Journal of Food Protection 86 (2023) 100136

Contents lists available at ScienceDirect

Journal of Food Protection

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

Research Note

Predicting Foodborne Disease Outbreaks with Food Safety Certifications: Econometric and Machine Learning Analyses

Yuqing Zheng [1,*], Azucena Gracia [2,3], Lijiao Hu [4]

[1] Department of Agricultural Economics, University of Kentucky, 313 Charles E. Barnhart Bldg., Lexington, KY 40546, United States
[2] Unidad de Economía Agroalimentaria, Centro de Investigación y Tecnología Agroalimentaria de Aragón (CITA), Avda. Montañana, 930, Spain
[3] Instituto Agroalimentario de Aragón–IA2, CITA-Universidad de Zaragoza, 50059 Zaragoza, Spain
[4] Agricultural Business, College of Agriculture, California State University, Chico, United States

ARTICLE INFO

Keywords:
BRC
Food safety certification
Foodborne disease outbreaks
GlobalG.A.P.
ISO 22000

ABSTRACT

Since the late 1990s, food safety certification has emerged as a prominent and influential regulatory mechanism in both the private and public spheres of the contemporary agri-food system. Food safety standards protect consumers from foodborne illnesses and help producers avoid the massive economic losses associated with food safety breaches. We empirically examine the relationship between foodborne disease outbreaks and certification adoption by utilizing the data on food safety certification adoption in the United States and Europe from 2015 through 2020. In our regression models, food safety certification along with select economic variables such as gross domestic product are used to explain the number of illnesses caused by foodborne disease outbreaks. For the United States at the state level, we found that certifications to SQF, PrimusGFS, BRC, or FSSC 22000 are negatively associated with the number of foodborne illnesses. For the case of Europe at the country level, certifications to ISO 22000 or FSSC 22000 are negatively associated with the number of foodborne illnesses. We then proceed to use machine learning techniques to examine how well we can use food safety certification data to predict foodborne disease outbreaks. Applying several algorithms (ordinary least squares, multinomial, decision tree, and random forest) to the U.S. data, we found that our models with food safety certification adoption can predict the number of U.S. foodborne illnesses or deaths with a relatively high degree of precision (testing accuracy at around 70% or better). Feature importance analysis allows us to inspect the relative importance of each explanatory variable (or feature) for making accurate predictions of the illness or death numbers. Through ranking the importance of explanatory variables, our study reveals that certification information could be the second most important variable (after gross domestic product) contributing to explain foodborne disease outbreaks.

Foodborne diseases pose a significant global health burden. For example, 299 foodborne disease outbreaks occurred in the United States in 2020, causing 5,987 illnesses, 641 hospitalizations, and fourteen deaths (CDC, 2022). In Europe, 3,166 foodborne disease outbreaks occurred during the same period, resulting in 22,010 illnesses, 1,838 hospitalizations, and 48 deaths (EFSA, 2022). Since the late 1990s, food safety certification has emerged as a prominent and influential regulatory mechanism in both the private and public spheres of the contemporary agri-food system. Food safety standards protect consumers from foodborne illnesses and help producers avoid the massive economic losses associated with food safety breaches. Faced with an increasing number of food recalls, many food retailers (e.g., Wal-Mart, Target) now demand that their suppliers obtain food safety certification. Government agencies have also adopted food safety certification procedures, e.g., the Food and Drug Administration's (FDA) new requirement of a credible food safety certification on high-risk imported foods.

The objective of this paper was twofold. First, by utilizing the recently compiled data on food safety certification adoption in the United States and Europe, we empirically investigate the association between foodborne disease outbreaks and the adoption of food safety certifications in these two regions separately. Food safety certifications intend to provide assurance to food safety.

Our research could provide the first insight into an important question of whether the use of food safety certification is associated with fewer disease outbreaks.

* Corresponding author.
E-mail addresses: yuqing.zheng@uky.edu (Y. Zheng), agracia@cita-aragon.es (A. Gracia), lhu@csuchico.edu (L. Hu).

https://doi.org/10.1016/j.jfp.2023.100136
Received 26 May 2023; Accepted 24 July 2023
Available online 27 July 2023

Second, we use machine learning techniques to examine how well we can use food safety certification data to predict foodborne disease outbreaks. It is difficult to predict disease outbreaks. However, food safety certifications have been increasingly popular among food producers worldwide in the last two decades, displaying tremendous variation in the degree of adoption among and within countries. For example, for the year 2020 (based on our own data), the number of sites certified to ISO 22000, a major (private) food safety standard, ranges from a few in Luxembourg and Malta to 929 and 2,069 in Italy and Greece, respectively. For the same year in the United States, the number of sites certified to PrimusGFS, the leading food safety standard in the United States by number (Hu et al., 2022), ranges from almost none in states such as Connecticut, New Hampshire, and Kansas to 1,261 in Arizona and 6,478 in California. We explore whether we can utilize such large variations across regions and over time to help predict foodborne disease outbreaks. In particular, we use several algorithms (decision tree and random forest) to assess how well these certifications can be used for prediction purpose and then compare the importance (feature importance) of certifications against other conventional economic control variables such as the gross domestic product.

