Pro$: A New Profit-Based Genetic Selection Index in Canada

B.J. Van Doormaal[1], G.J. Kistemaker[1], L. Beavers[1] and P.G. Sullivan[1]
[1]Canadian Dairy Network (CDN), Guelph, ON, Canada, N1K 1E5

Abstract

Pro$ (pronounced Pro Dollars) was recently developed by Canadian Dairy Network (CDN) as a second national index that targets dairy producers who generate essentially all of their farm revenue from milk sales. Actual cow profitability data provided to producers by dairy herd improvement (DHI) agencies in Canada, namely CanWest DHI and Valacta, was used as the basis for deriving the new profit-based genetic selection index. Economic parameters used to calculate profitability for each cow are updated annually by economists to reflect changes in milk pricing as well as the associated expenses, including overhead, maintenance feed costs, marginal feed costs and quota opportunity costs. Data used was the accumulated profit to 6 years of age for 672,254 registered Holstein cows with known sire identification, born from January 2005 to September 2008. For cows not surviving to 6 years of age, accumulated profit to the date they left the herd was considered as lifetime profit. For each sire, the average accumulated profit of daughters to 6 years of age was computed. A total of 830 sires with at least 100 daughters with profit data were used to conduct the two-step multiple trait regression analysis to determine the contribution of sire EBVs for three production, four major type, and eight functional traits in predicting the average daughter profit to 6 years of age. Adjusted R-squared of the Pro$ prediction equation was .6167, which can be applied to any dairy breed with the appropriate scaling factors. Relative to LPI, selection for Pro$ in Holsteins has an stronger expected response for Milk and Protein Yields as well as various functional traits, including Herd Life, while both indexes have similar selection responses for Fat Yield, Daughter Fertility, Mastitis Resistance and Rump. Effective August 2015, Pro$ will be available in the Holstein and Jersey breeds and will be expressed in dollar terms as a deviation from breed average. For other dairy breeds, the research behind the development of Pro$ was used to modify the LPI formula effective August 2015 to better reflect expected average daughter profit from milk sales.

Key words: profit, genetic selection, genetic progress, expected response

Introduction

Genetic improvement of dairy cattle involves the evaluation of many traits that, for Canada, currently include measures of production, conformation, udder health, reproduction, calving performance, milkability and longevity with new traits on the horizon. For decades, genetic evaluation services have included the publication of a national genetic selection index that combines important traits into a single value for ranking animals in each breed. Since 1991, the Lifetime Profit Index (LPI) has been published as the official index for all dairy cattle breeds. Over the course of time, several factors changed, which led to the development of a second national genetic selection index in Canada, named Pro$ (Pro Dollars). The main driving forces behind the development of Pro$ included (a) a high proportion of dairy producers now recognize that genetic selection affects cow and herd profitability, (b) dairy producers want to understand genetics in economic terms, and (c) a single national selection index no longer meets the needs of all dairy producers in Canada even if they all have a selection objective of maximizing cow profitability.

Data and Methods

Defining Cow Profitability

Dairy Herd Improvement (DHI) organizations in Canada, namely CanWest DHI and Valacta, jointly provide their clients, on an annual basis, with Cow Profitability values, which are also summarized in a Herd Summary Profitability Report that include national benchmarks. Cow

Profitability is calculated to first calving (i.e.: rearing cost) as well as to second, third and fourth calving for each cow that reaches that stage in their productive life. For cows that reach a fifth or later calving, the accumulated profit is presented as a single lifetime profitability value.

The profit equation used to calculate the DHI Cow Profitability values focuses on revenue associated with milk sales only and the primary associated expenses. Table 1 lists each revenue and expense item included in the profit equation and provides the economic values used for the Holstein breed in 2014. These values are updated annually by DHI staff, in consultation with external economists, based on the milk payment, component pricing and supply management system in each province as well as actual cost of production figures derived from herds in the Agritel database in Québec (http://agritel.gcaq.ca).

