when their trait can be improved significantly and their signal is informative. These two factors suggest that such a BA has not gained enough knowledge to be open to sharing information. As the BA improves its trait quality and the signal loses informativeness, we can expect this BA with a big market size to share information. For example, the Canadian Charolais Association, one of the four biggest BAs in Canada, has started transitioning to single nucleotide polymorphism (SNP) technology, suggesting they are just building up their genomic knowledge base, which may partly explain their reluctance to share information within an IS organization like IGS.

One important implication of Chapter 4 is that once coalitional sharing takes place, it may be difficult to bring the market to an equilibrium with full IS. From Chapter 4 we see that it is often advantageous for BAs in IS coalitions to withhold information from non-member BAs. This seems to be especially true when the signal is informative as selectively sharing information provides a competitive edge against non-members. Even though BAs typically operate in a differentiated market, even a moderate level of competition may be enough to trigger coalitional sharing. Mechanisms may need to be put in place by which member BAs are compensated for the foregone profit by admitting a new member. Further analysis can determine whether such compensation will be Pareto optimal, meaning that the new member will still be better off joining the coalition while paying the compensation. The exact mechanism of the compensation

in the real world will probably depend on the ownership structure of the IS coalition. For example, IGS is a subsidiary of the American Simmental Association (ASA) made in collaboration with the Red Angus Association of America (RAAA) and thus is not strictly a third-party organization. Thus, these BAs possibly have a larger say in admitting new members into the coalition. On the other hand, the Canadian Beef Breed Council's (CBBC) initiative, Canadian Breed Improvement Network (CBIN), is a third party organization and may not require individual compensation to each coalition member when admitting new BAs.

Another implication of Chapter 4 is that full IS in a market with IS coalitions may eventually result without a need for intervention. This relates to the finding that on the intervals of γ when coalitional IS is the equilibrium for BAs with a big market size, full IS can be achieved as the signal becomes less informative. The suggestion is that BAs with a big market will be willing to admit more members eventually. One caveat is that it can only happen for weak and moderate substitutes, but not for strong substitutes.

5.3 Future Research

Future research can expand the current model. One possibility is to consider each pair of BAs having its own market differentiation level γ corresponding to the breeds' breed complementarity. This would be a closer approximation to the real industry and make the result more generalizable. Another extension would explicitly consider BAs having asymmetric size and signal strengths. This specification would reflect an important reality of the beef industry in North America which is dominated by a few established breeds. We already considered asymmetric IS, so this specification can be an adaptation of our model in Chapter 4. The model can also be extended to make market differentiation γ endogenous in the correlation of signals. As BAs approach the optimal animal in their selection, their breeding products become less differentiated.

The model can also be extended to include vertical information transmission across the sectors of the supply chain. Initiatives exist currently to facilitate the collection and exchange of information across segments. For example, BIXS enables cow-calf producers to share information about its management, performance, health protocols, genetics and carcass data to downstream sectors using individual identification tags. It also facilitates the collection and systematization of data from the feedlot and packer segments to improve selection decisions at the seedstock level (BIXS 2021).

Bibliography

Ackert, Lucy F., Bryan K. Church, and Mandira Roy Sankar (2000). "Voluntary disclosure under imperfect competition: experimental evidence". In: *International Journal of Industrial Organization* 18.1, pp. 81–105.

Agricultural Business Research Institute (2019). *Understanding the EBVs, Selection Indexes and Accuracy*. Web Page. URL: http://abri.une.edu.au/online/pages/understanding_ebvs_char.htm.

Agriculture and Agri-Food Canada (2019). *Federally inspected slaughter plants – cattle and hog*. Web Page. URL: http://www.agr.gc.ca/eng/industry-markets-and-trade/canadian-agri-food-sector-intelligence/red-meat-and-livestock/red-meat-and-livestock-market-information/slaughter/federally-inspected-slaughter-plants-cattle-and-hog/?id=1415860000050.

Agriculture and Agri-Food Canada (Aug. 2021). *Beef Cattle Registrations Report: Total Registrations, 2020, Canada*. URL: https://aimis-simia.agr.gc.ca/rp/index-eng.cfm?action=gR&r=222&signature=B2D1D13F11860DF5D66EDD6BA17284B1&pdctc=&pTpl=1#wb-cont.

Akanno, E. C. et al. (July 2014). "Reliability of molecular breeding values for Warner-Bratzler shear force and carcass traits of beef cattle – An independent validation study". en. In: *J Anim Sci* 92.7, pp. 2896–2904. ISSN: 0021-8812. DOI: 10.2527/jas.2013-7374. URL: https://academic.oup.com/jas/article/92/7/2896/4702163 (visited on 11/20/2019).

Alberta Cattle Feeders Association (2019). *Facts and Stats: A Snapshot of Canada's Beef Industry*. en-CA. URL: https://www.cattlefeeders.ca/industry-overview/alberta-cattle-feeding-facts-and-stats/ (visited on 11/20/2019).

American Cattlemen (May 2014). *BRANGUS GENETICS OFFER HETEROSIS FOR REBUILDING THE NATION'S COWHERD*. URL: https://www.americancattlemen.com/articles/brangus-genetics-offer-heterosis-rebuilding-nations-cowherd#:~:text=BRANGUSGENETICSOFFERHETEROSISFORREBUILDINGTHENATION'SCOWHERD,-PublishedonTue&text=TheBrangusbreed,a3,twohighlysuccessfulparentbreeds..

Animal Pedigree Act (1985). URL: https://laws-lois.justice.gc.ca/eng/acts/A-11.2/page-1.html#h-2869.

Armantier, Olivier and Olivier Richard (2003). "Exchanges of cost information in the airline industry". In: *Rand J.Econ.* Journal Article, pp. 461–477.

Athwal, Rita K. (2002). "Integration of Canadian and U.S. Cattle Markets". In: *Agriculture and Rural Working Paper 53* Journal Article. URL: https://ageconsearch. umn.edu/record/28041/.

Atkins, Jackie (Feb. 2021). *Nobody Shares like IGS*. URL: https://simmental.org/site/ index.php/pub/article-topics/industry-events/340-nobody-shares-like-igs.

