Research in Veterinary Science 152 (2022) 333–353

Contents lists available at ScienceDirect

Research in Veterinary Science

journal homepage: www.elsevier.com/locate/rvsc

Udder health, veterinary costs, and antibiotic usage in free stall compared with tie stall dairy housing systems: An optimized matching approach in Switzerland

Armin van Aken[a,b,*], Daniel Hoop[a], Katharina Friedli[c], Stefan Mann[a]

[a] Agroscope, Tänikon 1, 8356 Ettenhausen, Switzerland
[b] University of Geneva, Geneva School of Social Sciences, Boulevard du Pont d' Arve 40, 1211 Geneva, Switzerland
[c] Federal Food Safety and Veterinary Office FSVO, Centre for Proper Housing of Ruminants and Pigs, Agroscope Tänikon, 8356 Ettenhausen, Switzerland

ARTICLE INFO

Keywords:
Dairy
Mastitis
Veterinary cost
Antibiotic usage
Housing system
Matching

ABSTRACT

Observational studies are important in livestock science. As treatment is not assigned randomly in such studies, selection bias can be a problem. This is often addressed by matching methods. However, if treatment and control groups differ considerably in their characteristics, it might be necessary to additionally prune observations that lack overlap in the opposite group. "Matching Frontier" method was developed because pruning observations manually often results in suboptimal solutions. The feasibility of the approach for animal health and welfare issues was tested in an observational study evaluating the effect of free stall housing and increased lying comfort on udder health, veterinary costs, and antibiotic usage in Swiss dairy farming.

Data were collected in a survey with 1835 Swiss dairy farmers (response rate 28.3%). The treatment group ($n = 179$) comprised farmers participating in a voluntary animal welfare program that, in addition to free stall housing, required increased lying comfort. Farmers in the control group ($n = 229$) kept their cows in tie stalls.

Using the Matching Frontier method, treated units were matched to control units based on five confounders. Subsequently, observations were pruned to achieve sufficient balance and overlap between the two groups. The effect of the program on the eight outcome variables was finally estimated using linear regression.

Farmers in the treatment group had a lower incidence of clinical mastitis (−3.66 per 100 cow-years, −25%, $p < 0.05$), a lower incidence of culled cows due to udder health problems (−1.61 per 100 cow-years, −30%, $p < 0.05$), fewer veterinary costs (−42.44 per cow-year, −22%, $p < 0.05$), a lower incidence of total intramammary antibiotic treatments (−15.88 per 100 cow-years, −23%, $p < 0.01$), a lower incidence of intramammary antibiotic treatments for mastitis therapy (−7.83 per 100 cow-years, −32%, $p < 0.01$), and a lower incidence of intramammary antibiotic treatments for dry-cow therapy (−8.80 per 100 cow-years, −21%, $p < 0.05$). No differences were found for the average somatic cell count and the number of cows with a cell count above 150.000.

The results suggest that free stall housing, in combination with increased lying comfort, can have a positive effect on udder health, animal welfare, and the economic situation of the farm. Additionally, fewer antibiotic treatments can be beneficial to public health.

The Matching Frontier method has proven to be a helpful tool that may also have added value for future observational studies in livestock science.

1. Introduction

Public concerns about animal welfare and specific production practices in livestock agriculture has increased over the last years (European Commission. Directorate-General for Health and Food Safety, 2016; Ly et al., 2021; Wolf et al., 2016). This also applies to the practice of keeping dairy herds in tie stalls, which are less accepted than free stalls (Robbins et al., 2019; Waldrop and Roosen, 2021). In align with these public values, the Swiss government promotes free stall housing through a voluntary direct payment program. This raises the question of whether the increase in animal welfare can also be shown based on measurable animal health indicators.

* Corresponding author at: Agroscope, Tänikon 1, 8356 Ettenhausen, Switzerland
E-mail address: armin.vanaken@agroscope.admin.ch (A. van Aken).

https://doi.org/10.1016/j.rvsc.2022.08.021
Received 25 June 2021; Received in revised form 13 August 2022; Accepted 24 August 2022
Available online 30 August 2022

In previous studies in Switzerland, positive effects of the voluntary free stall program on lameness, alterations of the skin around the hocks, callosities at the carpal joints, teat injuries, the incidence of medical treatments, and veterinary costs have been found (Odermatt et al., 2018; Regula et al., 2004). So far, one aspect that has limitedly been considered in program evaluations is udder health. However, udder health plays a major role in dairy farming, as udder health problems can lead to great suffering of cows, cause high costs for farmers and are the main reason for antibiotic usage (Halasa et al., 2007; Heikkilä et al., 2012; Menéndez González et al., 2010; Schaeren, 2006). Antibiotic usage in livestock is increasingly being viewed critically because it can provoke antibiotic resistance. This is a threat to public health at the global level (Talebi Bezmin Abadi et al., 2019).

Therefore, the present study has the aim to evaluate the effect of free stall housing on udder health, veterinary costs, and antibiotic usage in Swiss dairy farming.

2. Method

2.1. Voluntary free stall housing program

Switzerland promotes the system of free stalls through a voluntary program called "Particularly animal friendly stabling" (PAS). Farmers who participate have to keep cows in a free stall housing system as opposed to tie stalls. Also, the program requires a lying area separated from feeding area. This has to be bedded with deep straw or must comprise cubicles with an equivalent soft mattress. In the feeding area, solid flooring is mandatory. Farmers received a yearly amount of 90 CHF per livestock unit if they participated in the program (In 2018, one Swiss franc [CHF] corresponded € 0.87 and US $1.02; https://data.snb.ch). In 2018, 31.4% of Swiss dairy farmers participated in the program (FOAG, 2020). The majority of dairy farmers who kept their cows in free stalls participated in the program. Hence, in the analysis, only free stall farms that participated in the program were compared to farms with tie stalls.

2.2. Study design and data collection

In the analysis, a dataset was used generated through a survey in 2019. This survey was not conducted specifically to answer the present research question, but it was part of a larger project aiming at the evaluation of farmers' willingness to participate in voluntary antibiotic reduction programs (Swiss National Science Foundation, 2021). However, data on udder health, antibiotic use and veterinary costs was also collected in the survey, so the data set was appropriate to address the present research question.

In the survey of the larger project, 2250 livestock farms from all over Switzerland were contacted. Of these, 1835 were dairy farms and 415 were farms with other animal categories than dairy (Fig. 1). Both the total sample size of 2250 and the division into 1835 dairy farms and 415 farms with animal categories other than dairy was determined based on the research questions of the larger project and were not directly connected to the research question of the present paper. The 1835 dairy farms were randomly drawn from the population of all dairy farms in Switzerland, which comprises 25,007 farms. (FOAG, 2020). The 415 farms with animal categories other than dairy do not play a role in the further progress of this paper.

2.3. Questionnaire

The questionnaire used in this study included 43 questions on 12 DINA4 pages. It was initially developed in German and translated into French and Italian for the non-German-speaking regions with the help of a professional translation service provider. The questionnaire was pretested with five veterinarians and 15 farmers. Seventeen variables collected in the survey were relevant to this study. These are presented in Table 1. Farmers were asked to base all their responses in the survey exclusively on the year 2018.

Outcome variables collected through the survey include the incidence of clinical mastitis ("Mastitis"), the incidence of culled cows due to udder health problems ("Culled cows"), average herd somatic cell count over the year 2018 ("HSCC"), the number of cows with a somatic cell count level above 150.000—at least once during 2018 ("High cell"), cumulative veterinary costs ("Veterinary costs"), the incidence of intramammary antibiotic treatments for mastitis therapy ("Antibiotic mastitis"), and the incidence of antibiotic treatments for dry-cow therapy ("Antibiotic dry-cow").

The treatment variable of interest is whether farmers participated in the voluntary free stall program in 2018 ("Free stall program"). Furthermore, farm and farmer characteristics were collected in the survey. As it is possible to keep cows in free stalls without participating in the voluntary free stall program, farmers were additionally asked about the housing system ("Housing system"). Further questions related to the number of cows on farm ("Number of cows"), the average level of milk production per cow and year ("Milk production"), and whether the farm was conventional or organic ("Organic farming"). In Switzerland, different agricultural zones (plain zone, hilly zone, and mountain zone) were distinguished according to altitude and geographical conditions ("Agricultural zone"). Besides the free stall program, there are two other voluntary programs in Switzerland. One promotes regular outdoor access to livestock ("Outdoor program"). Farms participating in this program must give their animals access to pasture or an outdoor run for a minimum period per year. The other program promotes grassland-based production of milk and meat ("Grassland program"). For participating farms in the plain zone and hilly zone, 75% of the feed ration must be roughage, while 85% must be roughage for farms in the mountain zone. Program participation is therefore a feasible indicator of the feeding regime on the farm. Finally, farmers were asked about their level of agricultural education ("Agricultural education").

2.4. Data processing

Data collected through the survey was processed for further use. In all outcome variables, farmers had the option to check a "don't know" option. Because these responses were not considered in the statistical analysis, they were replaced by missing values.

The farm was the unit of the analysis. The within-herd incidences of clinical mastitis, culled cows, and intramammary antibiotic treatments were calculated as the number of reported cases or treatments reported by the farmer in 2018 divided by the average number of cows per farm throughout 2018 and expressed per 100 cows and year. Subsequently, for every farm, treatment incidence for mastitis therapy was added to the incidence of treatments for dry-cow therapy to obtain the incidence of total intramammary antibiotic treatments ("Antibiotic total").

For veterinary costs, the costs of insemination charged using the veterinarian were first subtracted from total veterinary costs and divided by the average number of cows per farm.

Finally, implausible values (entry errors) and values that deviated very strongly from the other values (outliers) were removed from the dataset.

2.5. Matching frontier method and statistical analysis

The decision to keep dairy cows in either a free stall or a tie stall is not random. It is influenced by the characteristics of the farmer and the farm. Estimated treatment effects could therefore be biased due to observed and unobserved confounding factors (Gelman et al., 2020). This is a common problem in observational studies. Remedies include methods that assume conditional independence (Imbens and Wooldridge, 2009). Under this assumption, unbiased estimates can be obtained after controlling for observed characteristics (Kreif et al., 2016). This can be done using parametric regression, but in situations with unbalanced groups, the estimated treatment effects can be highly

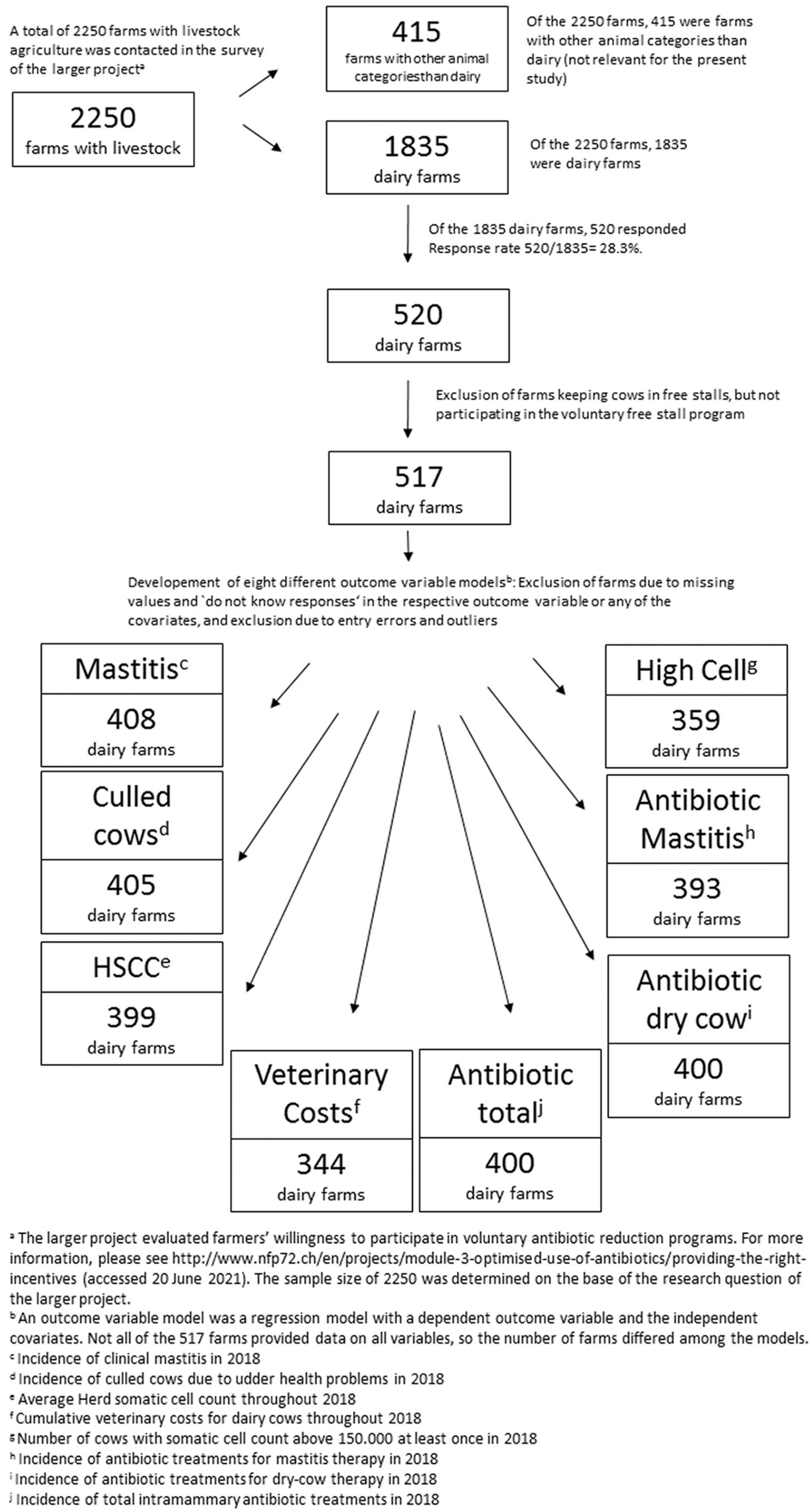

Fig. 1. The path from the 2250 farms contacted in the survey to the final samples of the outcome variable models.