Our study is closely related to three strands of literature. This first is a small but growing literature on private food safety certifications (Bovay, 2022; Hu et al., 2022; Rao et al., 2021), which can be divided into European-based ones (British Retail Consortium [BRC], Global Good Agricultural Practices [GlobalG.A.P.], Food Safety System Certification [FSSC] 22000, International Organization for Standardization [ISO] 22000, International Featured Standards [IFS]) and U.S.-based ones (Safe Quality Food [SQF] and PrimusGFS, Table 1 in Hu et al. (2022)). Rao et al. (2021) reviewed the 45 articles that discuss the European standards and found that these retailers initiated standards to change the agri-food supply chain in a way to allow the retailers to exert more influence without taking on additional legal and economic liability. Hu et al. (2022) provided an overview and conceptual framework of the private food safety certification market and reviewed 34 empirical studies devoted to this topic. They found that almost all of these studies focus either on the determinants of the adoption of certification or the impact of certification on farmers (e.g., financial performance, price, and quantity), firms (e.g., firm productivity and worker's welfare), industries (e.g., vertical integration), and exports (e.g., volume and value). Our research falls under the broad category of certification's impact but aims to enrich the literature by extending the outcome measure to disease outbreaks at the aggregate level.

The second related literature includes economics studies on foodborne diseases. A search of the keywords "foodborne disease" in the economics literature database EconLit yields about two dozens of published studies. The majority of these studies tend to examine consumers' risk perception of and response to foodborne disease outbreaks and the associated costs (Meagher, 2022; Roberts & Foegeding, 1991; Roberts & Marks, 2019; Shan et al., 2019; Sundström, 2018). For example, Sundström (2018) estimated the total cost for the five major foodborne illnesses (*Campylobacter*, *Salmonella*, *E. coli*, *Yersiniosis*, and *Shigellosis*) amount to 142 million euros a year for the case of Sweden. Consumers generally responded to news or advice about food contamination (Arnade et al., 2013). In the specific case of consumers in New York State, Zheng and Kaiser (2009) found that an additional person sickened due to the ingestion of tainted cheese products at home decreased per capita milk demand in New York State by 0.13 percent (or 0.07 pounds).

Several other studies in this literature discuss strategies against the outbreaks. Himmler et al. (2020) showed that most consumers surveyed in the United Kingdom, Denmark, Germany, Hungary, Italy, and the Netherlands were willing to pay for an early warning system (an increase in safety) for foodborne disease outbreaks. More recently, Anderson et al. (2022) examined the effects of many U.S. municipal-level efforts that were viewed important in fighting against food- and water-borne diseases. They found that water filtration was associated with a significant reduction in infant mortality during the period 1900–1940.

The last related strand is machine learning applications in the fields of economics, agricultural economics, and public health. Machine learning revolves around the problem of prediction and has found its own place in the statistical and econometric toolbox (Mullainathan & Spiess, 2017; Varian, 2014). For example, Hut and Oster (2022) used machine learning to predict households with significant dietary changes and found dietary concentration is a significant predictor of change, while demographics have little predictive power. Machine learning has also seen further adoption by agricultural economists, though applications seem limited. Empirical examples include predicting agricultural trade patterns (Gopinath et al., 2020), predicting hog inventory in China (Shao et al., 2021), estimating livestock transfer effect in Guatemala (Mullally et al., 2021), mapping croplands (Jia et al., 2019), etc. Storm et al. (2020) provide an excellent review of machine learning in agricultural and applied economics. As to public health, Wang et al. (2021) used machine learning methods to classify foodborne disease pathogens in China. They visually analyze several features of foodborne diseases, such as space, time, and exposed food to disease. They found the prediction accuracy approaches 69% in identifying pathogens (using the gradient boost decision tree model). In a similar fashion, Zhang et al. (2021) used features such as case

Table 1
Summary statistics for the United States (unit: state)

	Definition	Sample size	Mean	S.D.	Min.	Max.
Illnesses	Number of foodborne illnesses (cases) in a state	266	251.87	284.10	4 (Delaware[d])	1,537 (California)
Death	Number of death in a state	266	0.19	0.52	0	3
GDP	State real gross domestic product, billion dollars	266	390.09	473.22	29.12	2,739.34
FarmInc	Farm income, million dollars	266	1.68	2.78	−0.32	20.61
FoodEmp	Food manufacturing employment, thousand	266	36.00	31.79	0.93	174.65
SQF[a]	Number of sites certified to SQF	266	128.39	132.27	3 (Wyoming)	764 (California)
GFS	Number of sites certified to PrimusGFS	266	241.31	844.96	0 (Wyoming)	7,139 (California)
BRC[b]	Number of sites certified to BRC	266	50.00	52.13	2 (Vermont)	321 (California)
Enforcement	Number of USDA enforcement actions	266	21.48	34.16	0 (Kentucky)	238 (New York)
GAP[c]	Number of sites certified to USDA GAP	173	67.72	89.28	0 (Utah)	434 (Washington)
GlobalG.A.P.	Number of sites certified to GlobalG.A.P.	173	71.43	253.51	0 (Kansas)	1,780 (California)
FSSC22000	Number of sites certified to FSSC 22000	173	22.14	22.72	0 (Maine)	118 (California)

[a] Safe Quality Food standard.
[b] British Retail Consortium standard.
[c] USDA good agriculture practice.
[d] Many states have the same minimum values, e.g., Delaware and Wyoming both had 3 SQF adoptions in a year so we just list one state here.

information, exposure information, symptoms, and diagnosis results to classify whether a suspected foodborne disease outbreak is an actual outbreak, based on data in China.