Ballenger, Nicole et al. (June 2016). "30 and Daisy: Where's the Economics in Beef Cattle DNA Testing?" eng. In: *Choices* 31.2. DOI: 10.22004/ag.econ.236922. URL: http://ageconsearch.umn.edu/record/236922/files/cmsarticle_510.pdf.

Barkley, Melanie (2012). *Basic Beef Production Guidelines*. Web Page. URL: https: //extension.psu.edu/basic-beef-production-guidelines.

BEEF (July 2021). *Cattle losing adaptations to environment*. URL: https://www. beefmagazine.com/news/cattle-losing-adaptations-environment.

Beef Cattle Research Council (2017). *Genomics*. Web Page. URL: https://www. beefresearch.ca/research-topic.cfm/genomics-56.

Berry, D. P., J. F. Garcia, and D. J. Garrick (2016). "Development and implementation of genomic predictions in beef cattle". In: *af* 6.1, pp. 32–38. ISSN: 2160-6056. DOI: 10.2527/af.2016-0005. URL: https://doi.org/10.2527/af.2016-0005.

BIXS (2021). *Cow-calf*. URL: https://www.bixsco.com/cow-calf.

Boaitey, Albert (2017). "Three Essays on Beef Genomics: Economic and Environmental Impacts". Dissertation/Thesis. PhD thesis. University of Alberta Libraries. DOI: 10.7939/R32B8VR9H. URL: https://search.datacite.org/works/10.7939/ R32B8VR9H.

Bolormaa, S. et al. (2013). "Accuracy of prediction of genomic breeding values for residual feed intake and carcass and meat quality traits in Bos taurus, Bos indicus, and composite beef cattle". In: *J.Anim.Sci.* 91.7, pp. 3088–3104.

Briere, Karen (May 2021). *Beef network takes step forward*. URL: https://www. producer.com/livestock/beef-network-takes-step-forward/.

Brink, Tom (Apr. 2020). *Why Red Angus is Part of International Genetic Solutions*. en-gb. Library Catalog: simmental.org. URL: https://simmental.org/site/ index.php/pub/article-topics/industry-events/243-why-red-angus-is-part-of-international-genetic-solutions (visited on 04/02/2020).

Bullock, Darrh (2015). *Several Major Breeds Combine for Joint Genetic Evaluation*. en. Library Catalog: www.drovers.com. URL: https://www.drovers.com/article/ several-major-breeds-combine-joint-genetic-evaluation (visited on 04/01/2020).

Bulow, Jeremy I, John D Geanakoplos, and Paul D Klemperer (1985). "Multimarket oligopoly: Strategic substitutes and complements". In: *Journal of Political economy* 93.3, pp. 488–511.

Canovas, A. et al. (2017). "Applying new genomic technologies to accelerate genetic improvement in beef and dairy cattle". In: vol. 22. Conference Proceedings, pp. 377–383.

Cason, Timothy N. (1994). "The impact of information sharing opportunities on market outcomes: an experimental study". In: *Southern Economic Journal*, pp. 18–39.

CBBC (2018). *2018 Annual Report | Canadian Cattlemen's Association*. URL: http: //ar.cattle.ca/2018/cbbc.html (visited on 02/28/2020).

CBBC (2020). *Canadian Beef Improvement Network*. URL: http : / / www . canadianbeefbreeds.com/resource/canadian-beef-improvement-network-cbin/.

CCA (2019). *Canadian Charolais Association Annual Report 2018*. Tech. rep.

CCA and BIC (2010). *What Do Beef Cattle Eat?* URL: https://www.ontariobeef. com/uploads/userfiles/files/whatdobeefcattleeat.pdf.

CHA (Jan. 2021). *The Canadian Hereford Digest*. URL: https://en.calameo.com/ read/0060156950a5aff467274.

Christensen, Laurits Rolf and Richard E. Caves (1997). "Cheap talk and investment rivalry in the pulp and paper industry". In: *The Journal of Industrial Economics* 45.1, pp. 47–73.

CLA (2015). "Understanding the new base for Canadian Limousin EPD's". In: URL: http://www.limousin.com/assets/Uploads/performance/2015-Understanding-Limo-EPD.pdf.

Clarke, Richard N. (1983). "Collusion and the incentives for information sharing". In: *The Bell Journal of Economics* Journal Article, pp. 383–394.

De Roos, APW, B. J. Hayes, and M. E. Goddard (2009). "Reliability of genomic predictions across multiple populations". In: *Genetics* 183.4, pp. 1545–1553.

DeGroot, Morris H (2005). *Optimal statistical decisions*. Vol. 82. John Wiley & Sons.

DeVuyst, E. A. et al. (2011). "Relationships between fed cattle traits and Igenity panel scores". In: *J.Anim.Sci.* 89.5, pp. 1260–1269.

Doyle, Maura P. and Christopher M. Snyder (1999). "Information sharing and competition in the motor vehicle industry". In: *Journal of Political Economy* 107.6, pp. 1326–1364.

Duckworth, Barbara (Apr. 2018). *Beef improvement network remains work in progress*. URL: https://www.producer.com/livestock/beef-improvement-network-remains-work-in-progress/.

Dürr, João and Jan Philipsson (2012). "International cooperation: The pathway for cattle genomics". In: *Animal Frontiers* 2.1, pp. 16–21.

Falconer, D. S. (1960). "Introduction to quantitative genetics." In: *Introduction to quantitative genetics*. Journal Article. URL: https://www.cabdirect.org/cabdirect/ abstract/19601603365.

Fan, Jiaping (2018). "Canada's Beef Cattle Industry: Exchange Rates, Price Pass-Through and Feedlot Profitability Under Different Production Systems". Dissertation/Thesis. PhD thesis. University of Alberta Libraries. DOI: 10 . 7939 / R37D2QP9X. URL: https://search.datacite.org/works/10.7939/R37D2QP9X.

Farming Connect. *Genomic Breeding Values*. URL: https://businesswales.gov.wales/ farmingconnect/news-and-events/technical-articles/genomic-breeding-values#:~: text=GenomicBreedingValues(GEBVs)are,EstimatedBreedingValues(EBVs)..