Table 1
Variables collected in the survey.

Variable	Description
Mastitis	Incidence of clinical mastitis in 2018
Culled cows	Incidence of culled cows due to udder health problems in 2018
HSCC	Average Herd somatic cell count throughout 2018
High cell	Number of cows with somatic cell count above 150.000 at least once in 2018
Veterinary costs	Cumulative veterinary costs for dairy cows throughout 2018
Insemination costs	Amount of total veterinary costs spent on insemination
Antibiotic mastitis	Incidence of intramammary antibiotic treatments for mastitis therapy in 2018
Antibiotic dry-cow	Incidence of antibiotic treatments for dry-cow therapy in 2018
Free stall program	Dummy for whether farm took part in the voluntary free stall housing program "Particularly animal friendly stabling" (PAS) in the year 2018
Housing system	Dummy for whether farms keep their cows in a free stall housing system
Number of cows	Average number of cows on farm throughout 2018
Milk production	Average level of yearly milk yield per cow
Organic farming	Dummy for whether farm was organic
Outdoor program	Dummy for whether farm took part in the voluntary animal welfare program "Regular outdoor access of livestock" (ROEL) in 2018
Grassland program	Dummy for whether farm took part in the voluntary "Grassland-based milk and meat program" (GMF) in 2018
Agricultural zone	Agricultural zone in which farm is placed (plain zone, hill zone, and mountain zone)
Agricultural education	Farmer's level of agricultural education

sensitive to the choice of the parametric model specification (Ho et al., 2007). To tackle this problem, data can be pre-processed with the aim of gaining more balance and overlap between the two groups. This reduces model dependence and bias in subsequent regression (Ho et al., 2007). For this purpose, a matching method is suitable. In doing so, a treated observation is matched with one or more observations from the control group that are similar in characteristics. However, if there are observations in one of the groups that have no counterpart in the other group, balance can only be considerably achieved. A solution to this is the pruning of these observations from the dataset at the cost of reducing sample size. Therefore, a trade-off must be made between keeping observations while accepting some imbalance and pruning observations while achieving higher balance. In recent decades, several matching methods have been developed that have addressed this issue differently. These methods either try to reduce imbalance for a given sample size (e.g., Mahalinobis or Propensity score matching), fix balance and maximize sample size (e.g., Coarsened exact matching), or are arbitrary compromises between the two (such as calipers with ad hoc thresholds applied to other methods) (King et al., 2017). Since none of these methods optimize for both, researchers must settle for suboptimal solutions or manually optimize by iteratively tweaking their matching method and rechecking imbalance (King et al., 2017). The second approach is time-intensive. Moreover, it does not guarantee that an optimal solution is found.

To overcome the weaknesses of the manual approach, King et al. (2017) developed a procedure classified as a machine-learning approach (Sizemore and Alkurdi, 2019). Using this procedure, researchers can define, estimate, and visualize the Matching Frontier. The Matching Frontier comprises a set of matched subsamples that characterizes the trade-off between imbalance and sample. For each possible sample size, an algorithm is used to determine the matching solution with the lowest imbalance. All datasets on this frontier are optimal, meaning that there exists no matching solution with a lower imbalance for a given sample size or a higher sample size for a fixed imbalance. From this frontier, researchers can select one or more data subsets that reflect their preferences regarding imbalance and sample size for subsequent analysis (King et al., 2017).

To estimate the frontier in this study, the R package "MatchingFrontier" was used (King et al., 2015). The algorithm is based on a nearest-neighbor-approach. In a first step, every observation from the treated group is matched to the nearest observation in the control group, where the distance between observations is measured regarding the Mahalanobis distance. A matching with replacement approach was chosen, and the ratio of treated to control units was allowed to vary throughout the matched dataset. This is reflected in the assignment of weights[1] to the control units, which must be accounted for in the subsequent analysis steps. To obtain a sense of how much balance is achieved by this initial matching, the average Mahalanobis imbalance (AMI) metric is calculated. That is the Mahalanobis distance between each unit in the treatment group and the closest unit in the control group averaged over all units (King et al., 2017). In the next step of the algorithm, the unit or units with the largest difference from its matched partner gets pruned. To see how much balance is gained by pruning, the AMI is recalculated. This process is repeated until no imbalance between the treatment and control groups remains and the AMI equals zero. The various subsamples created by this process constitute the Matching Frontier. The Matching Frontier can be visualized by plotting the AMI of these subsamples as a function of the remaining sample size (or the number of observations pruned). After constructing and visualizing the frontier, researchers can select one or more datasets from the frontier that reflect their preference regarding the remaining imbalance and sample size. This selection can be made, for example, on the basis of AMI. However, in the literature, the standardized mean difference in the confounders between the treatment and control groups is more often used to measure the imbalance (Stuart, 2010). A maximum level of 0.1 is usually given as an acceptable cutoff value (Stuart, 2010), which we adopted in this study.

Based on the selected dataset, the treatment effect got estimated. We were interested in the average treatment effect on the treated (ATT). However, as it was not possible to achieve sufficient balance without pruning observations from the treatment group, the remaining observations were not representative of all treated units in the population. Therefore, the estimand was the average treatment effect in the remaining matched sample (ATM) (Greifer, 2021a) or sometimes called the feasible sample average treatment effect on the treated (FSATT) (King et al., 2017). The effect was estimated through a linear regression in which, in addition to the treatment variable, the confounders were included as control variables. This follows the common recommendation to run a regression after matching to account for the remaining imbalance (Abadie and Imbens, 2006; Rubin, 1973; Rubin and Thomas, 2000). The regression also included the weights that came from matching with replacement.

Estimated standard errors for the treatment effect have to consider the uncertainty present in the analysis through matching and estimation of the treatment effect (Greifer, 2021b). Overall, there are no analytic solutions to these issues; most studies have been done on uncertainty estimation after matching has relied on simulation studies. These have shown that robust standard errors and, depending on the matching method, cluster-robust standard errors were valid after matching (Austin, 2013, 2009; Austin and Small, 2014; Gayat et al., 2012; Wan, 2019) and tend to be conservative in situations with continuous outcomes (Hill and Reiter, 2006). The treatment effect and robust standard errors were

[1] When the ratio of treated to control unit is not allowed to vary throughout the matched data set (fixed-ratio matching), the sample average treatment effect can be estimated by a simple difference in means between the treated and control group (King et al., 2017). However, in variable-ratio matching, the approach we relied on in our analysis, treatment effects must first be estimated within each matched stratum by a simple difference in means. Aggregating up to the sample average treatment effect requires weighting, with the stratum-level treatment effect weighted according to the number of treated units (King et al., 2017). For further explanation of these weights and their calculation please see (King et al., 2017) and j.mp/CEM

estimated using the R packages “lmtest” (Zeileis and Hothorn, 2002) and “sandwich” (Zeileis, 2004).

The described steps with determination of the Matching Frontier, selection of a dataset, and calculation of the treatment effects were performed separately for each outcome variable. Overall, eight different models were calculated. Matching of these eight models was performed in the first step on the basis of the five confounders “Number of cows”, “Milk production”, “Organic farming”, “Outdoor program”, and “Grassland program”. These five confounders were, in addition to the treatment variable “Free stall program”, also included as independent variables in the linear regression. Subsequently, this solution was compared with a matching based on all seven covariates (including “Agricultural zone” and “Agricultural education”).

3. Results

3.1. Descriptive results

Fig. 1 shows the path from the initially 1835 dairy farms contacted in the survey to the final outcome variable models used for the analysis. Of the 1835 dairy farms, 520 responded. This corresponds to a return rate of 28.3% (= 520/1835). From these, three farmers reported keeping cows in free stalls but not participating in the voluntary free stall program. Since these fit neither into the treatment nor into the control group, these were dropped from the dataset.

The remaining 517 farms were then used as a starting point to develop the 8 linear regression models with different outcomes. These linear regression models each consist of a dependent outcome variable and the independent covariates. In each model, only those farms were included that had no missing values or do not know answers, outliers and entry errors in the specific outcome variable and all covariates. Thus, the number of farms differed slightly in the different models (Fig. 1). The total number of farms ranged from 359 to 408, depending on the outcome model. The number of farms with free stalls ranged from 152 to 179, and for tie stalls from 188 to 229 (Table 2). Table 2 provides the summary statistics for the outcome variables. Histograms for all outcome variables can be found in Appendix A.

For covariates, Tables 3–5 provide summary statistics. Numbers are presented as an example for the model where “Incidence of clinical mastitis” was the outcome. Numbers for the other 7 outcome models can be found in Appendix B, Appendix C, and Appendix D. As can be seen, the numbers for the covariates in the different models differ only minimally. Covariates are represented separately for the treatment and control group. *t*-tests were conducted to check whether the differences in the group was statistically different. Treatment farms had almost twice as many cows. Additionally, treatment farms had a significantly higher milk yield per cow and year. The treatment group had a higher proportion of farms located in the plain zone and a higher share of farms taking part in the outdoor exercise program. Farmers with tie stalls tended to participate more frequently in the grassland program. On the average, farmers with free stalls were higher educated.

Table 2
Summary statistics of outcome variables.

Variable	Unit	Treatment group	N	Mean	SD	Median	Min	Max
Mastitis[a]	per 100 cow-years	Free stall	179	11.00	7.54	9.43	0	46.67
		Tie stall	229	14.53	10.59	13.33	0	66.67
		Total[i]	408	12.98	9.53	11.54	0	66.67
Culled cows[b]	per 100 cow-years	Free stall	178	3.49	3.87	2.63	0	20.00
		Tie stall	227	5.03	6.41	3.85	0	38.46
		Total[i]	405	4.35	5.49	2.86	0	38.46
HSCC[c]	1000 cells/ml	Free stall	179	108.65	38.03	100.00	24	240.00
		Tie stall	220	90.97	39.06	82.50	20	300.00
		Total[i]	399	98.90	39.54	90.00	20	300.00
High cell[d]	per 100 cow-years	Free stall	152	33.67	21.15	28.57	0	100.00
		Tie stall	207	30.91	18.42	28.57	0	100.00
		Total[i]	359	32.08	19.64	28.57	0	100.00
Veterinary costs[e]	Swiss francs (CHF) per cow-year	Free stall	156	150.67	101.38	133.05	0	625.00
		Tie stall	188	195.70	133.67	190.08	0	733.33
		Total[i]	344	175.28	122.03	151.00	0	733.33
Antibiotic mastitis[f]	per 100 cow-years	Free stall	171	17.14	16.00	12.50	0	102.50
		Tie stall	222	22.53	23.21	20.00	0	238.10
		Total[i]	393	20.19	20.54	16.67	0	238.10
Antibiotic dry-cow[g]	per 100 cow-years	Free stall	177	34.98	31.11	27.50	0	100.00
		Tie stall	223	37.07	32.53	29.41	0	100.00
		Total[i]	400	36.15	31.88	28.57	0	100.00
Antibiotic total[h]	per 100 cow-years	Free stall	166	53.00	37.83	48.58	0	170.00
		Tie stall	215	58.82	39.87	54.17	0	190.91
		Total[i]	381	56.29	39.05	51.67	0	190.91

[a] Incidence of clinical mastitis in 2018.
[b] Incidence of culled cows due to udder health problems in 2018.
[c] Average Herd somatic cell count throughout 2018.
[d] Number of cows with somatic cell count above 150.000 at least once in 2018.
[e] Cumulative veterinary costs for dairy cows throughout 2018.
[f] Incidence of intramammary antibiotic treatments for mastitis therapy in 2018.
[g] Incidence of antibiotic treatments for dry-cow therapy in 2018.
[h] Incidence of total intramammary antibiotic treatments in 2018.
[i] “Total” is the sum of “Free stall” and “Tie stall” farms

Table 3
Summary statistics of continues covariate from treatment group ("Free stall") and control group ("Tie stall") and result of t-test to check for statistical significance of mean difference. Numbers are shown for the model where "Incidence of clinical mastitis" was the outcome variable. Numbers for the other 7 outcome models can be found in Appendix B.

Variable	Treatment group	Number of farms	Mean	SD	Median	Min	Max	t-value (*p*-value)
Number of cows	Free stall	179	37.06	18.57	35	4	92	−12.42 (0.000)
	Tie stall	229	18.64	7.88	18	3	45	

Table 4
Summary statistics of dummy covariates from treatment group ("Free stall") and control group ("Tie stall") and result of *t*-test to check for statistical significance of mean difference. The percentage reflects the proportion of organic farms and the proportion of participants in the outdoor and grassland program in the respective group. Numbers are shown for the model where "Incidence of clinical mastitis" was the outcome variable. Numbers for the other 7 outcome models can be found in Appendix C.

Variable	Treatment group	Number of farms	%	t-value (*p*-value)
Organic farming	Free stall	179	18.2%	−1.47 (0.142)
	Tie stall	229	12.9%	
Outdoor program[a]	Free stall	179	93.1%	−2.95 (0.003)
	Tie stall	229	83.5%	
Grassland program[b]	Free stall	179	66.5%	2.80 (0.005)
	Tie stall	229	79.9%	

[a] Farms participating in the outdoor program must give their animals access to pasture or an outdoor run for a minimum period per year. For more detailed information please see Odermatt et al. (2018).

