Audit Grades in Food Safety Certification[*]

Yuqing Zheng[†] and Talia Bar[‡]

Abstract: Retailers increasingly require suppliers to certify for food safety standards. Third party certification bodies audit manufacturing sites to determine compliance with a standard. We use panel data on U.S. manufacturing sites certified for the British Retail Consortium global standard food safety program to identify determinants of certification audit grades. We find that firms with more resources obtain higher grades. Also grades increase with experience which likely indicates that the certification process improves food safety practices. Investigating the relationship between audit grades and the degree of competition between third party certification bodies in the vicinity of a manufacturing site, we find that higher competition is associated with higher audit grades.

Key words: Certifiers, certification bodies, food safety, standard, audit grade, British Retail Consortium, voluntary provision.

JEL classifications: L66, L15, L13, H89

---

[*] We are grateful to Antony Harrison, Directory Services Manager at the BRC global standards, and Samuel Crome, Senior Directory Network Administrator for providing data and for their insights about the program. We are indebted to Haim Bar, Steven Buck, and Craig Landry for their comments. We thank Brendan Costello and Joe Roessler for outstanding research assistance.

[†] Assistant Professor, Department of Agricultural Economics, University of Kentucky, yuqing.zheng@uky.edu

[‡] Assistant Professor, Department of Economics, University of Connecticut, talia.bar@uconn.edu

1. Introduction

Grades are commonly used as a way for intermediaries to convey information about performance, proficiency or quality from one side of a market to another. Notably students receive grades from their teachers in schools, colleges and universities, businesses obtain credit ratings, and athletes sometimes get grades from judges about their achievements. In the food industry, standard programs (e.g., standards of food safety or for organic production) set minimal requirements for safety and quality. Certifying to a standard allows producers to convey to buyers unobserved information about their products. Certification bodies serve as intermediaries, they audit food producers to determine if they can be certified for the standard. In some programs, including the British Retail Consortium (BRC) global standard program that we study in this paper, certified sites are assigned an audit grade that signals to potential buyers further information about compliance with the standard.[1]

In this paper we empirically study the determinants of audit grades for food safety certification, using panel data from the BRC global standard program, a leading international food safety standard for food manufacturers. BRC global standard is the third most popular international, private voluntary food safety standard program in the United States, following PrimusGFS and Safe Quality Food (SQF). Food safety standards are typically not advertised directly to individual consumers, but retailers might require their suppliers to obtain certification. The BRC program originates in the United Kingdom, and is available in more than 120 countries.[2] To obtain BRC certification, a manufacturing site needs to be audited by one of BRC's accredited certification bodies.

The main data source for our study is a dataset that we obtained from the BRC standard program. The data lists U.S. certified sites between 2011 and 2015, their addresses, the certification bodies that audited the sites, as well as the sites' audit grades. Sites need to renew their certification at least once per year, and so we have repeated observations for many of the sites. We combined the BRC data with data we collected about the BRC certification bodies, as well as some firm characteristics including the number of employees, sales and whether the firm

[1] Because food safety certification is a business-to-business standard, the grades are intended to be seen by potential institutional buyers, mainly, food retailers.

[2] See http://www.brcglobalstandards.com/.

is publically owned (for which we used Lexis Nexis data), and information on food recalls from the Food and Drug Administration (FDA). Using ordered probit specifications, we examine which factors affect the audit grades of BRC certified manufacturing sites.

The first cost of certification includes a modest fee of £185 to the BRC program paid each time a firm certifies. Additionally, the price and expenses are set by the certification bodies with individual manufacturing sites. These mainly depend on the labor rate determined by the certification bodies, the length of the audit, distance between a site and a certification body (site pays for travel cost of auditors), the size of the manufacturing facility, and the conditions of the existing food safety practice of the site. We do not observe prices in our data. An additional, potentially large component of the cost of certification, is the number of non-conformities that the auditor identifies and that the manufacturing site would need to fix before obtaining certification.

Passing audit grades take the values C, B or A, with A being the highest and also the most common grade.[3] We do not observe failing grades in our data, thus our analysis of determinants of grades is conditional on the site obtaining BRC certification. Audit grades depend on the number of non-conformities that the auditors detected. Following the audit, sites are given about a month to fix non-conformities identified in the audit before they obtain certification. A higher audit grade suggests less non-conformities would need to be fixed and thus the manufacturing site is expected to have less expenses to fix non-conformities. A low grade C also implies that the site would need to re-certify after 6 months instead of 12. Since audit grades are posted in the BRC directory available online, higher grades might also serve as positive signals to buyers.