Field, T. G. (2017). *Beef Production and Management Decisions*. 6th ed. Book, Whole. New York, NY: Pearson Education.

Fried, Dov (May 1984). "Incentives for Information Production and Disclosure in a Duopolistic Environment". en. In: *Q J Econ* 99.2, pp. 367–381. ISSN: 0033-5533. DOI: 10.2307/1885531. URL: https://academic.oup.com/qje/article/99/2/367/ 1869930 (visited on 11/21/2019).

Future Beef (Sept. 2011). *Crossbreeding systems for beef cattle*. URL: https://futurebeef.com.au/knowledge-centre/crossbreeding-systems-for-beef-cattle/.

Gal-Or, Esther (1985). "Information sharing in oligopoly". In: *Econometrica: Journal of the Econometric Society* Journal Article, pp. 329–343.

Garrick, Dorian J. (2011). "The nature, scope and impact of genomic prediction in beef cattle in the United States". In: *Genetics Selection Evolution* 43.1, p. 17.

Genesove, David and Wallace Mullin (1999). "The Sugar Institute learns to organize information exchange". In: *Learning by doing in markets, firms, and countries.* University of Chicago Press, pp. 103–144.

Goddard, Ellen et al. (2016). "Improving sustainability of beef industry supply chains". In: *British Food Journal* 118.6, pp. 1533–1552. ISSN: 0007-070X. DOI: 10.1108/BFJ-10-2015-0411. URL: http://www.emeraldinsight.com/doi/abs/10.1108/BFJ-10-2015-0411.

Greiner, Scott Patrick (2005a). "Beef cattle breeds and biological types". In: Publisher: Virginia Cooperative Extension.

Greiner, Scott Patrick (2005b). "Understanding expected progeny differences (EPDs)". In:

Habier, David et al. (2010). "The impact of genetic relationship information on genomic breeding values in German Holstein cattle". In: *Genetics Selection Evolution* 42.1, p. 5.

Harris, B. L., D. L. Johnson, and R. J. Spelman (2009). "Genomic selection in New Zealand and the implications for national genetic evaluation." In: *ICAR Technical Series* 13, pp. 325–330.

Hayes, Ben J., Harris A. Lewin, and Michael E. Goddard (2013). "The future of livestock breeding: genomic selection for efficiency, reduced emissions intensity, and adaptation". In: *Trends in Genetics* 29.4, pp. 206–214.

Hölmstrom, Bengt (1979). "Moral hazard and observability". In: *The Bell journal of economics*, pp. 74–91.

Holstein Canada (2021). *The Holstein Breed.* URL: https://www.holstein.ca/Public/en/About_Us/The_Canadian_Dairy_Industry/The_Holstein_Breed#:~:text=TheHolsteincoworiginatedintheNetherlands.&text=Whilesomebreederswereskeptical,Holsteinshaveoutstandingmilkproduction..

Hough, Bob (Apr. 2019). *Management Perspective: The importance of Whole Herd Reporting.* en. Library Catalog: www.wlj.net. URL: https://www.wlj.net/opinion/management_perspectives/management-perspective-the-importance-of-whole-herd-reporting/article_26dc888e-66da-11e9-b206-93ada621f1a2.html (visited on 09/29/2020).

Hough, Bob and Kerry Halladay (2019). *Comparing Cattle for Crossbreeding.* en. Library Catalog: www.wlj.net. URL: https://www.wlj.net/top_headlines/magazines/comparing-cattle-for-crossbreeding/article_4c21d1b2-1f45-11e9-b47f-b79b96e8310b.html (visited on 04/02/2020).

Hueth, Brent and Philippe Marcoul (2006). "Information Sharing and Oligopoly in Agricultural Markets: The Role of the Cooperative Bargaining Association". In: *American Journal of Agricultural Economics* 88.4, pp. 866–881. ISSN: 0002-9092.

DOI: 10.1111/j.1467-8276.2006.00903.x. URL: http://www.ingentaconnect.com/content/bpl/ajae/2006/00000088/00000004/art00006.

IGS (2018). *FAQs: IGS Multi-breed Genetic Evaluation powered by BOLT*. en-gb. URL: https://www.internationalgeneticsolutions.com/index.php/component/k2/item/33-faqs-multi-breed-genetic-evaluation-powered-by-bolt (visited on 03/01/2020).

IGS (July 2021). *About IGS*.

Industry Voice by Neogen (July 2021). *Seedstock Producers Now in an "All Traits Matter" World*. URL: https://www.beefmagazine.com/genetics/seedstock-producers-now-œall-traits-matter-world.

Jelinski, Murray D. and Cheryl Waldner (Sept. 2018). "Changing demographics of the Canadian cow-calf industry for the period 2011 to 2016". In: *Can Vet J* 59.9, pp. 1001–1004. ISSN: 0008-5286. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6091134/ (visited on 02/26/2020).

Johnston, D. J., B. Tier, and H-U Graser (2012). "Beef cattle breeding in Australia with genomics: opportunities and needs". In: *Animal Production Science* 52.3, pp. 100–106.

Kachman, Stephen D. et al. (2013). "Comparison of molecular breeding values based on within- and across-breed training in beef cattle". In: *Genetics, selection, evolution : GSE* 45.1, p. 30. ISSN: 1297-9686. DOI: 10.1186/1297-9686-45-30. URL: https://www.ncbi.nlm.nih.gov/pubmed/23953034.

Kemp, Chip and Jackie Atkins (2019). *Interview with IGS*.

Khansefid, M. et al. (2014). "Estimation of genomic breeding values for residual feed intake in a multibreed cattle population". In: *J.Anim.Sci.* 92.8, pp. 3270–3283.

Kral, Linni (Aug. 2017). *The Fight for America's Disappearing Ancient Dairy Cows*. URL: https://www.theatlantic.com/science/archive/2017/08/america-heritage-cows/537507/.

Kuehn, Larry (2019). *Interview for Western Livestock Journal*.

Kuehn, Larry and Mark Thallman (2018). *USMARC releases 2017 across-breed EPD tables*. URL: https://www.beefmagazine.com/genetics/usmarc-releases-2017-across-breed-epd-tables (visited on 02/24/2020).