For heifer rearing expense, a base cost associated with age at first calving of 730 d (24 mo) is used in addition to a deviation of \$3.55/d (\$1.53/d for overhead plus \$2.02/d for maintenance feed cost for a Holstein heifer) for age at first calving deviated from this base value.

Once an animal calves for the first time, it has the opportunity to start generating revenue which, in Canada, is based on a payment for each kg of fat, protein and other solids produced as well as a deduction for the shipment of the fluid portion of milk (Table 1).

For cow expenses, a basic maintenance cost (\$2.02/d for Holstein) is considered for both lactating and dry cows while the overhead costs are higher for lactating (\$6.10/d) compared to dry (\$1.53/d) cows. Marginal feed costs are included based on a cost allocated per kg of fat (\$1.93) and per kg of protein (\$2.67) produced by the cow. Given the supply management system, a quota opportunity cost is also considered on the expense side of the profit equation, which is calculated as a function of the kg of fat produced (\$2.43/kg).

Table 1. Economic values for calculation of cow profitability in Holstein.

Heifer rearing expenses:	
Base rearing cost to 730 d (\$)	2560
Heifer overhead cost (\$/d)	1.53
Heifer maintenance feed cost (\$/d)	2.02
Revenue from milk sales:	
Fat (\$/kg)	9.91
Protein (\$/kg)	9.76
Other solids content (% of kg milk)	5.53
Other solids (\$/kg)	1.90
Fluid deductions (\$/kg milk)	0.042
Cow expenses:	
Maintenance - Lactating and dry cows (\$/d)	2.02
Overhead - Lactating cows (\$/d)	6.10
Overhead - Dry cows (\$/d)	1.53
Marginal feed cost (\$/kg fat)	1.93
Marginal feed cost (\$/kg protein)	2.67
Quota opportunity cost (\$/kg fat)	2.43

Source: Canadian DHI, 2014

In an effort to define lifetime profitability in a manner that reflects a cow's ability to be profitable across a series of production and reproduction cycles, the accumulated Cow Profitability to fourth calving, as calculated by DHI, was initially considered. Examining the distribution of Holstein cows by age at fourth calving the range was found to be very wide with some calving prior to 5 years of age while other reaching that reproductive stage after 7 years of age. For this reason, a new definition of Cow Profitability was established, namely accumulated profit to 6 years (72 mo) of age. Selecting a fixed age gives a greater focus on each cow's reproductive performance as well as its production performance compared to defining profit to a fixed number of calvings. In general, by six years of age most cows will have had the opportunity to achieve four cycles of breeding and pregnancy followed by a calving and then a lactation, during which time they are also at risk for various other reasons for disposal.

Calculation of Profit to 6 Years

Lactation, disposal and pedigree records were extracted from the Canadian Dairy Network (CDN) database for 690,553 Holstein cows born from January 2005 to September 2008, allowing every animal the opportunity to reach 6 years of age. After editing for known sire identification and age at first calving between 18 and 42 mo, 672,254 cows were analyzed. For each cow, the following variables were retained and/or calculated:

- Birth date and age at first calving
- Age at disposal if removed from the herd prior to 6 years of age
- Total number of days in lactation from first calving to 6 years of age, or disposal
- Total number of days dry from first calving to 6 years of age, or disposal
- Total kg milk, fat and protein produced to 6 years of age, or disposal

Since cows may be in lactation when reaching 6 years of age, an interpolation procedure was used to estimate the kg milk, fat and protein produced in that lactation to the point of reaching 6 years of age, based on the cumulative lactation yields at the nearest test day prior to and following the days in milk (DIM) at 6 years of age. The interpolation methodology used the shape of the average lactation curve for Holsteins in Canada between the DIM at the nearest test day on each side of the DIM at 6 years of age. Once the total kg milk, fat and protein to 6 years of age were estimated, the accumulated profit to 6 years was calculated using the economic values in Table 1. For cows removed from the herd prior to reaching 6 years of age, the accumulated profit to the age of disposal was considered as the profit to 6 years.