[b] For participating farms in the plain zone and hilly zone, 75% of the feed ration must be roughage, while 85% must be roughage for farms in the mountain zone.

3.2. Matching frontier and treatment effects

Intermediate steps of the analysis are shown by the example of the "Antibiotic dry-cow" model. For each of the other seven models, only treatment effects are represented. Intermediate results of these were almost identical, since the same confounders were used in all models. The only minimal differences were due to different structures in missing values.

Results are first represented for the matching solution based on five confounders. Afterward, they were compared to the seven confounder cases.

In the first step of the analysis, treatment units were matched to control units. This already reduced the imbalance between the treatment and control group (Fig. 2). Before matching, the absolute mean difference in the two dummy variables "Organic farming" and "Outdoor program" had already been low (0.05 and 0.10, respectively). After matching, they were perfectly balanced. The absolute mean difference in the variable "Grassland program" decreased from 0.13 to 0.03. The imbalance in "Number of cows" regarding standardized mean difference was clearly reduced from 1.01 to 0.58. However, the value was still well above the cut-off value of 0.1. This also applies to the variable "Milk production" (from 0.57 to 0.24). To achieve further balance between the two groups, observations were pruned. Fig. 3 illustrates the corresponding Matching Frontier. Before observations were pruned, the AMI was 0.8. By removing a few observations, much could be gained regarding balance. If 50 observations were removed, the AMI was reduced to 0.2. From 150 pruned observations, the imbalance was close to zero. The AMI is a measure of total imbalance regarding all covariates. Additionally, Fig. 4 shows the course of the mean differences of the individual covariates. Covariates "Organic farming" and "Outdoor program" remained perfectly balanced throughout the pruning process. After a few pruned observations, the covariate "Grassland program" was also balanced. From 51 pruned observations, the covariate "Milk production" is below the cut-off value of 0.1. The covariate "Number of cows" reached this point at 117 pruned observations. Based on Figs. 3 and 4, a dataset was subsequently selected. The two extremes, either the full sample or a perfectly balanced sample, would both be accompanied by major disadvantages. The former would forgo the reduction in imbalance that can be observed after only a few pruned observations. For the latter, the remaining number of observations would be small, resulting in a large variance of the treatment effect. Therefore, the dataset was selected in which the standardized mean difference for all covariates is below the cut-off value of 0.1. Depending on the outcome variable model, the number of observations that had to be pruned to reach a value below 0.1 in standardized mean difference, differed. It lied between 100 and 119 pruned observations (Table 6). Based on these datasets, the treatment effect was then calculated using linear regression, including the respective outcome variable model as depend variable and the treatment variable "Free stall program" and the five confounders "Number of cows", "Milk production", "Organic farming", "Outdoor program", and "Grassland program" as independent variable.

The resulting coefficients of the free stall program of all eight models are displayed in Table 6. Full model results of all regression models including estimated coefficients of the five confounders can be found in Appendix E. To check that assumptions of linear regression (e.g. linearity and homoscedasticity) were fulfilled, fitted values and residuals were visually inspected. Additionally, a potential problem of heteroscedasticity was adressed by relying on robust standard errors.

Treatment effects were significant at the 5% level for incidence of clinical mastitis (−3.66 per 100 cow-years), incidence of culled cows (−1.61 per 100 cow-years), veterinary costs (−42.44 per cow-year), and incidences of antibiotic treatments for dry-cow therapy (−8.80 per 100 cow-years). Incidence of antibiotic treatments for mastitis therapy (−7.83 per 100 cow-years) and the incidence of total intramammary antibiotic treatments (−15.88 per 100 cow-years) were significant at the 1%-level. No effect could be found for average herd somatic cell count and number of cows with somatic cell count level above 150.000.

The results presented in Table 6 were calculated based on one dataset selected from the Matching Frontier. However, it is also possible to calculate and visualize the treatment effect along the entire Matching Frontier. This allows us to see how the treatment effects changed with increasing balance, which is shown for the "Antibiotic dry-cow" model in Fig. 5. With increasing balance, the treatment effects decreased. An estimate based on the complete sample would thus overestimate the treatment effect. Additionally, the 90% confidence intervals are plotted, which increased with decreasing sample size. Fig. 6 again shows the same treatment effects. However, here, the model dependence, calculated according to Athey and Imbens (2015), is additionally plotted. It can be seen that the model dependence decreased with increasing balance.

Pruning of observations changes the quantity of interest. This is shown in Fig. 7, where the scaled covariate means of all remaining farms in the sample are plotted as a function of the number of observations pruned. The pruning decreased the average number of cows and the milk yield along the Matching Frontier. The reason for this is that large free stall farms with high milk production levels were pruned first. In the full sample, the maximum number of cows was 150 cows, and with 117 pruned observations, it was only 53 cows. In the case of the covariate "Grassland program", the proportion of farms not participating in the program decreased significantly. This was also true for the outdoor

Table 5
Summary statistics of categorical covariates from treatment group ("Free stall") and control group ("Tie stall"). To illustrate the differences between the groups, means based on categorical numbering of variables were calculated. A t-test was used to determine whether the difference between the means was statistically significant. Numbers are shown for the model where "Incidence of clinical mastitis" was the outcome variable. Numbers for the other 7 outcome models can be found in Appendix D.

Variable	Treatment group	Description	%	Mean[a]	t-value (p-value)
	Free stall (n = 179)	1 "< 5.000"	3.4%	4.02	
		2 "5.000–6.000"	11.2%		
		3 "6.001–7.000"	22.3%		
		4 "7.001–8.000"	27.4%		
		5 "8.001–9.000"	20.1%		
		6 "9.001–10.000"	10.1%		
		7 ">10.000"	5.6%		
Milk production[b]					−5.83 (0.000)
	Tie stall (n = 229)	1 "< 5.000"	5.2%	3.23	
		2 "5.000–6.000"	27.5%		
		3 "6.001–7.000"	27.5%		
		4 "7.001–8.000"	24.0%		
		5 "8.001–9.000"	11.4%		
		6 "9.001–10.000"	3.5%		
		7 "> 10.000"	0.9%		
	Free stall	1 Plain zone	51.4%	2.06	
		2 Hilly zone	21.1%		
		3 Mountain zone I	8.6%		
		4 Mountain zone II	9.7%		
		5 Mountain zone III	6.9%		
		6 Mountain zone IV	2.3%		
Agricultural zone[c]					5.19 (0.000)
	Tie stall	1 Plain zone	26.7%	2.81	
		2 Hilly zone	14.5%		
		3 Mountain zone I	26.2%		
		4 Mountain zone II	20.4%		
		5 Mountain zone III	8.1%		
		6 Mountain zone IV	4.1%		
	Free stall	1 None	2.8%	4.07	
		2 VET certificate[e]	4.5%		
		3 VET diploma[f]	33.0%		
		4 Diploma of higher education[g]	16.8%		
		5 Advanced diploma[h]	34.6%		
		6 College[i]	2.8%		
		7 University[j]	5.6%		
Agricultural education[d]					−4.25 (0.000)
	Tie stall	1 None	3.9%	3.54	
		2 VET certificate[e]	4.8%		
		3 VET diploma[f]	51.5%		
		4 Diploma of higher education[g]	18.3%		
		5 Advanced diploma[h]	18.3%		
		6 College[i]	0.4%		
		7 University[j]	2.6%		

[a] To calculate the mean, categorical numbering was used (1–7 for "Milk production" and" Agricultural education", 1–6 for "Agricultural zone").
[b] Average level of yearly milk yield per cow.
[c] Agricultural zone in which farm is placed (plain zone, hill zone, and mountain zone).
[d] Farmer's agricultural education level. For further information on the Swiss educational system, please see https://www.edk.ch/en/education-system/diagram
[e] Federal vocational education and training (apprenticeship) certificate.
[f] Federal vocational education and training (apprenticeship) diploma.
[g] Federal diploma of higher education.
[h] Advanced federal diploma of higher education.
[i] College of higher education.
[j] University incl. federal institute of technology /University of applied sciences.

program. Due to the change in the quantity of interest, found effects can therefore no longer be applied to all farms in the treatment group. For example, the eight treatment effects found above cannot be generalized to free stall farms having >53 cows.

Finally, matching based on five confounders was compared with matching based on seven confounders, including covariates "Agricultural zone" and "Agricultural education." The corresponding Matching Frontier and mean covariate plots are shown in Figs. 8 and 9. The AMI is higher along the entire frontier than in the five confounder case. Only with 300 pruned observations are all standardized mean differences below the cut-off of 0.1. This situation was unsatisfactory. Therefore, it was examined whether the matching based on five confounders is not advantageous overall, also regarding the two covariates "Agricultural zone" and "Agricultural education." For this purpose, the standardized mean differences from the matching based on five confounders are shown again in Fig. 10. Additionally, the mean differences for the two covariates "Agricultural zone" and "Agricultural education" are plotted, which were not included in the matching. As expected, these two covariates are not perfectly balanced. However, due to the matching based on the other five confounders and pruning of observations, the imbalance in these two covariates was also lower than in the unmatched full sample. At 115 pruned observations, all seven covariates are below or close to the cut-off value of 0.1 regarding standardized mean difference. The five confounder solution thus seems to have an advantage over

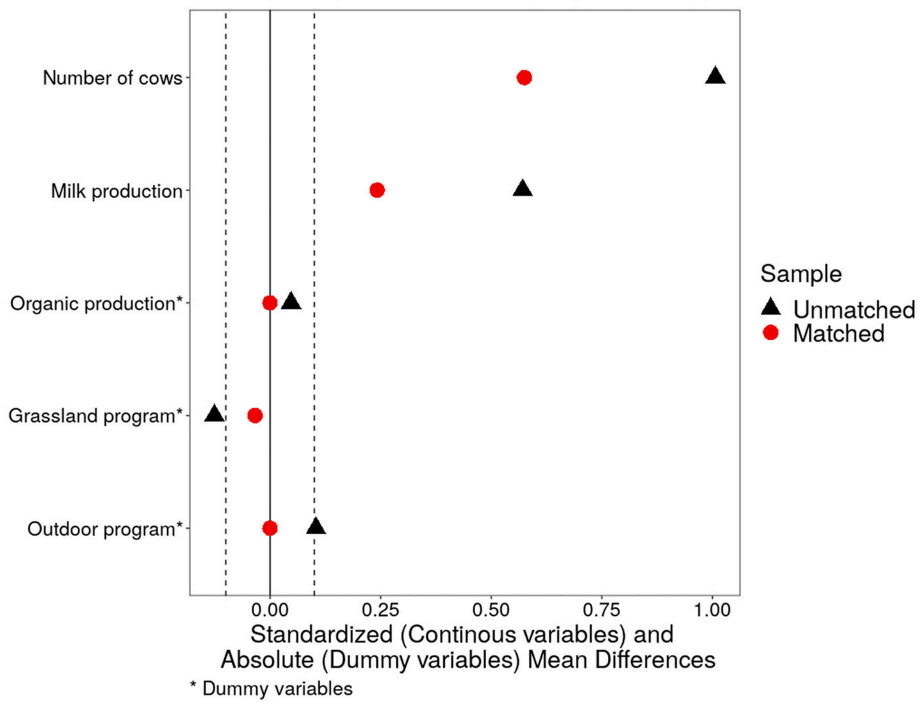

Fig. 2. Mean differences of covariates before (black triangle) and after matching (red dots). Results are shown for model with outcome variable "Incidence of antibiotic treatments for dry-cow therapy". Standardized mean difference for continuous variables was calculated dividing mean differences by standard deviation of the treated group. In case of dummy variables, absolute mean differences are displayed. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

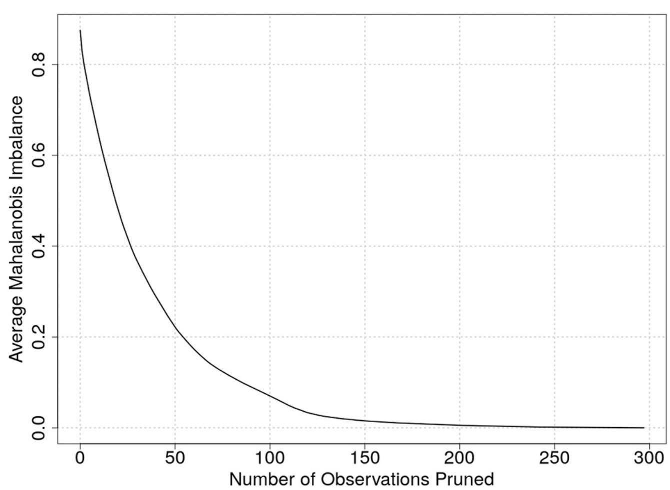

Fig. 3. Matching Frontier of outcome variable model "Incidence of antibiotic treatments for dry-cow therapy". The Matching Frontier comprises a set of matched subsamples that characterizes the trade-off between imbalance and sample size. For each possible sample size, an algorithm is used to determine the matching solution with the lowest imbalance. All datasets on this frontier are optimal, meaning that there exists no matching solution with a lower imbalance for a given sample size or a higher sample size for a fixed imbalance. Imbalance was measured by average Mahalanobis imbalance (AMI) metric, that is the Mahalanobis distance between each unit in the treatment group and the closest unit in the control group averaged over all units.