Kuehn, Larry and Mark Thallman (2019). *2019 Across-Breed EPD Table and Improvements*. Beef Improvement Federation. URL: https://beefimprovement.org/wp-content/uploads/2019/12/19_ABEPDpressreleaseandfactsheet.pdf.

Lemarié, Stéphane and Philippe Marcoul (2018). "Coordination and information sharing about pest resistance". In: *Journal of Environmental Economics and Management* 87, pp. 135–149.

Li, Lode (1985). "Cournot oligopoly with information sharing". In: *Rand J.Econ.* Journal Article, pp. 521–536.

Lourenco, Daniela C. et al. (2017). "BIF 2017: The Promise of Genomics for Beef Improvement". In: *A Steak in Genomics*. Conference Proceedings. ACI Information Group. URL: https://www.researchgate.net/profile/Daniela_Lourenco2/publication/321110779_The_Promise_of_Genomics_for_Beef_Improvement/links/5a0e06cbaca27244d285870a/The-Promise-of-Genomics-for-Beef-Improvement.pdf.

Lynch-Staunton, Tom (2020). *Interview.*

Maclachlan, I. R. and E. Stringham (2016). *Beef Cattle Farming.* Web Page. URL: https://www.thecanadianencyclopedia.ca/en/article/beef-cattle-farming.

Marcoul, Philippe (2020). "Fish Abundance Information and Fishery's Aggregate Behavior". In: *Journal of the Association of Environmental and Resource Economists* 7.5, pp. 933–964.

McClure, M. C. et al. (2010). "A genome scan for quantitative trait loci influencing carcass, post-natal growth and reproductive traits in commercial Angus cattle". In: *Anim.Genet.* 41.6, pp. 597–607.

Meat and Livestock Australia (Dec. 2020). *Multi breed evaluations on the way.* URL: https://www.mla.com.au/news-and-events/industry-news/multi-breed-evaluations-on-the-way/.

Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard (2001). "Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps". In: *Genetics* 157.4, pp. 1819–1829. ISSN: 0016-6731. URL: http://www.genetics.org/cgi/content/abstract/157/4/1819.

Meuwissen, Theo, Ben Hayes, and Mike Goddard (2013). "Accelerating improvement of livestock with genomic selection". In: *Annu. Rev. Anim. Biosci.* 1.1, pp. 221–237.

Meuwissen, Theo, Ben Hayes, and Mike Goddard (2016). "Genomic selection: A paradigm shift in animal breeding". In: *Anim Fron* 6.1, pp. 6–14. ISSN: 2160-6056. DOI: 10.2527/af.2016-0002. URL: https://academic.oup.com/af/article/6/1/6/4638797.

Neogen (2019). *Igenity Beef.* Web Page. URL: https://genomics.neogen.com/en/igenity-beef.

Notter, Dave et al. (2016). "Frequently Asked Questions about Estimated Breeding Values and NSIP Roxanne Newton, Kathy Bielek & James Morgan, PhD". In: Generic. URL: https://www.katahdins.org/wp-content/uploads/2016/02/faq.pdf.

Novshek, William and Hugo Sonnenschein (1982). "Fulfilled expectations Cournot duopoly with information acquisition and release". In: *The Bell Journal of Economics* Journal Article, pp. 214–218.

O'Hagan, Maureen (June 2019). *From Two Bulls, Nine Million Dairy Cows.* URL: https://www.scientificamerican.com/article/from-two-bulls-nine-million-dairy-cows/.

Rutherford, Burt (2016). *What's the future for beef cattle genomics?* Web Page. URL: https://www.beefmagazine.com/technology/what-s-future-beef-cattle-genomics.

Ryan, Janine (2016). "Understanding Estimated Breeding Values". In: *Farmer's Weekly* Online Discussion Forum/Blogs. URL: https://www.farmersweekly.co.za/farm-basics/how-to-livestock/understanding-estimated-breeding-values/.

Saatchi, Mahdi et al. (2014). "Large-effect pleiotropic or closely linked QTL segregate within and across ten US cattle breeds". In: *BMC genomics* 15.1, p. 442.

Scherer, Frederic M. and David Ross (1990). "Industrial market structure and economic performance". In: *University of Illinois at Urbana-Champaign's Academy for entrepreneurial leadership historical research reference in entrepreneurship.*

Schmid, Karin (2013). *EPDs: What do all those numbers mean? | BeefResearch.ca.* Web Page. URL: https://www.beefresearch.ca/blog/epds.

Schroeder, Ted C. (2003). "Enhancing Canadian beef industry value-chain alignment". In: *Kansas State University* Journal Article.

Shapiro, Carl (1986). "Exchange of Cost Information in Oligopoly". In: *restud* 53.3, pp. 433–446. ISSN: 0034-6527. DOI: 10.2307/2297638. URL: https://doi.org/10.2307/2297638.

Shavell, Steven (1979). "On moral hazard and insurance". In: *Foundations of insurance economics.* Springer, pp. 280–301.

Spangler, Matt (2019). *Interview for Western Livestock Journal.*

Speer, Nevil C (2011). "Crossbreeding: Considerations and Alternatives in an Evolving Market". en. In: p. 24.

StatCan (Sept. 2020). *Consumption of beef per capita in Canada from 2000 to 2021.* URL: https://www.statista.com/statistics/735166/consumption-of-milk-per-capita-canada/#statisticContainer.

Statistics Canada (2019). *Statistics Canada. Table 32-10-0045-01 Farm cash receipts, annual (x 1,000).* Web Page. URL: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3210004501.

Stoneking, M. (2001). "Single nucleotide polymorphisms. From the evolutionary past..." In: *Nature* 409.6822, pp. 821–822. ISSN: 0028-0836. DOI: 10.1038/35057279. URL: http://europepmc.org/abstract/MED/11236996.

Thompson, Nathanael M. et al. (2014). "Value of genetic information for management and selection of feedlot cattle". In: *J.Agric.Resour.Econ.* Journal Article, pp. 139–155.