Regression Analysis

By matching each cow's accumulated profit to 6 years of age to the sire identification, the average daughter profit to 6 years of age was calculated. Sires with at least 100 daughters with profit data were retained for analysis. Due to the likelihood of non-random usage across herds of sires that were first proven outside of Canada with subsequent semen importation, 174 sires that were first progeny proven in Canada after six years of age were excluded, leaving a total of 830 sires for final analysis.

A two-step multiple trait regression analysis was conducted. In the first step, the sire's official estimated breeding value (EBV or proof) for the following traits were used as input variables for predicting the average daughter profit to 6 years of age:

- Milk, Fat and Protein yields
- Mammary System, Feet & Legs, Dairy Strength and Rump as the four major scorecard type traits
- Somatic Cell Score
- Daughter Fertility
- Body Condition Score
- Milking Speed
- Milking Temperament
- Calving Ability
- Daughter Calving Ability

Fat and protein deviations are functions (i.e.: ratios) of the yield traits so were excluded as input variables. Similarly, since the major scorecard type traits are a function of the linear type traits associated with each scorecard section, including both the scorecard and linear traits is redundant. Since some linear type traits have an intermediate optimum for genetic selection, which would require both a linear and quadratic term for regression analysis, and given the producer attention to the major scorecard traits, the latter were included as the selected input variables to represent the contribution of type traits in predicting profit to 6 years of age. The sire's proof for overall conformation was excluded as an input variable due to its high correlation with the selected major scorecard type traits and the fact that cow final classification scores for overall conformation are determined based on a mathematical function of the cow's assessments for each of the four major scorecard traits. The seven functional traits included as input variables for the first step regression analysis include traits that have been evaluated for several years in Canada. A key trait excluded from the step 1 analysis was Herd Life, which is the genetic evaluation for longevity in Canada. Similar to previous arguments associated with the inclusion of highly correlated traits, Herd Life evaluations reflect how well a sire's daughters survive through to fourth calving, which is highly influenced by the ability of daughters to perform well for the various functional traits already included in the step 1 regression analysis. Excluding Herd Life forces the regression analysis to estimate the impact of each of the other functional traits directly on the sire's average daughter profit to 6 years of age.

The resulting regression equation from step 1, based on sire proofs for the 14 input traits, was applied to the group of 830 sires in the regression analysis to derive a predicted average daughter profit (Predicted Profit). For each sire the difference between the actual average daughter profit to 6 years minus Predicted Profit was calculated, referred to as Step 1 Residual Profit. The second step regression analysis used the sire's proof for only Herd Life or only Conformation as input variables to predict the Step 1 Residual Profit. This second analysis aimed to test the significance of including each of these two "compound" traits after already including the 14 traits included for the step 1 analysis. If significant, the resulting regression coefficient for each specific trait could simply be added to the 14-trait regression equation from step 1 to derive the final prediction equation used to calculate the new profit-based selection index, Pro$.

Expected Response

Over the course of time that LPI was used as the primary genetic selection index in Canada, the formula experienced regular updates, both by adding new traits as well as by modifying the relative weights applied to each trait. In general, outside of research conducted years ago for assessing the economic values for production traits, decisions regarding changes to the LPI formula were more intuitive than analytical. Prior to finalizing any new LPI formula, analysis focused on the realized rate of genetic progress per trait in the cow population during the previous 5- or 10-year period as well as the average genetic merit for each trait for the highest ranking proven sires for the proposed index.

The new approach of using regression analysis for deriving the Pro$ formula eliminates any discussion on traits to include and weights assigned to each trait. In addition, the concept of presenting the genetic response per trait, as expected from selection on the index, was favoured over focusing on relative weights on each trait included in the index. Software freely available for assessing multiple trait selection index response was used to calculate the expected response in standard units by trait for each standard unit gain for the index (van der Werf, 2014). As input, this software program requires the list of input traits, the unit of expression (i.e.: actual or standard units) and economic value (i.e.: relative weights or regression coefficients) by trait as well as the underlying correlation matrix among all traits included.