Overall, there seems a literature gap connecting food safety certification (which could be an important feature in predicting foodborne disease outbreaks) and foodborne disease outbreaks, especially for countries outside China. In particular, in an earlier study published in this journal, Crandall et al. (2017) surveyed thousands of food producers worldwide (mostly in North America and Europe) and reported almost 90% of the certified supplies perceived food safety certification were beneficial for addressing their food safety concerns. About one-fifth of the certified companies reported a decrease in the number of food safety recalls. We aim to fill this void by utilizing the most recently collected data on food safety certifications for the United States and Europe.

Materials and methods

Data. We describe the data used in the paper. For the United States, we obtained the numbers of foodborne illnesses (cases) and death in a state respectively from the National Outbreak Reporting System (NORS) Dashboard, Centers for Disease Control and Prevention (CDC, 2022). These are state-level data from 2015 through 2020, in annual intervals. The summary statistics are presented in Table 1. The number of illnesses ranges from four in Delaware to 1,537 in California displaying large variations. For the features (independent variables to be shown in the following model section), we have seven food safety-related features and three economic control variables.

The food safety-related features include the number of sites in a state certified to SQF, PrimusGFS, BRC (available for 2015–2020 for these three), U. S. Department of Agriculture GAP (USDA GAP or GAP for short), GlobalG.A.P., and FSSC22000 (available only for 2016, 2018–2020 for the latter three). We collected the certification data directly from standard holders' websites over the years, as most standard holders provide current certification on their websites to the public. We did not include certification information for ISO 22000 or IFS because ISO 22000 adoption in the United States has been close to nothing, and IFS does not disclose certification information. Therefore, we have certification adoption information for five of the seven major private food safety standards plus one government standard (USDA GAP). The certification data display enormous variation across states in terms of the degree of adoption.

The last feature of food safety measure applies only to the establishments under USDA's regulation, i.e., meat, poultry, and egg product producers. The USDA files a non-compliance if an establishment fails to meet any regulatory requirement stipulated under the Hazard Analysis and Critical Control (HACCP) system, sanitation standard operating procedures, and sanitation performance standards. When there are two or more non-compliances filed on an establishment, enforcement actions will take place. Enforcement actions are in the forms of regulatory control action, withholding, or suspension. We, therefore, use the number of enforcement actions taken in a state as a measure for the food safety practice in that state. The data source is Food Safety and Inspection Service's (FSIS) Quarterly Enforcement Report by the USDA (USDA, 2022).

We selected three variables for economic controls, which are the state gross domestic product (GDP), farm income, and food manufacturing employment. The data source is the U.S. Bureau of Economic Analysis (BEA, 2022). These variables could capture the impacts of the scale of the states and the associated agricultural/food sectors.

For European data, we obtained the disease outbreak data from the European Food Safety Authority Dashboard (EFSA, 2022) for 2015 through 2020 (country-level data in annual intervals). This dashboard reports foodborne outbreaks collected by the authority from the European Union member states and some other countries. Our data include the number of illnesses (cases) in a country caused by foodborne diseases (by all pathogens), covering Austria, Belgium, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, and United Kingdom.

Similar to the U.S. case, we also include country-level GDP, agricultural values, and agricultural labor forces as the economic control variables (EuroStats, 2022). For food safety standards, we obtained the country-level adoption to GlobalG.A.P., ISO 22000, and FSSC 22000. We collected the GlobalG.A.P. data by contacting the GlobalG.A.P. representative in Spain directly. These are the number of GlobalG.A.P. certificates in each country covering fruits, vegetables, and livestock (mainly consisting of the fruit and vegetable certifications). The FSSC data were collected by us over the years, which only cover the years from 2017 through 2020. The ISO data were obtained from the ISO Surveys on Certification (ISO, 2022). Table 2 presents the summary statistics for the European data. Note due to data availability, the GlobalG.A.P. is measured by certificates and ISO and FSSC are measured by sites, and as noted by the GlobalG.A.P. contact, a certificate could cover several sites.