Thoren, Mike (2020). *National Western Stock Show Board of Directors Meeting Speaker Recap.* URL: https://www.youtube.com/watch?v=GMR4Jk2t_t4 (visited on 04/02/2020).

Tirole, Jean (1986). "Hierarchies and bureaucracies: On the role of collusion in organizations". In: *JL Econ. & Org.* 2, p. 181.

Twine, Edgar E. (2014). "Canada's Beef Cattle Industry: Shocks, Cycles and Loan Guarantees". Dissertation/Thesis. PhD thesis. University of Alberta Libraries. DOI: 10.7939/R3K64B20R. URL: https://search.datacite.org/works/10.7939/R3K64B20R.

Van Eenennaam, A. L., JHJ Van Der Werf, and M. E. Goddard (2011). "The value of using DNA markers for beef bull selection in the seedstock sector". In: *J.Anim.Sci.* 89.2, pp. 307–320.

Van Eenennaam, A. L. et al. (2007). "Validation of commercial DNA tests for quantitative beef quality traits". In: *J.Anim.Sci.* 85.4, pp. 891–900.

Van Eenennaam, Alison L. and Daniel J. Drake (2012). "Where in the beef-cattle supply chain might DNA tests generate value?" In: *Animal Production Science* 52.3, p. 185. ISSN: 1836-0939. DOI: 10.1071/AN11060.

Van Eenennaam, Alison L. et al. (2014). "Applied Animal Genomics: Results from the Field". In: *Annual Review of Animal Biosciences* 2.1, pp. 105–139. DOI: 10.1146/annurev-animal-022513-114119. URL: https://doi.org/10.1146/annurev-animal-022513-114119.

Vestal, Mallory K. et al. (2013). "The value of genetic information to livestock buyers: a combined revealed, stated preference approach". In: *Agricultural Economics* 44.3, pp. 337–347. ISSN: 1574-0862. DOI: 10.1111/agec.12016. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/agec.12016.

Vives, Xavier (1984). "Duopoly information equilibrium: Cournot and Bertrand". In: *J.Econ.Theory* 34.1, pp. 71–94.

Vives, Xavier (1990). "Trade association disclosure rules, incentives to share information, and welfare". In: *Rand J.Econ.* Journal Article, pp. 409–430.

Vives, Xavier (2006). "Information sharing: economics and antitrust". In: *The Pros and Cons of Information Sharing* 83.Journal Article.

Wickham, Brian W. and João W. Dürr (2011). "A new international infrastructure for beef cattle breeding". In: *Animal Frontiers* 1.2, pp. 53–59.

Witte, Johanna (2018). "The use of market intelligence and producer decision-making as a driver of industry competitiveness: A producer perspective on the Canadian and Irish beef industries". Dissertation/Thesis. PhD thesis. University of Alberta Libraries. DOI: 10.7939/R3SB3XF31. URL: https://search.datacite.org/works/10.7939/R3SB3XF31.

Womach, Jasper (2005). *CRS report for Congress*. Tech. rep. Report. Washington, DC: LoC. URL: https://web.archive.org/web/20110810044532/http://ncseonline.org/nle/crsreports/05jun/97-905.pdf.

Appendix A: Proofs

Proof of Lemma 3.1: First, note that, in expectation, BI_i will increase from $BI_i = A - T/\tau_\theta$ to $BI_i = A$ with signal precision τ_ε. Thus, new information is valuable to the BA as can be shown by taking the derivative of the expectation of BI with respect to signal precision:

$$\frac{\partial}{\partial \tau_\varepsilon} \mathbb{E}\left[BI_i\right] = \frac{T}{\left(\tau_\theta + \tau_\varepsilon\right)^2} > 0. \tag{A.1}$$

However, we can see that each new piece of information is less valuable than the previous one since the the second derivative is negative:

$$\frac{\partial^2}{\partial^2 \tau_\varepsilon^2} \mathbb{E}\left[BI_i\right] = -2\frac{T}{\left(\tau_\theta + \tau_\varepsilon\right)^3} < 0. \tag{A.2}$$

Proof of Proposition 3.1: By inspection, polynomial (3.20) is always concave in T, which means that (3.20) is negative for high values of trait improvability T.

Proof of Proposition 3.2: First note that the polynomial in (3.20) is concave or convex with respect to parameter A depending on the sign of a concave polynomial in γ

$$- \left(n - 1\right)\left(\tau_\theta + \tau_\varepsilon + n\tau_\varepsilon\right)\gamma^2 + 4\left(1 - \gamma\right)\left(\tau_\theta + \tau_\varepsilon\right) \tag{A.3}$$

with roots

$$\gamma_i = \frac{2}{\sqrt{\rho n^2 + \left(1 - \rho\right)n} + 1} \in (0, 1) \tag{A.4}$$

and

$$\gamma_{ii} = -\frac{2}{\sqrt{\rho n^2 + \left(1 - \rho\right)n} - 1} < 0. \tag{A.5}$$

Since the polynomial in gamma (A.3) is concave, it is positive for $\gamma \in (\gamma_{ii}, \gamma_i)$ and negative for $\gamma \in (-1, \gamma_{ii})$ and $\gamma \in (\gamma_i, 1)$. Therefore, for high $\gamma \in (\gamma_i, 1)$ - low market differentiation - the polynomial in A in (3.20) will be concave and IS will be undesirable for big market size A and desirable for low market size A. Conversely, for

moderate $\gamma \in (\gamma_{ii}, \gamma_i)$ - high market differentiation - the polynomial in A in (3.20) will be convex and IS will be undesirable for low market size A and desirable for big market size A. In other words, when BAs sell close substitutes, there is no IS when BAs have a big market size A (big BAs will not share information but small BAs will). Conversely, when BAs sell weak substitutes, there is no IS when BAs have a small market size (big BAs will want to share information but small BAs will not).