Results & Discussion

Table 2 presents averages for general statistics associated with the calculation of accumulated profit to 6 years for the selected Holsteins born from 2005 to 2008, based on 2014 economic values. The average age at first calving for this data was 26.7 mo, yielding an average rearing cost of $2,861. Overall, in this data, 28.6% of cows were still in the herd at 6 years of age (2,191 d) and, given this maximum for those cows, the average days of life for accumulating profit for all cows was 1,688 d. Total days in milk and dry were 753 and 115, respectively, resulting in average total milk, fat and protein yields of 22,869, 874 and 739 kg, respectively. In this group of cows, the average cow income and expense after first calving was $17,320 and $12,308, respectively. Subtracting the rearing cost and cow expense after first calving from the cow revenue yielded an average accumulated profit to 6 years of $2,151 for this group of cows.

Table 2. Average statistics for Holstein cows included for the calculation of accumulated profit to 6 years of age (N=672,254).

Age at first calving (d/mo)	815/26.7
Cows reaching 6 yrs of age (%)	28.6
Days for profit to 6 yrs (max. 2191 d)	1688
Total days in lactation	753
Total days dry	115
Total milk (kg)	22869
Total fat (kg)	874
Total protein (kg)	739
Rearing cost ($)	2861
Cow income ($)	17320
Cow expense ($)	12308
Accumulated profit to 6 yrs ($)	2151

Table 3 presents the adjusted R-squared resulting from the various regression analyses conducted using sire proofs for various traits as predictors of the average daughter profit to 6 years of age.

Table 3. Adjusted R-squared (R^2) for various regression analyses used to derive the Pro$ equation (N=830 sires).

Model	Adjusted R^2
Step 1 including 14 traits	0.5778
Step 2 including Herd Life to predict the "Step 1 Residual Profit"	0.0818
Step 2 including Conformation to predict the "Step 1 Residual Profit"	0.0003
Regression of Pro$ equation on average daughter profit to 6 yrs	0.6221
Single regression using 14 traits in Step 1 as well as Herd Life	0.6717

The initial 14-trait model for the step 1 regression analysis resulted in an adjusted R-squared of .5778. The model testing the added significance of Herd Life, based on its value in predicting the Step 1 Residual Profit, yielded an adjusted R-squared of .0818 (*P<.0001*), whereas the same test for Conformation was not significant with an adjusted R-squared of .0003 (*P=0.6340*). The final prediction equation for Pro$, which therefore combined the regression coefficients from the step 1 analysis and the regression including Herd Life to predict the Step 1 Residual Profit, resulted in a total adjusted R-squared of .6221. This accuracy of prediction was lower than the adjusted R-squared of .6717, when the 14 traits included in step 1 and Herd Life were used in a single regression to predict average daughter profit to 6 years. While this 7.4% reduction in accuracy of prediction may seem important, the application of the Pro$ equation to rank bulls, both old and young, as well as heifers and cows means that maximum adjusted R-squared for only the older proven bulls is not the only important criterion. When examining lists of top ranking animals, users expect they are not only superior for Herd Life but also for the various functional traits that significantly contribute to longevity, especially Somatic Cell Score and Daughter Fertility. The group of 830 progeny proven sires included in the regression analysis all have actual daughter survival data in their Herd Life proof, but the Pro$ equation will also be applied to young bulls whose Herd Life evaluation will be based primarily on an indirect prediction of Herd Life from other functional traits and various type traits already included in the step 1 prediction of profit to 6 years.

The Holstein Pro$ equation was derived from standardized sire proofs. Therefore, the Pro$ regression coefficients can be applied to other dairy breeds using standardised genetic evaluations. Since Pro$ values for sires are scaled such that each point difference between two bulls equates to $1 CAD difference in the expected average profit to 6 years of age per daughter, application of the Holstein Pro$ formula to other breeds requires the proper scaling parameters (i.e.: mean and variance) as well. For expression of Pro$ within each breed, a rolling cow base is used such that the average Pro$ is set to 0 for cows born during a 3-year period centred seven years back from the current year (i.e.: birth years 2007 to 2009 for 2015). For the August 2015 genetic evaluation release, Pro$ will be published for the Holstein and Jersey breeds. For the other breeds in Canada, the LPI formula was modified to maximize the correlation with Pro$ rather than introducing Pro$ as a second national genetic selection index in addition to LPI.