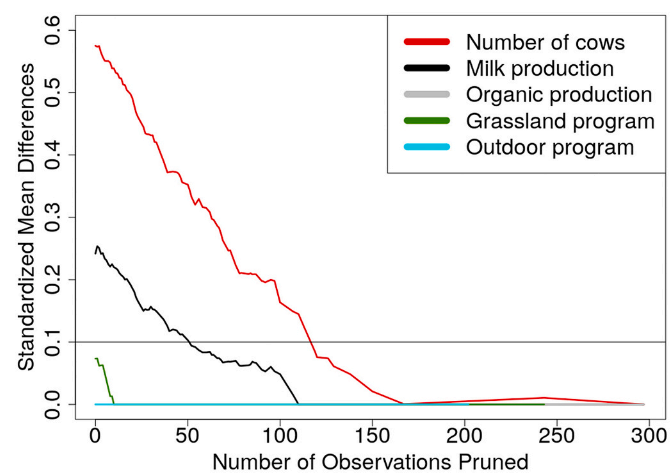

Fig. 4. Standardized mean differences of five confounders along the Matching Frontier.
Results are shown for model with outcome variable "Incidence of antibiotic treatments for dry-cow therapy". Standardized mean difference was calculated dividing mean differences by standard deviation of the treated group.

the seven confounder solution, also regarding the variables "Agricultural zone" and "Agricultural education." Significantly fewer pruned observations were sufficient to achieve a reasonable balance. Since the five confounder solution was more advantageous than the seven confounder solution, the treatment effects were not recalculated.

4. Discussion and conclusion

This study is the first application of the Matching Frontier method in the field of livestock science. The method proved to be a good complement to previous matching approaches in observational studies. It has especially added value in situations where the groups to be compared differ considerably in their characteristics. In such cases, matching alone is often insufficient to achieve sufficient balance and overlap, but additional pruning of observations is necessary.

The Matching Frontier method, developed by King et al. (2017), has several advantages over manual pruning approaches. First, all datasets on the Matching Frontier are optimal regarding balance and sample size. Researchers can thus focus on the trade-off between these two. Second, by visualizing the covariate mean difference along the Matching Frontier, the exact dataset can be selected where all covariates are below the desired cut-off value. It would be a great challenge to match exactly this dataset with manual approaches. Third, by calculating and visualizing

Table 6

Treatment effects of free stall program on outcome variables after matching and pruning observations. The treatment effects were determined using 8 different linear regression models. The dependent variable was the corresponding outcome variable, independent variables were the treatment variable "Free stall program" and the five confounders "Number of cows", "Milk production", "Organic farming", "Outdoor program", and "Grassland program". Linear regression was performed in each case with the largest possible sample at which all confounders at the same time had a maximum standardized mean difference of 0.1. In addition, means of treatment group ("Free stall") and control group ("Tie stall"), robust standard errors, t-values and p-values of statistical significance of treatment effects are displayed. For statistically significant treatment effects, the difference in percentage between treated and control group is reported. Full model results of all regression models including estimated coefficients of the five confounders can be found in Appendix E.

Outcome variable	Number of observations pruned	Remaining number of observations included in linear regression	Treatment group	Means after matching and pruning	Treatment effect of free stall program	% Difference	Robust standard error	t-value	p-value
Mastitis[a]	100	308	Free stall Tie stall	11.16 14.80	−3.66*	−25%	1.69	−2.16	0.031
Culled cows[b]	101	304	Free stall Tie stall	3.69 5.33	−1.61*	−30%	0.66	−2.42	0.016
HSCC[c]	111	288	Free stall Tie stall	97.95 88.93	8.37		5.57	1.50	0.134
High cell[d]	118	241	Free stall Tie stall	30.82 31.32	−0.91		3.26	−0.28	0.781
Veterinary costs[e]	119	225	Free stall Tie stall	152.86 197.06	−42.44*	−22%	17.94	−2.37	0.019
Antibiotic mastitis[f]	103	290	Free stall Tie stall	16.52 24.25	−7.83**	−32%	2.39	−3.28	0.001
Antibiotic dry-cow[g]	117	283	Free stall Tie stall	33.50 41.89	−8.80*	−21%	4.35	−2.02	(0.044)
Antibiotic total[h]	111	270	Free stall Tie stall	50.37 65.71	−15.88**	−23%	4.99	−3.18	0.002

Levels of significance: *$p < 0.05$; **$p < 0.01$.

[a] Incidence of clinical mastitis in 2018.
[b] Incidence of culled cows due to udder health problems in 2018.
[c] Average Herd somatic cell count throughout 2018.
[d] Number of cows with somatic cell count above 150.000 at least once in 2018.
[e] Cumulative veterinary costs for dairy cows throughout 2018.
[f] Incidence of intramammary antibiotic treatments for mastitis therapy in 2018.
[g] Incidence of antibiotic treatments for dry-cow therapy in 2018.
[h] Incidence of total intramammary antibiotic treatments in 2018.

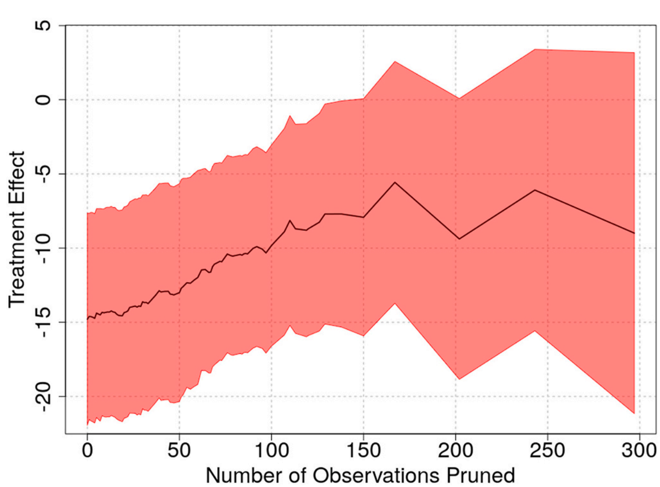

Fig. 5. Treatment effect of free stall program on incidence of antibiotic treatments for dry-cow therapy (per 100 cow-years) and 90%-confidence interval along the Matching Frontier.

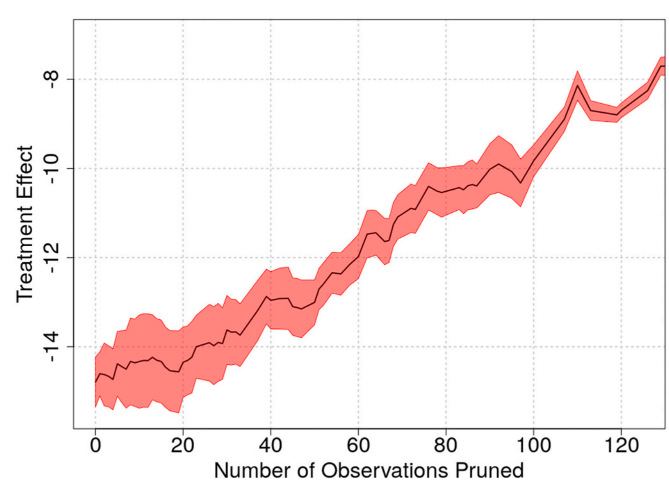

Fig. 6. Treatment effect of free stall program on incidence of antibiotic treatments for dry-cow therapy (per 100 cow-years) and model dependence according to Athey and Imbens (2015) along the Matching Frontier.

model dependence, researchers can see how much is actually gained regarding reduced model dependence by pruning observations. The fourth point is shown in the comparison between the matching solution based on five confounders and the matching solution based on seven confounders. Pruning of observations can also improve the balance in covariates that were not included in the matching. Thus, a matching based on just a subset of confounders can induce a better result regarding overall balance than a matching based on all confounders.

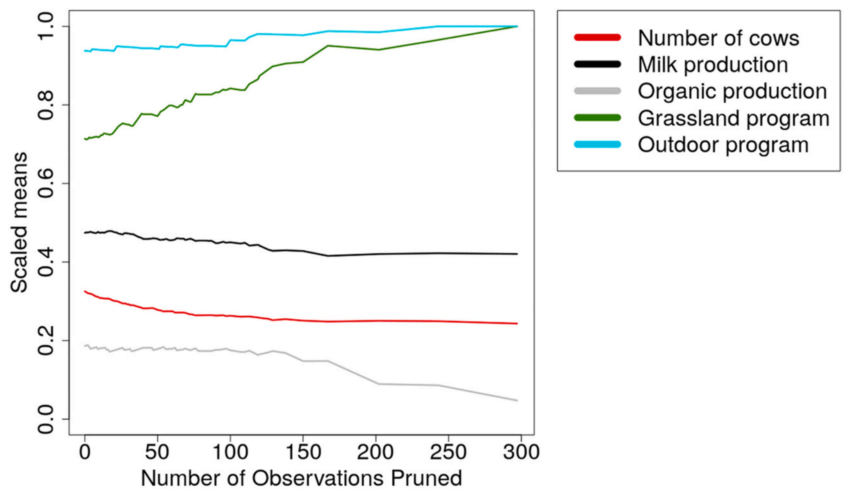

Fig. 7. Covariate Means along the Matching Frontier. Results are shown for model with outcome variable "incidence of antibiotic treatments for dry-cow therapy".

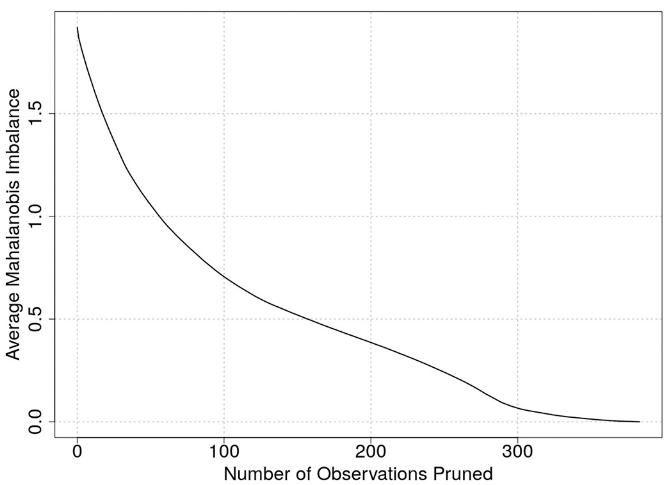

Fig. 8. Matching Frontier for matching based on seven confounders. The Matching Frontier comprises a set of matched subsamples that characterizes the trade-off between imbalance and sample size. For each possible sample size, an algorithm is used to determine the matching solution with the lowest imbalance. All datasets on this frontier are optimal, meaning that there exists no matching solution with a lower imbalance for a given sample size or a higher sample size for a fixed imbalance. Imbalance was measured by average Mahalanobis imbalance (AMI) metric, that is the Mahalanobis distance between each unit in the treatment group and the closest unit in the control group averaged over all units.

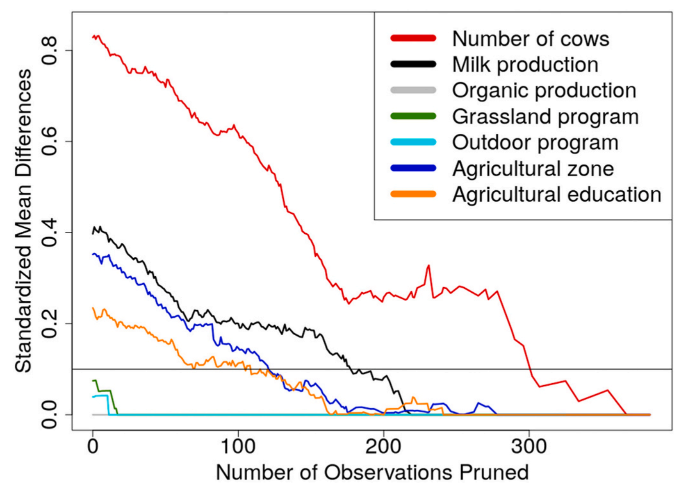

Fig. 9. Standardized mean difference along the Matching Frontier for the matching based on seven confounders. Standardized mean difference was calculated dividing mean differences by standard deviation of the treated group.

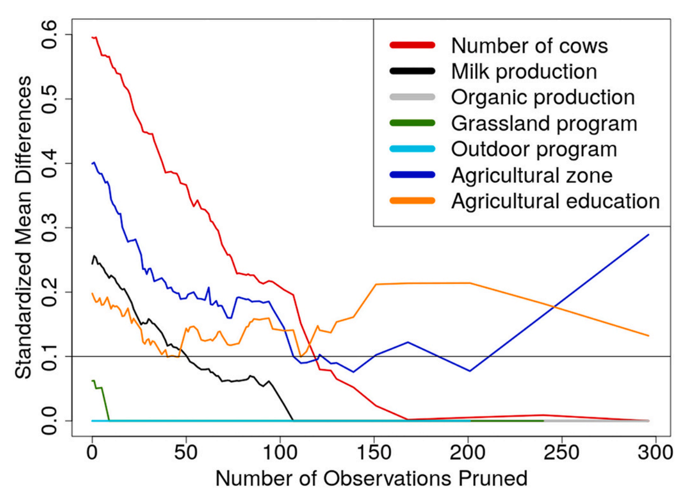

Fig. 10. Standardized mean difference along the Matching Frontier for the matching based on five confounders; in addition, standardized mean differences for the two covariates "Agricultural zone" and "Agricultural education".

Whether this is actually the case depends strongly on the structure of the dataset and the confounders used. However, checking this is facilitated by the Matching Frontier method.

The Matching Frontier method could have been useful in Odermatt et al. (2018), for example. As in this study, farms included in the free stall program were compared with non-participants, and balance was increased using a matching approach. Standardized mean differences of the matching solution are not explicitly given. However, from the given absolute means for the treatment and control groups, it is possible to conclude that the standardized mean differences in some covariates were above the cut-off value of 0.1. Pruning of single observations would probably have allowed a higher balance without sacrificing a considerable sample size.