Proof of Proposition 3.3: To recall, $\rho = \tau_\varepsilon / (\tau_\varepsilon + \tau_\theta)$ is the informativeness of a BA's private signal, or the quality of the new information received by the BA. A value of ρ close to 1 means that the signal adds significantly to the BA's understanding of the optimal animal. Inspecting the threshold γ_i for substitutes in (3.22) we notice that γ_i falls as ρ increases.

$$\frac{\partial \gamma_i}{\partial \rho} = -\frac{n(n-1)}{\left(\sqrt{n(\rho(n-1)+1)}+1\right)^2 \sqrt{n(\rho(n-1)+1)}} < 0.$$

We know that when $\gamma > \gamma_i$, BAs will not share information when the market size A is big. This means that when the signal is informative, big BAs will tend to decline IS for higher degrees of substitutability. On the other hand, when the signal is not very informative, big BAs will not mind sharing information when goods are strong substitutes.

Proof of Proposition 3.4: It can be easily shown that γ_i decreases as n increases

$$\frac{\partial \gamma_i}{\partial n} = -\frac{\rho(2n-1)+1}{\left(\sqrt{n(\rho(n-1)+1)}+1\right)^2 \sqrt{n(\rho(n-1)+1)}} < 0.$$

Proof of Proposition 4.1:

1. We assume that trait improvability T is small, making the difference in selection effects in (4.20) small and positive:

$$K_s^2 - K^2 = (K_s + K)(K_s - K)$$
$$= \frac{2T\tau_\varepsilon (2\tau_\theta + 4\tau_\varepsilon + \gamma\tau_\varepsilon)(K_s + K)}{(\tau_\theta + 3\tau_\varepsilon)(\tau_\theta + 2\tau_\varepsilon)(\tau_\theta + \tau_\varepsilon)(2 - \gamma)} > 0.$$

Setting (4.21) to zero yields a threshold

$$\mu(\bar{\gamma}) = \frac{\sqrt[3]{r} - (2-\gamma)\left(4\gamma - 28\gamma^2 - 9\gamma^3 + 32\right) + \frac{(2-\gamma)^2\left(81\gamma^6 + 312\gamma^5 + 352\gamma^4 + 40\gamma^3 - 96\gamma^2 - 32\gamma + 64\right)}{\sqrt[3]{r}}}{12(2-\gamma)^3(\gamma+1)}$$

$$\tag{A.6}$$

where

$$r = (384\gamma - 192\gamma^2 - 5632\gamma^3 - 9696\gamma^4 - 552\gamma^5 + 10\,684\gamma^6 + 10\,080\gamma^7 + 3780\gamma^8 + 513\gamma^9 - 512)\,(2 - \gamma)^3$$

$$+36\sqrt{-(2 - \gamma)^8\gamma^2(\gamma + 1)^5(\gamma + 2)^2\,(207\gamma^7 + 711\gamma^6 + 624\gamma^5 - 364\gamma^4 - 400\gamma^3 - 48\gamma^2 + 128\gamma - 64)}.$$

(A.7)

Then for any $\gamma > \bar{\gamma}$, the precision effect becomes negative. Since the selection effect is small, the profit difference is negative in (4.20) and so BA_1 is better off not joining the IS coalition. We can find intercepts by setting $\mu(\bar{\gamma}) = 0$, which yields solution $\gamma = \sqrt{3} - 1$ and setting $\gamma = 1$, which yields $\mu(1) = \frac{1}{24}\left(a + \frac{721}{a} + 1\right) \approx 2.13$, where $a = \sqrt[3]{8857 + 864i\sqrt{397}}$.

If we explore (4.21), we can show that the fourth power coefficient is positive for $0.37 \lessgtr \mu \lessgtr 2.22$ and negative for $\mu \gtrless 2.22$ or $\mu \lessgtr 0.37$. We can also show that the sign for the third power coefficient is positive for $\mu \gtrless 0.86$ and negative for $\mu \lessgtr 0.86$. By inspection, we also know that the second and first power coefficients are negative, and the constant is positive. Therefore, we end up with four cases of $P(\gamma)$ depending on the interval of prior precision relative to signal precision:

1	$\mu \gtrless 2.13$	$P(\gamma) = -a\gamma^4 + b\gamma^3 - c\gamma^2 - d\gamma + e$ (3 sign Δ)
		$P(-\gamma) = -a\gamma^4 - b\gamma^3 - c\gamma^2 + d\gamma + e$ (1 sign Δ)
2	$0.86 \lessgtr \mu \lessgtr 2.13$	$P(\gamma) = +a\gamma^4 + b\gamma^3 - c\gamma^2 - d\gamma + e$ (2 sign Δ)
		$P(-\gamma) = +a\gamma^4 - b\gamma^3 - c\gamma^2 + d\gamma + e$ (2 sign Δ)
3	$0.37 \lessgtr \mu \lessgtr 0.86$	$P(\gamma) = +a\gamma^4 - b\gamma^3 - c\gamma^2 - d\gamma + e$ (2 sign Δ)
		$P(-\gamma) = +a\gamma^4 + b\gamma^3 - c\gamma^2 + d\gamma + e$ (2 sign Δ)
4	$\mu \lessgtr 0.37$	$P(\gamma) = -a\gamma^4 - b\gamma^3 - c\gamma^2 - d\gamma + e$ (1 sign Δ)
		$P(-\gamma) = -a\gamma^4 + b\gamma^3 - c\gamma^2 + d\gamma + e$ (3 sign Δ)

Table A.1: Four cases of P(γ) depending on the interval of prior precision relative to signal precision

where a,b,c,d are 4th, 3rd, 2nd, and 1st power coefficients depending on the interval of relative precisions and e is the constant in (4.21), respectively. Let us explore each case. Using Descartes' rule of signs we find the number of positive, negative and complex roots for each case:

case 1: $\mu \gtrless 2.22$. By Descartes' rule, $P(\gamma)$ can have 3 or 1 positive real roots, 1 negative root and 2 or 0 complex roots. However, we also can show that $P(\gamma)$ is

always positive for any $\mu \gtrsim 2.13$. Therefore, $P(\gamma) > 0$ for all γ in the first case, and the upper bound in case 2 changes to 2.13.