Econometric model. For the United States, our empirical model is as follows:

$$\begin{aligned}\mathrm{Illness}_{st} = {} & \beta_0 + \beta_1 SQF_{st} + \beta_2 GFS_{st} + \beta_3 BRC_{st} + \beta_4 ENF_{st} + \beta_5 GAP_{st} \\ & + \beta_6 GlobalGAP_{st} + \beta_7 FSSC_{st} + \beta_8 GDP_{st} + \beta_9 FarmInc_{st} \\ & + \beta_{10} FoodEmp_{st} + \varnothing_s + \alpha_t + \varepsilon_{st} \qquad (1)\end{aligned}$$

where subscripts s and t stand for state and year, respectively, $\beta_0 - \beta_{10}$ are parameters to be estimated, $\varnothing_s$ and α_t are state and yearly fixed effects (dummy variables), respectively, and ε_{st} is the error term. Therefore, the number of illnesses caused by foodborne diseases in a state is modeled as a function of state certification adoption to the six food safety standards (SQF, PrimusGFS, BRC, GAP, GlobalG.A.P., and FSSC22000), federal government food safety enforcement action in a state (ENF), three state economic control variables (GDP, FarmInc, and FoodEmp), and fixed effects. We expect that the number of illnesses is negatively correlated with food safety certification adoptions.

The European model is first specified as:

$$\begin{aligned}\mathrm{Illness}_{ct} = {} & \gamma_0 + \gamma_1 GlobalGAP_{ct} + \gamma_2 ISO_{ct} + \gamma_3 FSSC_{ct} + \gamma_4 GDP_{ct} \\ & + \gamma_5 AgValue_{ct} + \gamma_6 AgLabor_{ct} + \mu_r + \omega_t + \theta_{ct} \qquad (2)\end{aligned}$$

where the subscripts c stands for country, t still stands for year, $\gamma_0 - \gamma_6$ are parameters to be estimated, μ_r and ω_t are region (Central and Eastern, Northern, Southern, and Western Europe) and yearly fixed effects, respectively, and θ_{ct} is the error term. The number of illnesses caused by foodborne diseases in a country is modeled as a function of country adoption to three food safety standards that we have data (GlobalG.A.P., ISO 22000, and FSSC), three country-wise economic control variables (GDP, AgValue, and AgLabor), and fixed effects.

To utilize the more granular pathogen-level data (and increase the sample size), we also estimate a variant of equation (2) as follows:

$$\begin{aligned}\mathrm{Illness}_{pct} = {} & \gamma_0 + \gamma_1 GlobalGAP_{ct} + \gamma_2 ISO_{ct} + \gamma_3 FSSC_{ct} + \gamma_4 GDP_{ct} \\ & + \gamma_5 AgValue_{ct} + \gamma_6 AgLabor_{ct} + \mu_r + \omega_t + \vartheta_p + \theta_{pct} \qquad (3)\end{aligned}$$

where p (*Salmonella*, norovirus, bacterial, *Campylobacter*, and other pathogens combined) indexes the four leading pathogens that cause foodborne diseases and other types combined, ϑ_p is the pathogen fixed effects.

Statistical methods. For the above linear regression models specified in equations (1)–(3), we use the ordinary least squares (OLS) method. All models use the robust standard errors to account for potential heterogeneity in the error term. For the machine learning analysis, we focus on the U.S. data (to conserve space) to conduct machine learning from two perspectives: prediction and feature impor-

Table 2
Summary statistics for European countries (unit: country)

	Definition	Sample Size	Mean	S.D.	Min.	Max.
Illnesses	Number of foodborne illnesses (cases) in a country[a]	178	1607.51	2655.87	2 (Luxembourg)	15,677 (France)
AgVaue	Real agricultural values, million euros	178	7,517.58	10,375.65	39.00	44,051.62
AgLabor	Agricultural labor force, thousands	178	315.84	435.43	3.43	1,937.10
GDP	Country real gross domestic product, billion euros	178	506.52	730.11	9.17 (Malta)	2,987.19 (Germany)
GlobalG.A.P.	Number of GlobalG.A.P. certificates	178	750.86	1,253.64	0 (Cyprus, Estonia)	4,397 (Netherlands)
ISO22000	Number of sites certified to ISO 22000	178	282	397.40	3 (Luxembourg, Malta)	2,285 (Greece)
FSSC22000	Number of sites certified to FSSC 22000	119	171.74	176.98	1 (Luxembourg)	870 (Netherlands)

[a] Countries include Austria, Belgium, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, and the United Kingdom.

tance. Machine learning by definition is programming computers to optimize a performance criterion using example data or past experience. With a model defined with some parameters (regression coefficients to be estimated plus additional parameters such as weights), learning optimizes the parameters of the model using the training data. Then, the trained model intends to make predictions in the future, or descriptive to gain knowledge from data, or both. Specifically, in the first step, we randomly distribute 80% of the U.S. dataset (respectively for the illness and death equations [note in both equations we model the numbers, not if death or illness occurred]) into training set and the remaining 20% into test set (as a standard practice), and train the regression model on the training dataset. For the illness equation, the data to be split include the number of illnesses, certifications, and economic variables. Separately for the death equation, the data to be split include the number of deaths, certifications, and economic variables. In the next step, we predict the test set result based on several criteria including the R^2 measuring the goodness of model fit (proportion of variance in the dependent variable that can be explained by the independent variables).