Table 4 compares the expected response by trait in Holsteins for Pro$ and the new LPI formula effective August 2015, which is renamed Lifetime Performance Index. Six traits, namely Fat Yield, Mastitis Resistance, Rump, Herd Life, Daughter Calving Ability and Daughter Fertility have essentially equal expected responses (i.e.: difference within ±.02) with the two national indexes. Differences in selection response in favour of Pro$ are highest for Milk Yield and Body Condition Score, but also favourable for Protein Yield, Somatic Cell Score, Milking Speed, Milking Temperament and Calving Ability. Traits that are expected to have more response with selection for LPI instead of Pro$ include Conformation, Feet & Legs, Dairy Strength and Mammary System as well as Fat and Protein Deviations.

Table 4. Genetic response by trait expected from selection for Pro$ compared to LPI in Holsteins (standard units).

Trait	Pro$	LPI*	Diff.
Milk Yield	0.60	0.51	0.09
Fat Yield	0.68	0.67	0.01
Protein Yield	0.73	0.68	0.05
Fat Deviation	0.12	0.20	-0.07
Protein Deviation	0.07	0.15	-0.07
Mastitis Resistance	0.31	0.32	-0.01
Somatic Cell Score	0.46	0.41	0.05
Herd Life	0.64	0.62	0.02
Milking Speed	0.03	-0.02	0.05
Milking Temperament	0.10	0.05	0.05
Calving Ability	0.22	0.19	0.03
Daughter Calving Ability	0.25	0.24	0.02
Daughter Fertility	0.11	0.13	-0.02
Body Condition Score	0.09	0.02	0.07
Conformation	0.56	0.66	-0.11
Mammary System	0.55	0.63	-0.08
Feet & Legs	0.48	0.59	-0.11
Dairy Strength	0.19	0.30	-0.11
Rump	0.23	0.22	0.01

** New formula for LPI effective August 2015 and renamed to Lifetime Performance Index.*

Table 5 presents the same comparison as Table 4 but for the Jersey breed. Even with identical regression coefficients used for Pro$ in Holstein and Jersey breeds, expected response by trait varies between breeds due to differences in the underlying correlation structure among traits. In addition, the LPI formula for Holstein and Jersey breeds are substantially different to coincide with the different breed goals. For Jerseys, traits with little difference in selection response (i.e.: less than ±.02) from the two indexes include Milk Yield, Calving Ability, Conformation, Mammary System, Feet & Legs and Rump. Differences in selection response in favour of Pro$ are largest for Herd Life and Daughter Fertility but also significant for other functional traits, namely Somatic Cell Score, Mastitis Resistance, Daughter Calving Ability, Body Condition Score, Milking Speed and Milking Temperament. Expected response for LPI is significantly higher than Pro$ for Protein and Fat, both in terms of deviations and yields, as well as Dairy Strength.

Table 5. Genetic response by trait expected from selection for Pro$ compared to LPI in Jerseys (standard units).

Trait	Pro$	LPI*	Diff.
Milk Yield	0.53	0.54	-0.01
Fat Yield	0.66	0.78	-0.11
Protein Yield	0.65	0.74	-0.09
Fat Deviation	0.13	0.26	-0.13
Protein Deviation	0.15	0.33	-0.18
Mastitis Resistance	0.18	0.09	0.09
Somatic Cell Score	0.52	0.41	0.11
Herd Life	0.27	0.04	0.23
Milking Speed	0.15	0.07	0.07
Milking Temperament	0.14	0.09	0.05
Calving Ability	-0.22	-0.21	-0.01
Daughter Calving Ability	0.24	0.14	0.10
Daughter Fertility	0.24	0.03	0.22
Body Condition Score	-0.29	-0.38	0.09
Conformation	0.32	0.33	-0.01
Mammary System	0.38	0.38	0.00
Feet & Legs	0.12	0.11	0.01
Dairy Strength	0.14	0.24	-0.10
Rump	0.15	0.14	0.01