case 2: $0.86 \lesssim \mu \lesssim 2.13$. Using Descartes' rule of signs we find that $P(\gamma)$ can have 2 or 0 positive real roots 2 or 0 negative real roots and 4,2, or 0 complex roots. The graph of $P(\gamma)$ opens upward since $P(\gamma)$ goes to positive infinity as γ goes to either positive or negative infinity. $P(\gamma)$ is positive for $\gamma = -1$ and 0. We also know that $P(\gamma) > 0$ for any $\gamma \lesssim 0.732$. Thus, we know there are either 0 or 2 negative roots, but both are less than $\gamma = -1$. $P(\gamma)$ is negative for $\gamma = 1$, therefore $P(\gamma)$ changes sign between $\gamma = 0$ and $\gamma = 1$.

case 3: $0.37 \lesssim \mu \lesssim 0.86$. By Descartes rule, there can be 2 or 0 positive real roots, 2 or 0 negative real roots, and 4,2 or 0 complex roots. Similarly to case 2, we know that $P(\gamma) > 0$ for any $\gamma \lesssim 0.732$. $P(\gamma)$ is negative for $\gamma = 1$, therefore $P(\gamma)$ changes sign between $\gamma = 0$ and $\gamma = 1$

case 4: $\mu \lesssim 0.37$. By Descartes rule there can be 1 positive real root, 3 or 1 negative real roots and 2 or 0 complex roots. $P(\gamma)$ is negative for $\gamma = 1$, therefore $P(\gamma)$ changes sign between $\gamma = 0$ and $\gamma = 1$.

To summarize, $P(\gamma)$ is always positive for any $\mu \gtrsim 2.13$ or any $\gamma < \sqrt{3} - 1$. Also, with regards to cases 2, 3 and 4, $P(\gamma)$ is negative when $\gamma = 1$, therefore $P(\gamma)$ changes sign when goods are substitutes, more precisely between $\gamma = \sqrt{3} - 1$ and $\gamma = 1$.

2. If we set $P(\gamma = 1) = 0$ we find expression $9\tau_\theta^3 - 8\tau_\varepsilon^3 + \tau_\theta\tau_\varepsilon^2 + 30\tau_\theta^2\tau_\varepsilon = 0$ which yields two (discarded) negative roots and one positive root such that $P(\gamma = 1) > 0$ for any $\mu \gtrsim 2$. We can then show that $P(\gamma) > 0$ for any $\mu \gtrsim 2$.

Proof of Proposition 4.2: $V(\gamma)$ can be expressed as a fourth-degree expression in γ:

$$
\begin{aligned}
V(\gamma) = & \left(-2\tau_\varepsilon^4\right)\gamma^4 + 4\tau_\varepsilon^2\left(\tau_\theta + \tau_\varepsilon\right)\left(\tau_\theta + 2\tau_\varepsilon\right)\gamma^3 \\
& + \left(-\left(\tau_\theta + \tau_\varepsilon\right)^2\left(\tau_\theta + 2\tau_\varepsilon\right)\left(\tau_\theta + 3\tau_\varepsilon\right)\right)\gamma^2 + \left(-4\left(\tau_\theta + \tau_\varepsilon\right)^3\left(\tau_\theta + 2\tau_\varepsilon\right)\right)\gamma \quad \text{(A.8)} \\
& + 4\left(\tau_\theta + \tau_\varepsilon\right)^3\left(\tau_\theta + 2\tau_\varepsilon\right)
\end{aligned}
$$

or

$$
V(\gamma) = -a\gamma^4 + b\gamma^3 - c\gamma^2 - d\gamma + e \quad (3\Delta)
$$
$$
V(-\gamma) = -a\gamma^4 - b\gamma^3 - c\gamma^2 + d\gamma + e \quad (1\Delta)
$$

where a, b, c, d, e are absolute values of coefficients in front of γ. By Descartes rule, we can have 3 or 1 positive roots, and there is exactly 1 negative root and 2 or 0 complex roots.

If we let $\gamma = 1$,

$$V(1) = -\tau_\theta \left(\tau_\theta^3 + 7\tau_\theta^2 \tau_\varepsilon + 13\tau_\theta \tau_\varepsilon^2 + 5\tau_\varepsilon^3 \right) < 0$$

and if we let $\gamma = -1$,

$$V(-1) = \tau_\theta \left(7\tau_\theta^3 + 33\tau_\theta^2 \tau_\varepsilon + 51\tau_\theta \tau_\varepsilon^2 + 27\tau_\varepsilon^3 \right) > 0. \tag{A.9}$$

Thus, as $\gamma \to \infty$, $V(\gamma) \to -\infty$ and as $\gamma \to -\infty$, $V(\gamma) \to -\infty$, so $V(\gamma)$ opens downwards. Taking the first derivative with respect to γ, we get

$$\begin{aligned}
\frac{\partial}{\partial \gamma} V(\gamma, \mu) &= \left(-8\mu^4\right) \gamma^3 + \left(24\mu^4 + 36\mu^3 + 12\mu^2\right) \gamma^2 \\
&\quad + \left(-12\mu^4 - 34\mu^3 - 34\mu^2 - 14\mu - 2\right) \gamma \\
&\quad - \left(8\mu^4 + 28\mu^3 + 36\mu^2 + 20\mu + 4\right),
\end{aligned} \tag{A.10}$$

where $\mu = \tau_\varepsilon / \tau_\theta$.

Graphing it implicitly yields Figure A.1 below. We see that $V(\gamma)$ decreases for all $\gamma > 0$, which suggests only one root, not three roots, on the interval $0 < \gamma < 1$. Thus, there is 1 positive, 1 negative, and 2 complex roots to $V(\gamma)$. This proves the proposition that there is only one threshold of γ on the interval $0 < \gamma < 1$ for which $V(\gamma)$ changes sign. We can show three instances of γ and the sign of the derivative of $V(\gamma)$ with respect to γ:

$$\frac{\partial}{\partial \gamma} V(\gamma = -1, \mu) = 2\left(3\mu + 1\right)\left(5\mu^2 + 6\mu^3 - 1\right) > 1 \tag{A.11}$$

$$\frac{\partial}{\partial \gamma} V(\gamma = 0, \mu) = -\left(4\tau_\theta^4 + 20\tau_\theta^3 \tau_\varepsilon + 36\tau_\theta^2 \tau_\varepsilon^2 + 28\tau_\theta \tau_\varepsilon^3 + 8\tau_\varepsilon^4\right) < 0 \tag{A.12}$$

$$\frac{\partial}{\partial \gamma} V(\gamma = 1, \mu) = -2\left(3\tau_\theta^4 + 2\tau_\varepsilon^4 + 29\tau_\theta^2 \tau_\varepsilon^2 + 13\tau_\theta \tau_\varepsilon^3 + 17\tau_\theta^3 \tau_\varepsilon\right) < 0. \tag{A.13}$$