The adequacy of the indicators used for the areas studied in the present study has been demonstrated several times. The four variables, "Mastitis", "Culled cows", "HSCC", and "High cell", have proven to be useful udder health indicators in various studies (Alvåsen et al., 2012; Bartlett et al., 2001; Bradley et al., 2007; Gordon et al., 2013; Madouasse et al., 2010; Miller et al., 2008; Olde Riekerink et al., 2008; O'Reilly et al., 2006; Pantoja et al., 2009; Peeler et al., 2002; Schukken et al., 2003; Valde et al., 2005, 1997; van den Borne et al., 2010) Veterinary costs were included in the study because they can affect the profitability of dairy farming, which can be an important consideration for farmers (Odermatt et al., 2018). Farmers were additionally asked how much of the total veterinary costs were spent on insemination. This was because, depending on the farm, insemination was performed by the veterinarian, breeding association, or farm manager. Therefore, to make veterinary costs comparable between farms, they had to be adjusted for artificial insemination costs. Antibiotic treatment incidence has proven to be a useful indicator of the amount of antibiotic use in three studies in Switzerland (Menéndez González et al., 2010; Schaeren, 2006; Spycher et al., 2002).

The present analysis, based on the Matching Frontier, suggests that the free stall program with increased lying comfort has a positive effect on udder health, veterinary costs, and antibiotic usage. Thus, free stall housing is not only more accepted by non-producers, but also seems to be associated with benefits for farmers and public health.

Incidence rates for mastitis, with a median of 11.65 cases per 100 cow-years (mean 12.99), were in the range of another Swiss questionnaire study, which reported a median of 11.6 cases per 100 cow-years (mean 14.7) (Gordon et al., 2013). Studies in Denmark (Bartlett et al., 2001), Norway (Valde et al., 2005), England and Wales (Bradley et al., 2007), Canada (Olde Riekerink et al., 2008), and the Netherlands (van den Borne et al., 2010) reported higher values. However, it is important to distinguish that study farmers in countries other than Switzerland were asked to report the number of quarters affected rather than the number of cows affected (Gordon et al., 2013). The positive effect of the free stall program on the incidence of mastitis is in accordance with results found in previous studies (Hultgren, 2002; Olde Riekerink et al., 2008; Valde et al., 1997). Mastitis incidences in free stalls were, on average, lower by 25%. The size of the effect is in the range of other studies, where incidences in free stalls were lower by 25% (Valde et al., 1997) and 28% (Olde Riekerink et al., 2008). The average number of cows in the matched treatment group was 27 cows. A treated farm of this size would have one mastitis case less per year compared to a tie stall farm of the same size. Costs of mastitis are complex to estimate, and no standardized approach exists (Halasa et al., 2007; Heikkilä et al., 2012). Heikkilä et al. (2012) estimated the average costs of a clinical mastitis case based on Finnish dairy cows to 485 €, with a range from 209 € to 1006 €. On Canadian farms, the median cost for a clinical mastitis case was 744 Canadian dollars (CAD) (range: 50 CAD to 5.349 CAD) (Aghamohammadi et al., 2018). For Switzerland, costs of 209 CHF per cow-year at risk for clinical and subclinical mastitis were estimated (Heiniger et al., 2014). For the latter, no estimates per mastitis case were reported. Although the exact costs of mastitis are difficult to estimate, the difference found in one mastitis case per year for an average sized farm seems to be relevant for the profitability of the farm.

The mean of 4.31 culled cows per 100 cow-years (median 2.82) was lower than reported in a Norwegian study (mean of 7.6, median of six culled cows per 100 cow-years) (Valde et al., 2005). A study showed that premature culling increased the already high costs of mastitis by 28% (Heikkilä et al., 2012). The positive effect of the free stall program (−2.80 cows per 100 cow-years) thus suggests that farms can save costs caused by premature culling by keeping cows in free stalls. The effect cannot be compared to other studies, as it has not been investigated before.

No significant difference between the two groups could be found for average herd somatic cell count and the number of cows with cell counts level above 150.000. This is in accordance with two studies from Norway and Sweden (Hultgren, 2002; Valde et al., 1997). However, two studies in Switzerland found higher HSCC levels in free stalls (Bielfeldt et al., 2004; Gordon et al., 2013), whereas results in Dufour et al. (2011) indicated the opposite.

The result that farms with free stalls had significantly low veterinary costs follows a Swiss study (Odermatt et al., 2018). The effect in this study was higher, with a value of −42.44 CHF per cow-year compared to −19.32 CHF. This effect could have a significant impact on the profitability of milk production, which could be an argument for farmers switching to free stall housing. However, Odermatt et al. (2018) specified the limitations of veterinary costs as an indicator of the level of animal health. Low costs could be related to healthy animals, but also to the non-treatment of animals that need veterinary care.

Incidences of antibiotic treatments with a mean of 55.97 total intramammary treatments per 100 cow-years (20.09 treatments per 100 cow-years due to mastitis and 36.04 treatments per 100 cow-years for dry-cow therapy) were lower than reported in previous studies in Switzerland. In Menéndez González et al. (2010), the mean incidence of total intramammary antibiotic treatments was 76 per 100 cow-years (37 for mastitis and 39 for dry-cow therapy). Schaeren (2006) reported a mean incidence of 61 treatments per 100 cow-years (25 due to mastitis and 36 for dry-cow therapy). In Spycher et al. (2002), the mean incidence of treatments due to mastitis was 26 per 100 cow-years. The difference from previous studies is mainly in the treatments due to mastitis, whereas treatments for dry-cow therapy are in a similar range. The quantity of antibiotics sold for veterinary medicine has decreased in Switzerland in recent years (FOPH, 2020). However, it is unclear whether this can explain the difference from previous studies, as the amount of antibiotics sold for both mastitis and dry-cow therapy has decreased. The trend in this study was only evident in mastitis treatments.

The difference in antibiotic treatments found between farms with free stalls and farms with tie stalls is significant. Free stall farms thus had a 23% lower incidence of total intramammary antibiotic treatments, a 32% lower incidence of antibiotic treatments for mastitis therapy, and a 21% lower incidence of treatments for dry-cow therapy. The treatment effect of −7.83 per 100 cow-years in the incidence of antibiotic treatments for mastitis therapy is slightly lower than in Spycher et al. (2002), who reported 9.3 fewer treatments per 100 cow-years. To the best of the authors' knowledge, no other study has examined the effect of free stalling on antibiotic usage, to which the effect could be compared to. Fewer treatments with antibiotics in free stalls can induce lower costs for farmers and can thus be associated with economic benefits. However, free stalls also seem beneficial to the public, as the reduced use of antibiotics can have a positive impact on public health. This may further increase the acceptance of free stalls. However, it is important to note that the link between the incidence of intramammary treatments and public health is complex. First, intramammary antibiotic treatments represent only a part of antibiotic treatments in dairy cattle. With 75% (Schaeren, 2006) and 71% (Menéndez González et al., 2010), respectively, they took up a considerable portion of the treatments, but other treatments were not considered in this study. Additionally, it is difficult to draw conclusions from the number of treatments regarding the quantities, dosages, and duration of therapy used. Also, no active substances used were collected, which means that no statement can be made about possible critical antibiotics used, which could be important for human medicine. Also, the link between exposure to antibiotics and the emergence of resistance is complex and still under investigation. However, some studies have suggested that the higher the antibiotic use, the more likely it is that resistance will occur (Helke et al., 2017; Oliver et al., 2020, 2011; Omwenga et al., 2021; Pol and Ruegg, 2007).

Six of the eight effects studied are significant and indicate that free stall farming is associated with better udder health and, thus, cost benefits for the farmer. Notably, the effects found are not additive. The eight independent variables correlate with each other because they all represent some aspect of udder health. For example, the costs saved by fewer cases of mastitis are already reflected in veterinary costs, and the same is true for fewer treatments.

Through the combination of free-walking and increased lying comfort in the voluntary program, specific requirements responsible for the treatment effects cannot be determined from the results of this study. When applying the results to other countries, the prevailing specific housing conditions must therefore be considered.

This study has some limitations. Participation in the survey was voluntary. Therefore, a selection bias must be assumed. It is possible that more motivated farmers participated. Also, all data was self-reported. It cannot be verified whether all farmers correctly recorded treatments. Studies have shown that poorly kept treatment journals can be a problem, leading to an underestimation of the number of treatments (Carson et al., 2008; Menéndez González et al., 2010). A further limitation of the study is that omitted variable bias cannot be completely ruled out. Central differences between the two groups were controlled for by five and seven confounders, respectively. It was also shown that matching and pruning of observations reduced the balance in variables that were not included in the matching. However, it is possible that unobservable effects that could not be controlled for influence both the outcome and treatment variables. Therefore, further studies on this topic should include more covariates to address potential omitted-variable bias.

Overall, although no final conclusion can be drawn, the study results suggest that free stall housing, in combination with increased lying comfort, is associated with better udder health, lower veterinary costs, and public health benefits. Promoting free stall housing with public funding, as in Switzerland, therefore seems sensible. The Matching Frontier method applied in the study may facilitate future observational studies in which the treatment and control groups differ considerably in their characteristics.

Declarations of interest

None.

Funding

This work was supported by the National Research Program “Antibiotic Resistance” (NRP 72) of the Swiss National Science Foundation (Grand no. 407240_166997). The funding source had no involvement in study design, data collection and analysis, and interpretation of data.

Appendix A. Appendix

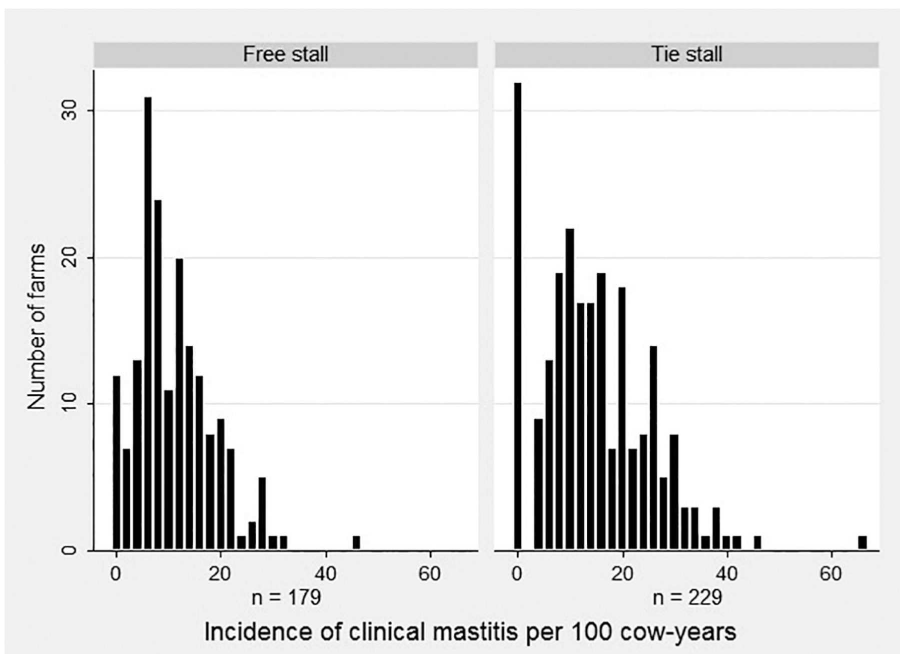

Fig. A1. Histogram of incidence of clinical mastitis.

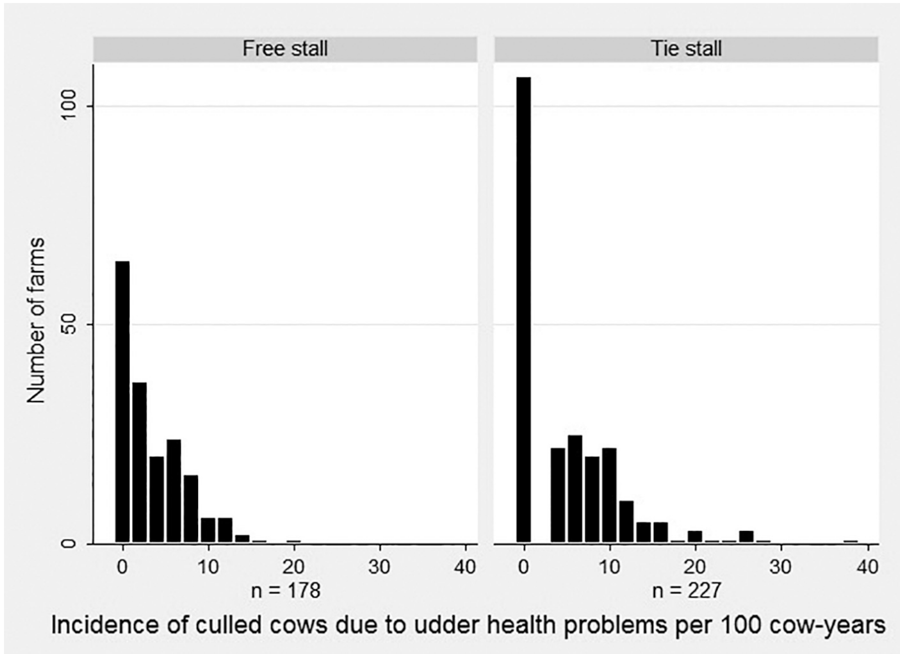

Fig. A2. Histogram of incidence of culled cows due to udder health problems.