We also apply k-fold cross-validation (fivefold in our case) to all three algorithms to obtain the train accuracy, test accuracy, and associated standard errors for test accuracy. The accuracy is the k-fold cross-validation average accuracy. The accuracy measure represents the share of the total outcome variance explained by the model, as such it is closely similar to an adjusted R^2. Some important hyperparameter values used are k equals 5, seed equals 3, number of estimators equals 30, and tree depth is set at 12. The prediction part is achieved by utilizing the r_ml_stata_cv and c_ml_stata_cv (for continuous and limited dependent variables, respectively) packages written for Stata. It uses the Stata/Python integration capability of Stata 16 and above to implement a variety of regression algorithms including the OLS, decision trees, random forest, neural network, etc. K-fold cross-validation is also available in this package. We present the prediction results based on three algorithms: OLS, decision tree, and random forest. These are supervised machine learning algorithms. Tree can capture non-linear relationships (Storm et al., 2020). Random forest combines the results of multiple trees and automatically detects interactions to improve prediction (Hut & Oster, 2022; Storm et al., 2020). All algorithms for the illness machine learning use the variables contained in equation (1). For example, the OLS algorithm uses the exact model specification in equation (1). A decision tree model builds on equation (1) but combines some decision, e.g., adopting certification, and if so, which types. A random forest model combines several decision trees and allows variables in equation (1) to interact.

As to feature importance, we programmed directly in Python 3.1 using the RandomForestRegressor and its associated feature importance function. Our objective is to provide a visual inspection regarding the importance of the food safety-related variables versus economic control variables in predicting disease outbreaks.

Results and discussion

Regression models. We first report the regression results for equation (1) in Table 3. Column (1) presents the model with the full data for 2015–2020 but with less standard (sample size 266) and Column (2) includes results for the years 2016, 2018–2020 where all standards are included (sample size 173). Both models show a strong R^2 at around 0.8. For specification (1), we also tried a model that only includes the first four variables (SQF, GFS, BRC, and Enforcement, results not reported here), and the R^2 reaches 0.78. That is, the food safety-related variables can provide a high prediction power for the foodborne disease outbreaks.

Focusing on the full specification in column (2), which includes all standards, we found that the coefficients are negative and statistically significant at 10% or better for four of the safety standards (SQF, GFS, BRC, and FSSC 22000). For example, given the estimated coefficient for PrimusGFS is -0.099, ten state-level additional adoption to the PrimusGFS standard is associated with one illness reduction (-0.099x10) in that state, while the association for the other three standards is even much larger. Such results of the negative association are consistent with our a priori expectation. All the three economic control variables were found statistically significant at the 5% level or better, also displaying the importance of accounting for the size of the economy. For example, the state with a higher GDP or food employment is found to have more foodborne illnesses. We also run the full specification for the number of deaths, applying the same model specified in equation (1) but using the tobit estimator. Column (3) of Table 3 shows, in this case, only the PrimusGFS and FSSC 22000 standards remain statistically significant at the 10% level or better (and negative).

For the European model (reported in Table 4), we first present the results combining all pathogens (equation [2]), with and without FSSC 22000. The R^2 stays above 0.6 in both specifications. In particular, we found that disease outbreaks are negatively associated with ISO 22000 and FSSC 22000. Ten country-level additional adoption to the ISO 22000 standard is associated with eight illness reduction in that country, and that association with FSSC 22000 is three times larger. Surprisingly, we fail to find any statistically significant impact for the GlobalG.A.P. certification, considering its large popularity in Europe. Columns (3) and (4) present the results corresponding to equation (3), the model that breaks down pathogen types and increases the sample size significantly (fivefold). The results corroborate earlier findings reported in columns (1) and (2).

Machine learning. In Table 5, we present the results for predicting the number of foodborne disease-caused illnesses, based on the variables used in equation (1). We present the results covering the whole 2015–2020 period but with less standards (upper panel), and the results using four years of data (2016, 2018–2020) but with the full six standards (lower panel). These two panels show similar results, and we focus on interpreting the results using all standards. We found

Table 3
Regression results for the United States

	Dependent Variable: Number of Illness		Dependent Variable: Number of Deaths
	(1) 2015–2020	(2) 2016, 2018–2020, More Standards	(3) 2016, 2018–2020, More Standards
SQF (β_1)	−0.34	−1.989*	0.001
	(0.79)	(1.02)	(0.00)
GFS (β_2)	−0.169**[a]	−0.099**	−0.001***
	(0.07)[b]	(0.04)	(0.0002)
BRC (β_3)	−1.398	−5.580*	0.007
	(2.97)	(3.19)	(0.01)
Enforcement (β_4)	−3.440***	−0.825	−0.003
	(1.19)	(1.51)	(0.003)
Gap (β_5)		0.57	−0.001
		(0.53)	(0.001)
GlobalG.A.P. (β_6)		0.175	0.001
		(0.21)	(0.001)
FSSC22000 (β_7)		−4.974*	−0.009*
		(2.92)	(0.005)
GDP (β_8)	−0.215	1.748**	−0.009***
	(0.79)	(0.73)	(0.002)
FarmInc (β_9)	−33.296*	−53.374**	0.030
	(19.66)	(21.27)	(0.07)
FoodEmp (β_{10})	16.966	39.121***	0.019
	(10.47)	(13.16)	(0.03)
R^2	0.790	0.804	0.42
Adjusted R^2	0.730	0.699	
Log-likelihood	−1672.25	−1072.00	−68.99
Sample size	266	173	173

[a] *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$. State and yearly fixed effects are included in all specifications.
[b] Robust standard errors are in parentheses.