Proof of Proposition 4.3: $Z(\gamma, \mu)$ can be rewritten as

$$\begin{aligned}
Z(\gamma, \mu) &= -2\left(\gamma + 1\right)\left(5\gamma - 1\right)\left(\gamma - 2\right)^2 \mu^3 - \left(2 - \gamma\right)\left(29\gamma + 28\gamma^2 - 2\gamma^3 - 10\right) \mu^2 \\
&\quad + 2\left(-16\gamma - 11\gamma^2 + 4\gamma^3 + 8\right) \mu - \left(4\gamma + 5\gamma^2 - 4\right).
\end{aligned} \tag{A.14}$$

Solving $Z(\gamma, \mu) = 0$ we find the unique root (see below) such that
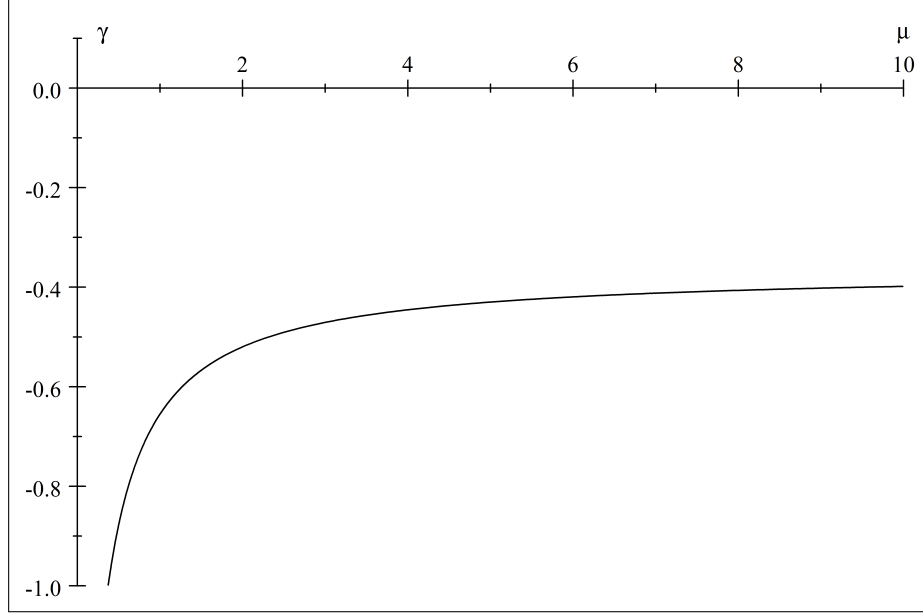
Figure A.1: The first derivative of $V(\gamma, \mu)$ changes sign at a threshold between $-1 < \gamma < 0$, such that the derivative is negative for any γ above the threshold. This suggests that $V(\gamma, \mu)$ is falling on the interval of interest, $0 < \gamma < 1$. Therefore, there can only be one root, not three, of $V(\gamma, \mu)$ on this interval. Thus, there is a unique threshold of γ on $0 < \gamma < 1$ for which $V(\gamma, \mu)$ changes sign.

1. (4.26) will be positive for $\gamma < 1/5$.

2. On the interval $\hat{\gamma} \in \left(\frac{1}{5}, \frac{2}{5}\left(\sqrt{6} - 1\right)\right)$, (4.26) will be positive if $\gamma < \hat{\gamma}$ and negative for $\gamma > \hat{\gamma}$.

3. (4.26) will be negative for $\gamma > 2/5\left(\sqrt{6} - 1\right)$.

To see this, solve $\lim_{\mu \to \infty} Z(\gamma, \mu)$ to get

$$\left((\gamma + 1)(\gamma - 2)^2 (5\gamma - 1)\right)(-\infty) - 5\gamma^2 - 4\gamma + 4 = 0$$

which is zero for $\gamma \approx 0.2$. In fact $\gamma = 0.2$ is the asymptote. Solving $\lim_{\mu \to 0} Z(\gamma, \mu)$ we find

$$-5\gamma^2 - 4\gamma + 4$$

with a (positive) solution $\frac{2}{5}\left(\sqrt{6} - 1\right) \approx 0.58$. This is the γ-intercept. The root in μ can also be found explicitly:

$$\bar{\mu} = \frac{1}{6\,(2-\gamma)} \frac{-29\gamma - 28\gamma^2 + 2\gamma^3 + 10}{(\gamma+1)\,(5\gamma-1)} - \frac{\left(i\sqrt{3}-1\right)}{12\,(\gamma+1)\,(\gamma-2)^2\,(5\gamma-1)} \sqrt[3]{k}$$
$$+ \frac{\left(i\sqrt{3}+1\right)\left(4\gamma^8 + 112\gamma^7 - 296\gamma^6 - 160\gamma^5 + 413\gamma^4 + 8\gamma^3 + 520\gamma^2 - 32\gamma + 16\right)}{12r\,(2-\gamma)^2\,(5\gamma^2 + 4\gamma - 1)}$$

(A.15)

where

$$k = 7328\gamma^3 - 7152\gamma^2 - 192\gamma + 13\,932\gamma^4 + 25\,668\gamma^5 - 23\,517\gamma^6 - 19\,728\gamma^7$$
$$+ 15\,012\gamma^8 + 2032\gamma^9 - 2586\gamma^{10} + 336\gamma^{11} + 8\gamma^{12}$$
$$+ 18\sqrt{-\gamma^2\,(\gamma+1)^5\,(\gamma-2)^6\,(5\gamma-1)^2\,(8\gamma^7 + 344\gamma^6 + 123\gamma^5 - 453\gamma^4 + 108\gamma^2 + 256\gamma + 64)} + 64.$$

(A.16)