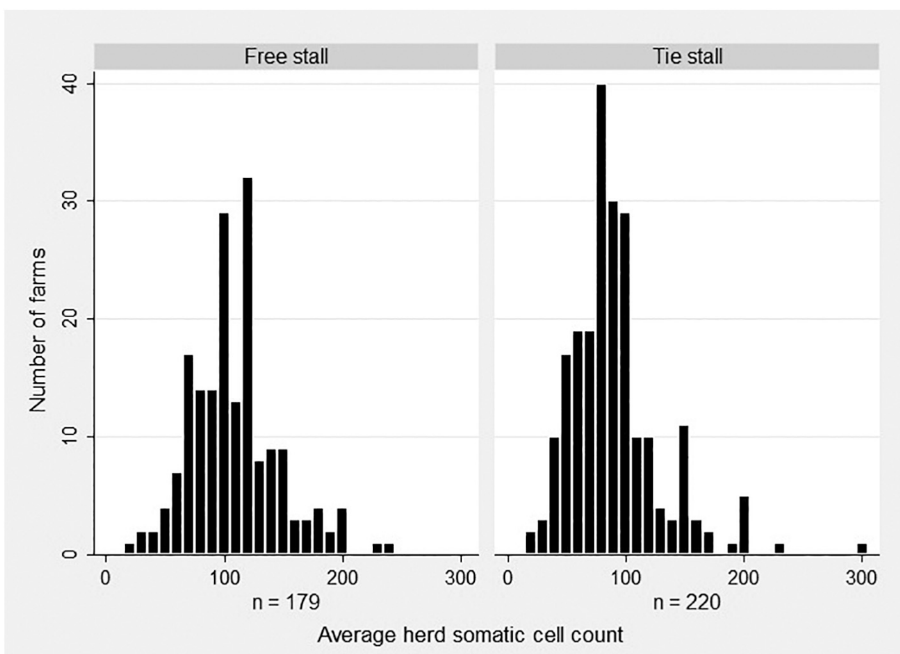

Fig. A3. Histogram of average herd somatic cell count.

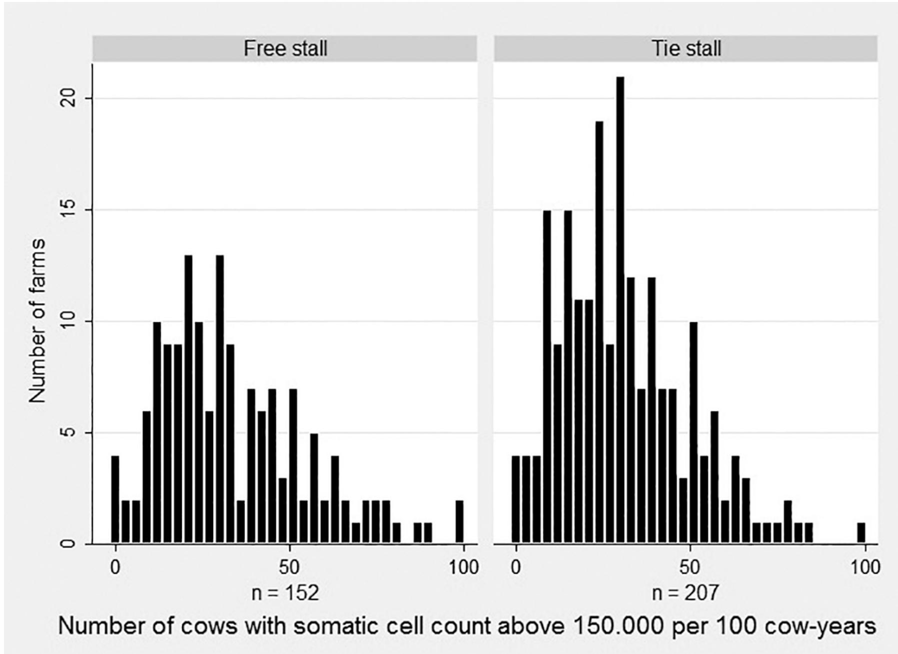

Fig. A4. Histogram of number of cows with somatic cell count above 150.000.

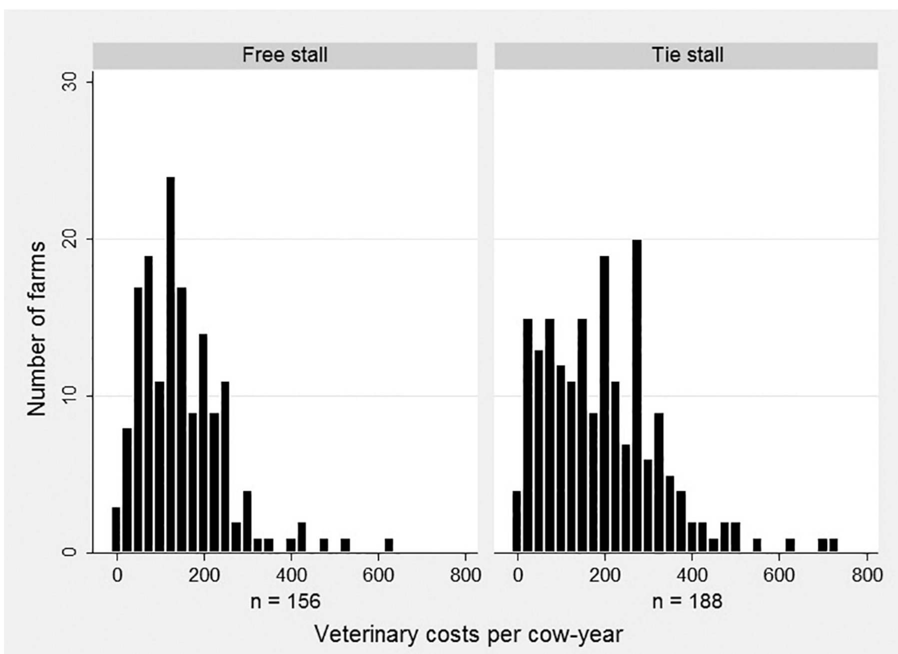

Fig. A5. Histogram of veterinary costs (in Swiss franc (CHF)).

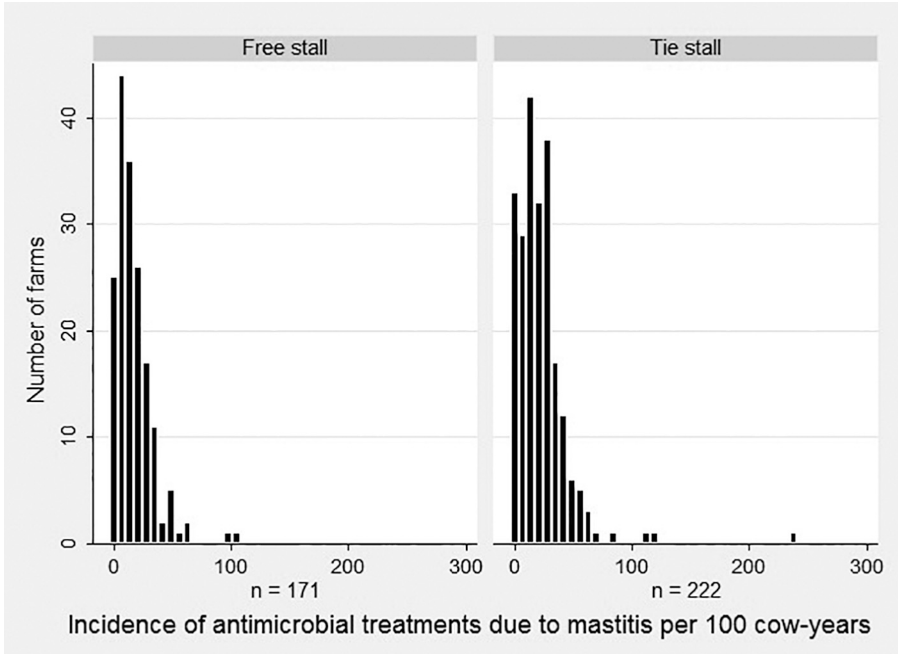

Fig. A6. Histogram of intramammary antibiotic treatments for mastitis therapy.

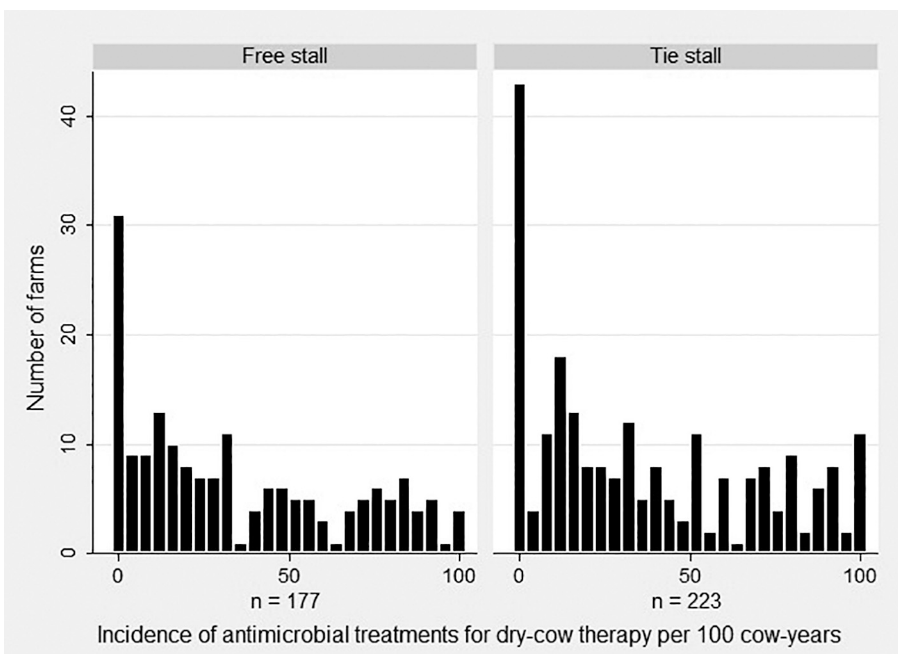

Fig. A7. Histogram of antibiotic treatments for dry-cow therapy.

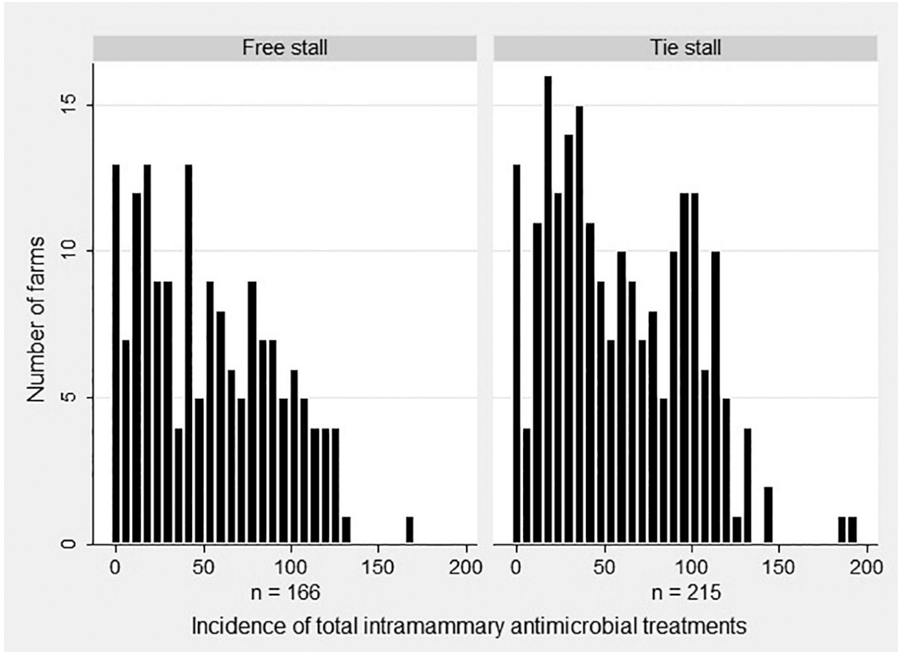

Fig. A8. Histogram of total intramammary antibiotic treatments.

Appendix B. Appendix

Summary statistics of continues covariate "Number of cows" from treatment group ("Free stall") and control group ("Tie stall") for 7 different outcome variable models. Numbers for the model where "Incidence of clinical mastitis" was the outcome variable can be found in Table 3.

Outcome variable model	Variable	Treatment group	N	Mean	SD	Median	Min	Max
Culled cows[a]	Number of cows	Free stall	178	37.03	18.67	35	4	92
		Tie stall	227	18.85	7.94	19	3	45
HSCC[b]	Number of cows	Free stall	179	38.45	18.78	35	5	92
		Tie stall	220	19.12	7.91	19	3	45
High cell[c]	Number of cows	Free stall	152	37.85	19.01	35	4	92
		Tie stall	207	18.58	7.85	18	3	45
Veterinary costs[d]	Number of cows	Free stall	156	38.92	18.81	36	6	92
		Tie stall	188	18.82	8.20	18	4	45
Antibiotic mastitis[e]	Number of cows	Free stall	171	37.19	18.89	35	4	92
		Tie stall	222	18.36	7.87	18	3	45
Antibiotic dry-cow[f]	Number of cows	Free stall	177	38.16	19.39	35	4	92
		Tie stall	223	18.65	7.94	18	3	45
Antibiotic total[g]	Number of cows	Free stall	166	37.61	18.96	35	4	92
		Tie stall	215	18.41	7.89	18	3	45

[a] Incidence of culled cows due to udder health problems in 2018.
[b] Average Herd somatic cell count throughout 2018.
[c] Number of cows with somatic cell count above 150.000 at least once in 2018.
[d] Cumulative veterinary costs for dairy cows throughout 2018.
[e] Incidence of intramammary antibiotic treatments for mastitis therapy in 2018.
[f] Incidence of antibiotic treatments for dry-cow therapy in 2018.
[g] Incidence of total intramammary antibiotic treatments in 2018.

Appendix C. Appendix

Summary statistics of dummy covariates from treatment group ("Free stall") and control group ("Tie stall"). Numbers for the model where "Incidence of clinical mastitis" was the outcome variable can be found in Table 4. The percentage reflects the proportion of organic farms and the proportion of participants in the outdoor and grassland program in the respective group.