Table 4
Regression results for Europe

Dependent variable: Number of illnesses	All pathogens combined		Broken down by pathogens	
	(1) 2015–2020	(2) 2017–2020	(3) 2015–2020	(4) 2015–2020
GlobalG.A.P.(γ_1)	−0.173	0.099	−0.035	0.02
	(0.19)	(0.23)	(0.05)	(0.05)
ISO22000 (γ_2)	−1.165***[a]	−0.780*	−0.233***	−0.156**
	(0.38)[b]	(0.46)	(0.06)	(0.07)
FSSC22000 (γ_3)		−2.527*		−0.505*
		1.38		0.29
GDP (γ_4)	−1.750***	−1.551***	−0.350***	−0.310**
	(0.45)	(0.53)	(0.12)	(0.12)
Ag. Value (γ_5)	0.339***	0.329***	0.068***	0.066***
	(0.07)	(0.09)	(0.02)	(0.02)
Ag. Labor (γ_6)	−0.619	−0.941	−0.124	−0.188
	(0.85)	(1.02)	(0.14)	(0.16)
R^2	0.618	0.659	0.315	0.352
Adjusted R^2	0.587	0.62	0.302	0.334
Log-likelihood	−1569.95	−1039.83	−7118.65	−4693.07
Sample size	178	119	890	595

[a] *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$. Region and yearly fixed effects are included in all specifications. Columns (3) and (4) also include pathogen fixed effects.
[b] Robust standard errors are in parentheses.

with the base OLS model, the training and testing accuracies are 88% and 73%, respectively. These numbers improve to 99% and 74% for the decision tree algorithm and 97% and 76% for the random forest model, at the cost of higher standard errors.

In Table 6, we present the results for predicting the number of foodborne disease-caused deaths. As the summary statistics in Table 1 show, the deaths range from zero to three, thus making this prediction largely a classification issue. In this case, the baseline is a multinomial model (available in the c_ml_stata_cv) instead of an OLS model, which produces a training accuracy of 89% and testing accuracy of 81% (for either model with three or six standards). For the model with three standards, the decision tree model improves the testing accuracy over the multinomial model to 85%, with a much smaller standard error (1% vs. 14%). The training accuracy for the random forest model is 80%, while its training accuracy is perfect. For the model with six standards, the results are similar. While the decision tree model and random forest model produce a higher training accuracy, the testing accuracy is similar to that of the multinomial model. While the standard errors of testing accuracy are lower for the decision tree and random forest models, their classification error rate is higher for testing.

Finally, we turn to the results of the feature importance. In Figure 1, we present the results on feature importance based on random forest for the U.S. model. This allows us to visually inspect the relative importance of each attribute (feature) for making accurate predictions. Python (Scikit-Learn package) measures the feature importance by examining how much the tree nodes that use a feature reduce impurity

Table 5
Machine learning results for the number of illnesses in the United States

Methods	Training Accuracy	Testing Accuracy	Standard Errors (Testing Accuracy)
2015–2020, 3 Standards			
OLS	85%	60%	28%
Decision Tree	84%	63%	38%
Random Forest	97%	76%	31%
2016, 2018–2020, 6 Standards			
OLS	88%	73%	20%
Decision Tree	99%	74%	36%
Random Forest	97%	76%	31%

Table 6
Machine learning results for the number of deaths in the United States

Methods	Training Accuracy	Testing Accuracy	Standard Errors (Testing Accuracy)	Classification Error Rate	
				Training	Testing
2015–2020, 3 Standards					
Multinomial	89%	81%	14%	43%	8%
Decision Tree	87%	85%	1%	14%	17%
Random Forest	100%	80%	10%	35%	17%
2016, 2018–2020, 6 Standards					
Multinomial	89%	81%	14%	12%	8%
Decision Tree	92%	81%	12%	9%	19%
Random Forest	100%	80%	10%	0%	17%