** New formula for LPI effective August 2015 and renamed Lifetime Performance Index.*

Ranking by Pro$ versus LPI

Pro$ and the new LPI effective August 2015 have a correlation among progeny proven sires of .957 in Holstein and .925 in Jersey. Such correlations are generally considered very high but a significant re-ranking among the top animals may still arise, which is observable to users of genetic evaluations. Table 6 shows the percentage of bulls, either progeny proven or genomic young bulls, that are in common for the top LPI lists in Holstein and Jersey and the top 15, 50, 100 and 250 Pro$ bulls. Official evaluations for April 2015 were used to calculate Pro$ as well as LPI values based on the formula in each breed effective August 2015. Among the high ranking Pro$ proven sires in each breed, there is a significant overlap with LPI, having two-thirds and 60% in common for the top 15 Pro$ in Holstein and Jersey, respectively. The percentage in common grows higher faster in Jersey compared to Holstein (i.e.: 94.8% vs 78.8%) as the count of top Pro$ bulls includes 250, primarily due to the relatively limited number of Jersey proven sires in Canada (N=541) compared to Holsteins (N=9,531).

Table 6. Percentage of high ranking bulls (Proven vs Young) in common for Pro$ and LPI in Holstein and Jersey breeds.

Top "N" for Pro$	% in common with LPI*			
	Holstein		Jersey	
	Proven	Young	Proven	Young
15	66.7%	33.3%	60.0%	26.7%
50	66.0%	60.0%	76.0%	46.0%
100	77.0%	50.0%	83.0%	54.0%
250	78.8%	60.8%	94.8%	72.0%

** New LPI formula effective August 2015 applied to official evaluations of April 2015.*

For genomic young bulls, the degree of re-ranking between Pro$ and LPI is more significant than for proven sires, which is expected due to the tighter range at the high end for each index. Since genomic young sires currently represent approximately 60% of the market share in terms of semen sales in Canada, the difference in ranking for Pro$ versus LPI is expected to lead to a more genetically diversified group of popular genomic young bulls used, which may help to control the rate of increase in the average inbreeding levels of the female population in each breed.

Conclusions

Actual cow profitability data, based on economic values used by DHI agencies to provide herd management information to clients across Canada, was used to develop Pro$ as a new profit-based genetic selection index. Regression analysis of Holstein proofs involving 15 traits for 830 sires to predict the average daughter profit accumulated to 6 years of age was conducted, yielding an adjusted R-squared of .6221. Using regression analysis to derive the contribution of each trait (i.e.: regression coefficient) in the Pro$ equation accounts for correlations among traits. The derived Pro$ equation based on Holstein data can be applied to other breeds with the appropriate scaling factors to maintain the scale of expression such that each point difference in Pro$ values between sires equates to an expected difference of $1 CAD in accumulated profit to 6 years of age per daughter. Expected response by trait resulting from selection for either Pro$ or LPI varies depending on the underlying correlation structure among traits in each breed.

Effective August 2015, Pro$ will be introduced as a second national genetic selection index in the Holstein and Jersey breeds while the other breeds opted to modify the LPI formula, effective August 2015, to reflect the research results associated with Pro$. Even though correlations between Pro$ and LPI are high (i.e.: .957 in Holstein and .925 in Jersey), re-ranking among the list of top proven sires and genomic young bulls is significant to the industry and is expected to help reduce the rate of increase in inbreeding levels in the Holstein and Jersey populations in Canada.

Acknowledgements

Appreciation is expressed to Canadian DHI partners, CanWest DHI and Valacta, for providing details regarding the calculation of Cow Profitability values used by their clients for herd management purposes. Special thanks to Dr. Robert Moore of Valacta and Canadian DHI for his time and assistance.

References

van der Werf, J. 2014. Software program entitled Multiple Trait Selection Index - 20 Trait Version. Accessed September 2014. *http://www.personal.une.edu.au/~jvanderw/software.htm*.