Outcome variable model	Variable	Treatment group	Number of farms	%
Culled cows[c]	Organic farming	Free stall	178	19.1%

(continued on next page)

(*continued*)

Outcome variable model	Variable	Treatment group	Number of farms	%
		Tie stall	227	13.7%
	Outdoor program[a]	Free stall	178	93.3%
		Tie stall	227	84.1%
	Grassland program[b]	Free stall	178	66.9%
		Tie stall	227	80.6%
HSCC[d]	Organic farming	Free stall	179	18.4%
		Tie stall	220	12.7%
	Outdoor program[a]	Free stall	179	93.3%
		Tie stall	220	85.9%
	Grassland program[b]	Free stall	179	66.5%
		Tie stall	220	81.4%
High cell[e]	Organic farming	Free stall	152	20.4%
		Tie stall	207	14.0%
	Outdoor program[a]	Free stall	152	94.7%
		Tie stall	207	86.0%
	Grassland program[b]	Free stall	152	67.8%
		Tie stall	207	82.1%
Veterinary costs[f]	Organic farming	Free stall	156	20.5%
		Tie stall	188	14.4%
	Outdoor program[a]	Free stall	156	93.6%
		Tie stall	188	85.6%
	Grassland program[b]	Free stall	156	65.4%
		Tie stall	188	82.4%
Antibiotic mastitis[g]	Organic farming	Free stall	171	18.7%
		Tie stall	222	14.0%
	Outdoor program[a]	Free stall	171	93.6%
		Tie stall	222	83.3%
	Grassland program[b]	Free stall	171	66.1%
		Tie stall	222	81.1%
Antibiotic dry-cow[h]	Organic farming	Free stall	177	18.6%
		Tie stall	223	13.9%
	Outdoor program[a]	Free stall	177	93.8%
		Tie stall	223	83.4%
	Grassland program[b]	Free stall	177	66.7%
		Tie stall	223	81.6%
Antibiotic total[i]	Organic farming	Free stall	166	19.3%
		Tie stall	215	14.4%
	Outdoor program[a]	Free stall	179	94.0%
		Tie stall	229	83.3%
	Grassland program[b]	Free stall	179	66.9%
		Tie stall	229	81.4%

[a] Farms participating in the outdoor program must give their animals access to pasture or an outdoor run for a minimum period per year. For more detailed information please see Odermatt et al. (2018)
[b] For participating farms in the plain zone and hilly zone, 75% of the feed ration must be roughage, while 85% must be roughage for farms in the mountain zone.
[c] Incidence of culled cows due to udder health problems in 2018.
[d] Average Herd somatic cell count throughout 2018.
[e] Number of cows with somatic cell count above 150.000 at least once in 2018.
[f] Cumulative veterinary costs for dairy cows throughout 2018.
[g] Incidence of intramammary antibiotic treatments for mastitis therapy in 2018.
[h] Incidence of antibiotic treatments for dry-cow therapy in 2018.
[i] Incidence of total intramammary antibiotic treatments in 2018.

Appendix D. Appendix

Summary statistics of categorical covariates from treatment group ("Free stall") and control group ("Tie stall") for 7 different outcome variable models. Numbers for the model where "Incidence of clinical mastitis" was the outcome variable can be found in Table 5. To illustrate the differences between the groups, means based on categorical numbering of variables were calculated.

Outcome variable model	Variable	Treatment group	Number of farms	Mean[a]
Culled cows[e]	Milk production[b]	Free stall	178	4.02
		Tie stall	227	3.22
	Agricultural zone[c]	Free stall	178	2.01
		Tie stall	227	2.83
	Agricultural education[d]	Free stall	178	3.55
		Tie stall	227	4.11

(continued on next page)

(*continued*)

Outcome variable model	Variable	Treatment group	Number of farms	Mean[a]
HSCC[f]	Milk production[b]	Free stall	179	4.10
		Tie stall	220	3.28
	Agricultural zone[c]	Free stall	179	1.98
		Tie stall	220	2.81
	Agricultural education[d]	Free stall	179	4.11
		Tie stall	220	3.58
High cell[g]	Milk production[b]	Free stall	152	4.01
		Tie stall	207	3.27
	Agricultural zone[c]	Free stall	152	2.88
		Tie stall	207	2.10
	Agricultural education[d]	Free stall	152	4.11
		Tie stall	207	3.54
Veterinary costs[h]	Milk production[b]	Free stall	156	4.09
		Tie stall	188	3.28
	Agricultural zone[c]	Free stall	156	1.87
		Tie stall	188	2.86
	Agricultural education[d]	Free stall	156	4.29
		Tie stall	188	3.54
Antibiotic mastitis[i]	Milk production[b]	Free stall	171	4.03
		Tie stall	222	3.20
	Agricultural zone[c]	Free stall	171	2.00
		Tie stall	222	2.85
	Agricultural education[d]	Free stall	171	4.13
		Tie stall	222	3.51
Antibiotic dry-cow[j]	Milk production[b]	Free stall	177	4.04
		Tie stall	223	3.22
	Agricultural zone[c]	Free stall	177	1.97
		Tie stall	223	2.83
	Agricultural education[d]	Free stall	177	4.15
		Tie stall	223	3.53
Antibiotic total[k]	Milk production[b]	Free stall	166	3.19
		Tie stall	215	4.05
	Agricultural zone[c]	Free stall	166	1.99
		Tie stall	215	2.85
	Agricultural education[d]	Free stall	166	4.17
		Tie stall	215	3.52

[a] To calculate the mean, categorical numbering was used (1–7 for "Milk production" and" Agricultural education", 1–6 for "Agricultural zone"). Please see Table 5 for more information on categorical numbering.
[b] Average level of yearly milk yield per cow.
[c] Agricultural zone in which farm is placed (plain zone, hill zone, and mountain zone).
[d] Farmer's agricultural education level. For further information on the Swiss educational system, please see https://www.edk.ch/en/education-system/diagram
[e] Incidence of culled cows due to udder health problems in 2018.
[f] Average Herd somatic cell count throughout 2018.
[g] Number of cows with somatic cell count above 150.000 at least once in 2018.
[h] Cumulative veterinary costs for dairy cows throughout 2018.
[i] Incidence of intramammary antibiotic treatments for mastitis therapy in 2018.
[j] Incidence of antibiotic treatments for dry-cow therapy in 2018.
[k] Incidence of total intramammary antibiotic treatments in 2018.

Appendix E. Appendix

Results of the full regression models used to determine the treatment effect of the "Free stall program". The dependent variable was the corresponding outcome variable, independent variables were the treatment variable "Free stall program" and the five confounders "Number of cows", "Milk production", "Organic farming", "Outdoor program", and "Grassland program". Linear regression was performed in each case with the largest possible sample at which all confounders at the same time had a maximum standardized mean difference of 0.1. In addition, robust standard errors, t-values and *p*-values of statistical significance of all confounders are displayed.

Outcome variable model	Variable	Coefficient	Robust Standard Error	t-value	p-value
Mastitis[a]	Free stall program	−3.66	1.69	−2.16	0.031
$n = 308$	Number of cows	0.02	0.11	0.15	0.884
	Milk production[i]	1.55	0.65	2.40	0.017
	Organic farming	3.20	3.47	0.92	0.357
	Grassland program[j]	−2.70	3.80	−0.71	0.478
	Outdoor program[k]	3.35	2.80	1.19	0.233
	Intercept	7.08	3.40	2.08	0.038
Culled cows[b]	Free stall program	−1.61	0.66	−2.42	0.016

(*continued on next page*)

(continued)

Outcome variable model	Variable	Coefficient	Robust Standard Error	t-value	p-value
$n = 304$	Number of cows	−0.05	0.04	−1.18	0.237
	Milk production[i]	0.75	0.32	2.34	0.020
	Organic farming	0.51	1.05	0.48	0.630
	Grassland program[j]	1.73	0.87	2.00	0.047
	Outdoor program[k]	−2.41	2.60	−0.93	0.354
	Intercept	4.73	3.16	1.50	0.136
HSCC[c] $n = 288$	Free stall program	8.37	5.57	1.50	0.134
	Number of cows	0.66	0.27	2.43	0.016
	Milk production[i]	−1.08	2.97	−0.36	0.716
	Organic farming	10.92	7.11	1.53	0.126
	Grassland program[j]	12.88	11.51	1.12	0.264
	Outdoor program[k]	−19.02	7.62	−2.50	0.013
	Intercept	81.34	13.51	6.02	0.000
High cell[d] $n = 241$	Free stall program	−0.91	3.26	−0.28	0.782
	Number of cows	0.38	0.23	1.66	0.098
	Milk production[i]	−0.43	1.76	−0.24	0.807
	Organic farming	5.79	3.59	1.61	0.109
	Grassland program[j]	6.24	5.58	1.12	0.265
	Outdoor program[k]	3.93	5.48	0.72	0.474
	Intercept	12.92	9.52	1.36	0.176
Veterinary costs[e] $n = 225$	Free stall program	−42.44	17.94	−2.37	0.019
	Number of cows	−1.81	1.07	−1.69	0.093
	Milk production[i]	34.38	11.81	2.91	0.004
	Organic farming	13.77	25.97	0.53	0.596
	Grassland program[j]	18.13	42.75	0.42	0.672
	Outdoor program[k]	19.03	69.55	0.27	0.785
	Intercept	82.63	84.59	0.98	0.330
Antibiotic mastitis[f] $n = 290$	Free stall program	−7.83	2.39	−3.28	0.001
	Number of cows	0.10	0.14	0.75	0.456
	Milk production[i]	3.70	1.39	2.65	0.008
	Organic farming	−0.80	3.11	−0.26	0.797
	Grassland program[j]	2.60	4.60	0.57	0.572
	Outdoor program[k]	−7.85	10.81	−0.73	0.469
	Intercept	13.66	14.16	0.97	0.335
Antibiotic dry-cow[g] $n = 283$	Free stall program	−8.80	4.35	−2.02	0.044
	Number of cows	0.46	0.29	1.55	0.121
	Milk production[i]	1.87	2.60	0.72	0.473
	Organic farming	−25.25	4.27	−5.91	0.000
	Grassland program[j]	−12.76	6.31	−2.02	0.044
	Outdoor program[k]	14.65	9.70	1.51	0.132
	Intercept	24.12	13.74	1.76	0.080
Antibiotic total[h] $n = 270$	Free stall program	−15.88	4.99	−3.18	0.002
	Number of cows	0.57	0.31	1.84	0.068
	Milk production[i]	6.31	2.84	2.22	0.027
	Organic farming	−25.39	5.09	−4.99	0.000
	Grassland program[j]	−8.42	9.18	−0.92	0.360
	Outdoor program[k]	14.78	12.09	1.22	0.223
	Intercept	25.10	17.35	1.45	0.149

[a] Incidence of clinical mastitis in 2018.
[b] Incidence of culled cows due to udder health problems in 2018.
[c] Average Herd somatic cell count throughout 2018.
[d] Number of cows with somatic cell count above 150.000 at least once in 2018.
[e] Cumulative veterinary costs for dairy cows throughout 2018.
[f] Incidence of intramammary antibiotic treatments for mastitis therapy in 2018.
[g] Incidence of antibiotic treatments for dry-cow therapy in 2018.
[h] Incidence of total intramammary antibiotic treatments in 2018.
[i] Average level of yearly milk yield per cow.
[j] For participating farms in the plain zone and hilly zone, 75% of the feed ration must be roughage, while 85% must be roughage for farms in the mountain zone.
[k] Farms participating in the outdoor program must give their animals access to pasture or an outdoor run for a minimum period per year. For more detailed information please see Odermatt et al. (2018).

References

Abadie, A., Imbens, G.W., 2006. Large sample properties of matching estimators for average treatment effects. Econometrica 74, 235–267. https://doi.org/10.1111/j.1468-0262.2006.00655.x.

Aghamohammadi, M., Haine, D., Kelton, D.F., Barkema, H.W., Hogeveen, H., Keefe, G.P., Dufour, S., 2018. Herd-level mastitis-associated costs on Canadian dairy farms. Front. Vet. Sci. 5 https://doi.org/10.3389/fvets.2018.00100.

Alvåsen, K., Jansson Mörk, M., Hallén Sandgren, C., Thomsen, P.T., Emanuelson, U., 2012. Herd-level risk factors associated with cow mortality in Swedish dairy herds. J. Dairy Sci. 95, 4352–4362. https://doi.org/10.3168/jds.2011-5085.

Athey, S., Imbens, G., 2015. A measure of robustness to misspecification. Am. Econ. Rev. 105, 476–480. https://doi.org/10.1257/aer.p20151020.

Austin, P.C., 2009. Type I error rates, coverage of confidence intervals, and variance estimation in propensity-score matched analyses. Int. J. Biostat. 5 https://doi.org/10.2202/1557-4679.1146.

Austin, P.C., 2013. The performance of different propensity score methods for estimating marginal hazard ratios. Stat. Med. 32, 2837–2849. https://doi.org/10.1002/sim.5705.

Austin, P.C., Small, D.S., 2014. The use of bootstrapping when using propensity-score matching without replacement: a simulation study. Stat. Med. 33, 4306–4319. https://doi.org/10.1002/sim.6276.

Bartlett, P.C., Agger, J.F., Houe, H., Lawson, L.G., 2001. Incidence of clinical mastitis in Danish dairy cattle and screening for non-reporting in a passively collected national surveillance system. Prev. Vet. Med. 48, 73–83. https://doi.org/10.1016/S0167-5877(00)00192-6.