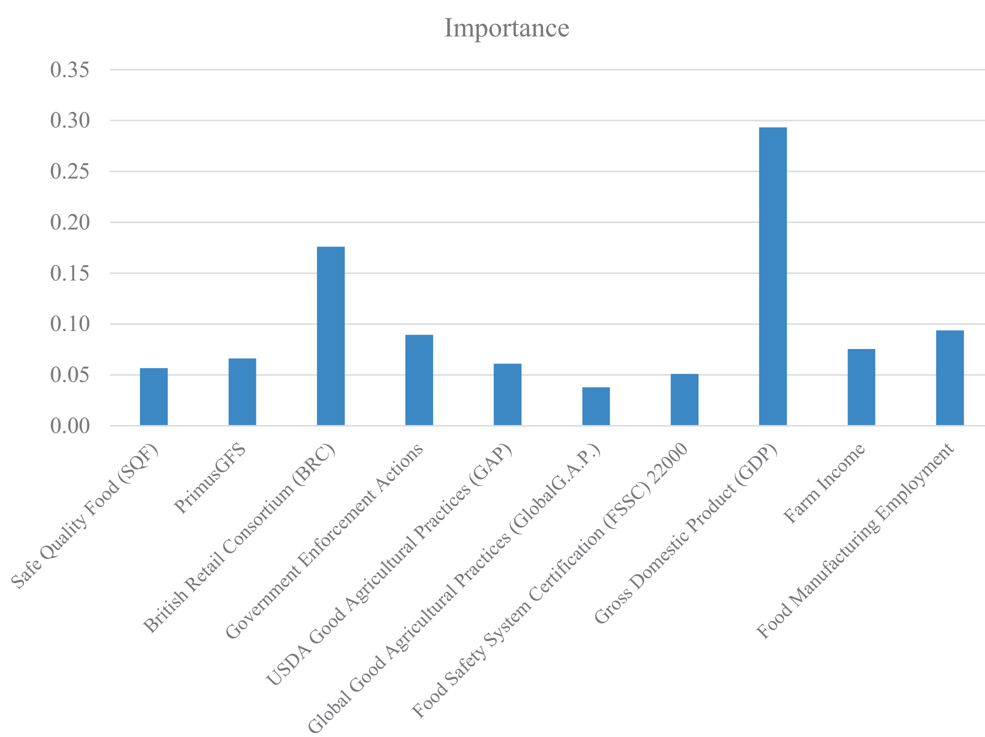

Figure 1. Feature importance (U.S. Model). Obtained through random forest in python.

across all trees in the forest on average (Géron, 2022). The feature importance in Figure 1 is scaled so that the sum of all importance is one.

Figure 1 illustrates the feature importance from the highest to the lowest is GDP (29%), BRC (18%), food employment (9%), enforcement (9%), farm income (8%), GFS (7%), GAP (6%), SQF (6%), FSSC 22000 (5%), and GlobalG.A.P. (4%). It is not surprising to see GDP ranks the highest among all the features. However, it is worth noting the adoption to a certain food safety standard, namely BRC, can rank second in the importance in predicting food safety disease outbreaks, with a feature importance index much higher than other economic variables such as employment.

Overall, several results emerge from our empirical analyses, with potentially important policy implications. First, we found a negative association between food safety certification association and foodborne disease outbreaks. In particular, we found that certifications to SQF, PrimusGFS, BRC, or FSSC 22000 are negatively associated with the number of illnesses related to foodborne diseases in the United States. For the case of Europe, we found that certifications to ISO 22000 or FSSC 22000 are negatively associated with the number of illnesses related to foodborne diseases.

Second, through machine learning (mainly decision tree and random forest algorithms), we found that our models with food safety certification adoption can predict the U.S. state-level illnesses caused by foodborne diseases with a relatively high degree of precision (testing accuracy at around 75%). Such testing accuracy reaches over 80% for predicting the number of deaths (U.S. state level) caused by foodborne diseases, which is a classification problem. Through further analysis of feature importance, we found that certification information (i.e., BRC adoption) could be the second most important variable (after GDP) contributing to explain foodborne disease outbreaks. These results highlight the potential importance of utilizing certification data in future government efforts to monitor or predict foodborne disease outbreaks.

Though our results provide the first preliminary evidence that food safety certification could potentially promote food safety, we emphasize association and do not intend to claim causality in our study design. This is the first limitation of the study. Second, we predict the test set result not future disease outbreaks. Third, our results are confined to the United States and Europe and may not extend to other countries.

Funding and competing interests

This study was funded by USDA National Institute of Food and Agriculture under the Grant/Award Number No. 1011683. The authors have no other relevant financial or non-financial interests to disclose.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Anderson, D. M., Charles, K. K., & Rees, D. I. (2022). Reexamining the contribution of public health efforts to the decline in urban mortality. *American Economic Journal: Applied Economics, 14*(2), 126–157. https://doi.org/10.1257/app.20190034.

Arnade, C., Kuchler, F., & Calvin, L. (2013). Consumers' response when regulators are uncertain about the source of foodborne illness. *Journal of Consumer Policy, 36*(1), 17–36. https://doi.org/10.1007/s10603-012-9217-6.

BEA. (2022). Bureau of Economic Analysi. https://www.bea.gov/data

Bovay, J. (2022). Food safety, reputation, and regulation. *Applied Economic Perspectives and Policy, 45*(2), 684–704. https://doi.org/10.1002/aepp.13315.

CDC. (2022). National Outbreak Reporting System (NORS) Dashboard, Centers for Disease Control and Prevention. https://wwwn.cdc.gov/norsdashboard/

Crandall, P. G., Mauromoustakos, A., O'Bryan, C. A., Thompson, K. C., Yiannas, F., Bridges, K., & Francois, C. (2017). Impact of the global food safety initiative on food safety worldwide: Statistical analysis of a survey of international food processors. *Journal of Food Protection, 80*(10), 1613–1622. https://doi.org/10.4315/0362-028X.JFP-16-481.

EFSA. (2022). European Food Safety Authority Dashboard. https://www.efsa.europa.eu/en/topics/topic/monitoring-foodborne-diseases

EuroStats. (2022). Eurostat - European Commission. https://ec.europa.eu/eurostat

Géron, A. (2022). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media, Inc.

Gopinath, M., Batarseh, F. A., & Beckman, J. (2020). *Machine learning in gravity models: An application to agricultural trade*, NBER working paper. https://www.nber.org/system/files/working_papers/w27151/w27151.pdf

Himmler, S., van Exel, J., Perry-Duxbury, M., & Brouwer, W. (2020). Willingness to pay for an early warning system for infectious diseases. *The European Journal of Health Economics, 21*(5), 763–773. https://doi.org/10.1007/s10198-020-01171-2.

Hu, L., Zheng, Y., Woods, T. A., Kusunose, Y., & Buck, S. (2022). The market for private food safety certifications: Conceptual framework, review, and future research directions. *Applied Economic Perspectives and Policy, 45*(1), 197–220. https://doi.org/10.1002/aepp.13226.

Hut, S., & Oster, E. (2022). Changes in household diet: Determinants and predictability. *Journal of Public Economics, 208*. https://doi.org/10.1016/j.jpubeco.2022.104620 104620.

ISO. (2022). ISO Surveys on Certification. https://committee.iso.org/sites/jtc1sc40/home/news/content-left-area/news-and-updates/iso-survey-of-certifications.html

Jia, X., Khandelwal, A., Mulla, D. J., Pardey, P. G., & Kumar, V. (2019). Bringing automated, remote-sensed, machine learning methods to monitoring crop landscapes at scale. *Agricultural Economics, 50*, 41–50. https://doi.org/10.1111/agec.12531.

Meagher, K. D. (2022). Policy responses to foodborne disease outbreaks in the United States and Germany. *Agriculture and Human Values, 39*(1), 233–248. https://doi.org/10.1007/s10460-021-10243-9.

Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives, 31*(2), 87–106. https://doi.org/10.1257/jep.31.2.87.

Mullally, C., Rivas, M., & McArthur, T. (2021). Using machine learning to estimate the heterogeneous effects of livestock transfers. *American Journal of Agricultural Economics, 103*(3), 1058–1081. https://doi.org/10.1111/ajae.12194.

Rao, M., Bast, A., & De Boer, A. (2021). European private food safety standards in global agri-food supply chains: A systematic review. *International Food and Agribusiness Management Review, 24*(5), 739–754. https://doi.org/10.22434/IFAMR2020.0146.

Roberts, T., & Foegeding, P. M. (1991). In *Risk assessment for estimating the economic costs of foodborne disease caused by microorganisms* (pp. 103–129). Springer https://doi.org/10.1007/978-94-011-7076-5_6.

Roberts, T., & Marks, S. (2019). In *Valuation by the cost of illness method: the social costs of Escherichia coli O157: H7 foodborne disease* (pp. 173–205). Routledge https://doi.org/10.4324/9780429267031.

Shan, L., Wang, S., Wu, L., & Tsai, F.-S. (2019). Cognitive biases of consumers' risk perception of foodborne diseases in China: Examining anchoring effect. *International Journal of Environmental Research and Public Health, 16*(13). https://doi.org/10.3390/ijerph16132268 2268.

Shao, Y., Xiong, T., Li, M., Hayes, D., Zhang, W., & Xie, W. (2021). China's Missing Pigs: Correcting China's Hog Inventory Data Using a Machine Learning Approach. *American Journal of Agricultural Economics, 103*(3), 1082–1098. https://doi.org/10.1111/ajae.12137.

Storm, H., Baylis, K., & Heckelei, T. (2020). Machine learning in agricultural and applied economics. *European Review of Agricultural Economics, 47*(3), 849–892. https://doi.org/10.1093/erae/jbz033.

Sundström, K. (2018). Cost of illness for five major foodborne illnesses and sequelae in Sweden. *Applied Health Economics and Health Policy, 16*(2), 243–257. https://doi.org/10.1007/s40258-017-0369-z.

USDA. (2022). Quarterly Enforcement Reports. https://www.fsis.usda.gov/inspection/regulatory-enforcement/quarterly-enforcement-reports

Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives, 28*(2), 3–28. https://doi.org/10.1257/jep.28.2.3.

Wang, H., Cui, W., Guo, Y., Du, Y., & Zhou, Y. (2021). Machine learning prediction of foodborne disease pathogens: Algorithm development and validation study. *JMIR Medical Informatics, 9*(1). https://doi.org/10.2196/24924 e24924.

Zhang, P., Cui, W., Wang, H., Du, Y., & Zhou, Y. (2021). High-efficiency machine learning method for identifying foodborne disease outbreaks and confounding factors. *Foodborne Pathogens and Disease, 18*(8), 590–598. https://doi.org/10.1089/fpd.2020.2913.

Zheng, Y., & Kaiser, H. M. (2009). Dairy-borne disease outbreak and milk demand: A study using outbreak surveillance data. *Agricultural and Resource Economics Review, 38*(3), 330–337. https://doi.org/10.1017/S1068280500009588.