Bielfeldt, J.C., Badertscher, R., Tölle, K.-H., Krieter, J., 2004. Factors influencing somatic cell score in Swiss dairy production systems. Schweiz. Arch. Für Tierheilkd. 146, 555–560. https://doi.org/10.1024/0036-7281.146.12.555.

Bradley, A.J., Leach, K.A., Breen, J.E., Green, L.E., Green, M.J., 2007. Survey of the incidence and aetiology of mastitis on dairy farms in England and Wales. Vet. Rec. 160, 253–258. https://doi.org/10.1136/vr.160.8.253.

Carson, C.A., Reid-Smith, R., Irwin, R.J., Martin, W.S., McEwen, S.A., 2008. Antimicrobial use on 24 beef farms in Ontario. Can. J. Vet. Res. 72, 109–118.

Dufour, S., Fréchette, A., Barkema, H.W., Mussell, A., Scholl, D.T., 2011. Invited review: effect of udder health management practices on herd somatic cell count. J. Dairy Sci. 94, 563–579. https://doi.org/10.3168/jds.2010-3715.

European Commission. Directorate-General for Health and Food Safety, 2016. Attitudes of Europeans towards Animal Welfare: Report. European Commission. https://data.europa.eu/doi/10.2875/884639.

FOAG, 2020. Agricultural Report 2020. Federal Office for Agriculture, Bern.

FOPH, 2020. Swiss Antibiotic Resistance Report 2020. Usage of Antibiotics and Occurrence of Antibiotic Resistance in Switzerland. Federal Office of Public Health and Federal Food Safety and Veterinary Office.. FOPH publication number: 2020-OEG-6.

Gayat, E., Resche-Rigon, M., Mary, J.-Y., Porcher, R., 2012. Propensity score applied to survival data analysis through proportional hazards models: a Monte Carlo study. Pharm. Stat. 11, 222–229. https://doi.org/10.1002/pst.537.

Gelman, A., Hill, J., Vehtari, A., 2020. Regression and Other Stories, 1st ed. Cambridge University Press. https://doi.org/10.1017/9781139161879.

Gordon, P.F., van den Borne, B.H., Reist, M., Kohler, S., Doherr, M.G., 2013. Questionnaire-based study to assess the association between management practices and mastitis within tie-stall and free-stall dairy housing systems in Switzerland. BMC Vet. Res. 9, 200. https://doi.org/10.1186/1746-6148-9-200.

Greifer, N., 2021a. MatchIt: Getting Started [WWW Document]. URL. https://cran.r-project.org/web/packages/MatchIt/vignettes/MatchIt.html (accessed 6.20.21).

Greifer, N., 2021b. Estimating Effects After Matching [WWW Document]. URL. https://cran.r-project.org/web/packages/MatchIt/vignettes/MatchIt.html (accessed 6.20.21).

Halasa, T., Huijps, K., Østerås, O., Hogeveen, H., 2007. Economic effects of bovine mastitis and mastitis management: a review. Vet. Q 29, 18–31. https://doi.org/10.1080/01652176.2007.9695224.

Heikkilä, A.-M., Nousiainen, J.I., Pyörälä, S., 2012. Costs of clinical mastitis with special reference to premature culling. J. Dairy Sci. 95, 139–150. https://doi.org/10.3168/jds.2011-4321.

Heiniger, D., van den Borne, B.H.P., Lechner, I., Tschopp, A., Strabel, D., Steiner, A., Meier, H., 2014. Kosten-Nutzen-Analyse einer Intervention zur Verbesserung der Eutergesundheit in Schweizer Milchviehbetrieben. Schweiz. Arch. Für Tierheilkd. 156, 473–481. https://doi.org/10.1024/0036-7281/a000634.

Helke, K.L., McCrackin, M.A., Galloway, A.M., Poole, A.Z., Salgado, C.D., Marriott, B.P., 2017. Effects of antimicrobial use in agricultural animals on drug-resistant foodborne salmonellosis in humans: a systematic literature review. Crit. Rev. Food Sci. Nutr. 57, 472–488. https://doi.org/10.1080/10408398.2016.1230088.

Hill, J., Reiter, J.P., 2006. Interval estimation for treatment effects using propensity score matching. Stat. Med. 25, 2230–2256. https://doi.org/10.1002/sim.2277.

Ho, D.E., Imai, K., King, G., Stuart, E.A., 2007. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. Polit. Anal. 15, 199–236. https://doi.org/10.1093/pan/mpl013.

Hultgren, J., 2002. Foot/leg and udder health in relation to housing changes in Swedish dairy herds. Prev. Vet. Med. 53, 167–189. https://doi.org/10.1016/S0167-5877(01)00279-3.

Imbens, G.W., Wooldridge, J.M., 2009. Recent developments in the econometrics of program evaluation. J. Econ. Lit. 47, 5–86. https://doi.org/10.1257/jel.47.1.5.

King, G., Lucas, C., Nielsen, R., 2015. MatchingFrontier: R Package for Computing the Matching Frontier.

King, G., Lucas, C., Nielsen, R.A., 2017. The balance-sample size frontier in matching methods for causal inference. Am. J. Polit. Sci. 61, 473–489. https://doi.org/10.1111/ajps.12272.

Kreif, N., Gruber, S., Radice, R., Grieve, R., Sekhon, J.S., 2016. Evaluating treatment effectiveness under model misspecification: a comparison of targeted maximum likelihood estimation with bias-corrected matching. Stat. Methods Med. Res. 25, 2315–2336. https://doi.org/10.1177/0962280214521341.

Ly, L.H., Ryan, E.B., Weary, D.M., 2021. Public attitudes toward dairy farm practices and technology related to milk production. PLoS One 16, e0250850. https://doi.org/10.1371/journal.pone.0250850.

Madouasse, A., Huxley, J.N., Browne, W.J., Bradley, A.J., Green, M.J., 2010. Somatic cell count dynamics in a large sample of dairy herds in England and Wales. Prev. Vet. Med. 96, 56–64. https://doi.org/10.1016/j.prevetmed.2010.05.005.

Menéndez González, S., Steiner, A., Gassner, B., Regula, G., 2010. Antimicrobial use in Swiss dairy farms: quantification and evaluation of data quality. Prev. Vet. Med. 95, 50–63. https://doi.org/10.1016/j.prevetmed.2010.03.004.

Miller, R.H., Kuhn, M.T., Norman, H.D., Wright, J.R., 2008. Death losses for lactating cows in herds enrolled in dairy herd improvement test plans. J. Dairy Sci. 91, 3710–3715. https://doi.org/10.3168/jds.2007-0943.

Odermatt, B., Keil, N., Lips, M., 2018. Animal welfare payments and veterinary and insemination costs for dairy cows. Agriculture 9, 3. https://doi.org/10.3390/agriculture9010003.

Olde Riekerink, R.G.M., Barkema, H.W., Kelton, D.F., Scholl, D.T., 2008. Incidence rate of clinical mastitis on Canadian dairy farms. J. Dairy Sci. 91, 1366–1377. https://doi.org/10.3168/jds.2007-0757.

Oliver, S.P., Murinda, S.E., Jayarao Bhushan, M., 2011. Impact of antibiotic use in adult dairy cows on antimicrobial resistance of veterinary and human pathogens: a comprehensive review. Foodborne Pathog. Dis. 8, 337–355. https://doi.org/10.1089/fpd.2010.0730.

Oliver, J.P., Gooch, C.A., Lansing, S., Schueler, J., Hurst, J.J., Sassoubre, L., Crossette, E. M., Aga, D.S., 2020. Invited review: fate of antibiotic residues, antibiotic-resistant bacteria, and antibiotic resistance genes in US dairy manure management systems. J. Dairy Sci. 103, 1051–1071. https://doi.org/10.3168/jds.2019-16778.

Omwenga, I., Aboge, G.O., Mitema, E.S., Obiero, G., Ngaywa, C., Ngwili, N., Wamwere, G., Wainaina, M., Bett, B., 2021. Antimicrobial usage and detection of multidrug-resistant Staphylococcus aureus, including methicillin-resistant strains in raw milk of livestock from northern Kenya. Microb. Drug Resist. 27, 843–854. https://doi.org/10.1089/mdr.2020.0252.

O'Reilly, K.M., Green, M.J., Peeler, E.J., Fitzpatrick, J.L., Green, L.E., 2006. Investigation of risk factors for clinical mastitis in British dairy herds with bulk milk somatic cell counts less than 150,000 cells/ml. Vet. Rec. 158, 649–653. https://doi.org/10.1136/vr.158.19.649.

Pantoja, J.C.F., Hulland, C., Ruegg, P.L., 2009. Dynamics of somatic cell counts and intramammary infections across the dry period. Prev. Vet. Med. 90, 43–54. https://doi.org/10.1016/j.prevetmed.2009.03.012.

Peeler, E.J., Green, M.J., Fitzpatrick, J.L., Green, L.E., 2002. Study of clinical mastitis in British dairy herds with bulk milk somatic cell counts less than 150,000 cells/ml. Vet. Rec. 151, 170–176. https://doi.org/10.1136/vr.151.6.170.

Pol, M., Ruegg, P.L., 2007. Relationship between antimicrobial drug usage and antimicrobial susceptibility of gram-positive mastitis pathogens. J. Dairy Sci. 90, 262–273. https://doi.org/10.3168/jds.S0022-0302(07)72627-9.

Regula, G., Danuser, J., Spycher, B., Wechsler, B., 2004. Health and welfare of dairy cows in different husbandry systems in Switzerland. Prev. Vet. Med. 66, 247–264. https://doi.org/10.1016/j.prevetmed.2004.09.004.

Robbins, J.A., Roberts, C., Weary, D.M., Franks, B., von Keyserlingk, M.A.G., 2019. Factors influencing public support for dairy tie stall housing in the U.S. PLoS One 14, e0216544. https://doi.org/10.1371/journal.pone.0216544.

Rubin, D.B., 1973. The use of matched sampling and regression adjustment to remove bias in observational studies. Biometrics 29, 185–203. https://doi.org/10.2307/2529685.

Rubin, D.B., Thomas, N., 2000. Combining propensity score matching with additional adjustments for prognostic covariates. J. Am. Stat. Assoc. 95, 573–585. https://doi.org/10.2307/2669400.

Schaeren, W., 2006. Antibiotikaverbrauch 2003 und 2004 in der Milchproduktion. Agrarforsch. Schweiz 13, 234–239.

Schukken, Y.H., Wilson, D.J., Welcome, F., Garrison-Tikofsky, L., Gonzalez, R.N., 2003. Monitoring udder health and milk quality using somatic cell counts. Vet. Res. 34, 579–596. https://doi.org/10.1051/vetres:2003028.

Sizemore, S., Alkurdi, R., 2019. Matching Methods for Causal Inference: A Machine Learning Update [WWW Document]. URL. https://humboldt-wi.github.io/blog/research/applied_predictive_modeling_19/matching_methods// (accessed 6.20.21).

Spycher, B., Regula, G., Wechsler, B., Danuser, J., 2002. Gesundheit und Wohlergehen von Milchkühen in verschiedenen Haltungsprogrammen. Schweiz. Arch. Für Tierheilkd. 144, 519–530. https://doi.org/10.1024/0036-7281.144.10.519.

Stuart, E.A., 2010. Matching methods for causal inference: a review and a look forward. Stat. Sci. 25, 1–21. https://doi.org/10.1214/09-STS313.

Swiss National Science Foundation, 2021. Providing the Right Incentives - Project Description [WWW Document]. Antimicrob. Resist. - Natl. Res. Programme, 72. URL. http://www.nfp72.ch/en/projects/module-3-optimised-use-of-antibiotics/providing-the-right-incentives (accessed 6.20.21).

Talebi Bezmin Abadi, A., Rizvanov, A.A, Haertlé, T., Blatt, N.L, 2019. World Health Organization Report: Current Crisis of Antibiotic Resistance. J. Bionanosci. 9, 778–788. https://doi.org/10.1007/s12668-019-00658-4.

Valde, J.P., Hird, D.W., Thurmond, M.C., Østerås, O., 1997. Comparison of ketosis, clinical mastitis, somatic cell count, and reproductive performance between free stall and tie stall barns in Norwegian dairy herds with automatic feeding. Acta Vet. Scand. 38, 181–192. https://doi.org/10.1186/BF03548498.

Valde, J.P., Østerås, O., Simensen, E., 2005. Description of herd level criteria for good and poor udder health in Norwegian dairy cows. J. Dairy Sci. 88, 86–92. https://doi.org/10.3168/jds.S0022-0302(05)72665-5.

van den Borne, B.H.P., van Schaik, G., Lam, T.J.G.M., Nielen, M., 2010. Variation in herd level mastitis indicators between primi- and multiparae in Dutch dairy herds. Prev. Vet. Med. 96, 49–55. https://doi.org/10.1016/j.prevetmed.2010.05.010.

Waldrop, M.E., Roosen, J., 2021. Consumer acceptance and willingness to pay for cow housing systems in eight European countries. Q Open 1. https://doi.org/10.1093/qopen/qoab001.

Wan, F., 2019. Matched or unmatched analyses with propensity-score–matched data? Stat. Med. 38, 289–300. https://doi.org/10.1002/sim.7976.
Wolf, C.A., Tonsor, G.T., McKendree, M.G.S., Thomson, D.U., Swanson, J.C., 2016. Public and farmer perceptions of dairy cattle welfare in the United States. J. Dairy Sci. 99, 5892–5903. https://doi.org/10.3168/jds.2015-10619.
Zeileis, A., 2004. Econometric computing with HC and HAC covariance matrix estimators. J. Stat. Softw. 11, 1–17. https://doi.org/10.18637/jss.v011.i10.
Zeileis, A., Hothorn, T., 2002. Diagnostic checking in regression relationships. R News 2, 7–10.