Our analysis shows that producers who own multiple manufacturing sites tend to have higher audit grades. This might suggest that such larger scale producers have better food safety practices because they have more resources that allow them to comply with the standard, or because they care more about their reputation. However, another plausible reason why multisite businesses obtain higher grades could be that certification bodies are particularly interested in

[3] A very small number of A grade observations have a + sign after the grade (3 A+ in 2014 and 2015, respectively), indicating an unannounced audit, 190 audit grades are AA, which represents fewer than 5 non-conformities. But this AA grade was only instituted starting from 2015 when BRC updated to version 7. We treat A+ and AA grades as A grades.

attracting and keeping these businesses, and thus grade them more leniently. Producers who certify for the first time have significantly lower grades than more experienced ones. This likely suggests learning in the certification process and durable investments in quality and safety, so that producers are less likely to exhibit non-conformities after they became certified for the first time. A negative association between food recalls and grades seems to confirm that grades are indeed associated with safety and quality. Even after controlling for other observed determinants of grades, we find variations in average assigned grades between certification bodies, suggesting certification bodies are heterogeneous in their degree of audit intensity or grading leniency. This finding is consistent with a 2011 FDA study which found a wide variation in the performance of certification bodies for aquacultured shrimp producers. While we observe grade disparities in our data and a large share of A grades, during the period we study we do not see a clear grade inflation trend, which has been widely documented in the context of higher education.[4]

Dranove and Jin (2010) survey theoretical and empirical literature on quality disclosure and certification. They argue that third party certification bodies may have incentives to bias quality ratings, and that theoretically, the role of competition in determining the information content of ratings is ambiguous. In particular, they explain that "the presence of multiple certifiers encourages sellers to shop around, […]". We empirically examine the hypothesis that competition between certifiers affect grades. For each manufacturing site we define a measure of the degree of certifier competition around that site. We find that more intense competition is associated with higher audit grades.

Our paper contributes to the economic literature on certification and standards. Much of this literature has been theoretical.[5] Empirical analyses covered different markets, such as financial markets,[6] health care, and restaurants (see Dranove and Jin, 2010, for an excellent survey of this literature). Hubbard (1998) finds that the probability of a vehicle failing emissions inspection is slightly lower when the inspector has a close geographic competitor. More specifically, in the context of food safety certification, a number of authors offer insightful

[4] See for example Johnson (2006) and Bar et al. (2009).

[5] See for example Lizzeri (1999), Albano and Lizzeri (2001), Jahn et al. (2005), Baron (2011), Farhi et al. (2013), Podhorsky (2013) and Gaigné and Larue (2016).

[6] Cohen and Manuszak (2013) and Bae et al. (2015) examine the effects of competition on credit rating. There are three major credit rating agencies. These papers used the "Fitch's market share," which is the market share of the third largest credit rating company, as a measure of competition. While Cohen and Manuszak (2013) find a positive correlation between competition and ratings, Bae et al. (2015) do not.

discussions on food safety standards and third party certification. Henson and Caswell (1999) consider food safety regulations in developing countries. Hatanaka et al. (2005) and Fulponi (2006) discuss the role of retailers in the rise of voluntary food safety standards and third party certification. Tanner (2000) focuses on the Hazard Analysis Critical Control Point (HACCP) management system in food safety and Fagotto (2014) examines the pros and cons of private food safety standards. Only a few studies have offered empirical analyses of food safety certification. In one study, Anders et al. (2007) consider entry and competition in the market for certification bodies using data from the EurepGAP standard. Bar and Zheng (2016) use BRC data to examine the determinants of producers' choices of certification bodies. They provide evidence that food producers are attracted to "leniently grading" certification bodies and to ones that are geographically closer to their manufacturing site.

With the unique panel data of a popular food safety standard, our study offers insight into what determine audit grades and discusses possible policy implications based on our empirical findings. This paper is organized as follows: in Section 2 we present the data. In Section 3 we explain our empirical approach and results. Section 4 concludes.

2. Data

In our dataset there are 7,061 observations of 2,027 U.S. food manufacturing sites that certified to the BRC food standard and have expiration dates between May 2011 and December 2016.[7] Manufacturers need to renew their certification every year, and so for many of the manufactures we have repeated occurrences in our data. Each observation includes the manufacturing site name, address, product category (to be define later in this section), site owner, certification body that audited the site, dates of certification and expiration, and audit grade.

We augmented the (manufacturing site)×date level data that we obtained from the BRC program with information we collected from other data sources. We recorded whether a manufacturing site had a recall in the year of certification using FDA Archive for Recall, Market

[7] We do not include manufacturers located in Hawaii, Alaska, and Puerto Rico. Additionally, the data was given to us in one year batches. The first year was extracted in March, and later years were extracted as of November of each year. Therefore, there was a gap in the early period. Our results are robust to excluding the early "pre-gap" 2011 observations.

Withdrawals & Safety Alerts data. We searched for the producers in Lexis Nexis and recorded whether the site's owner is a publically traded company, the numbers of employees and revenues for the site owner when available. We used certification body websites to get the location of certification body offices. We used data from the census bureau and site's zip code to determine if the site is located in a metropolitan area (which has a population of at least 50,000).[8]

Using the data we have, we generated a few additional variables of interest. "Distance," is defined as the distance between each manufacturing site and the closest office of each certification body. "Multisites," is an indicator for whether there are other sites in the data owned by the same owner in the same year.[9] Multisites are likely larger more established companies (e.g., Tyson Foods and Archer Daniels Midland Company known as ADM) than those that only have one certified manufacturing site under that owner. Finally, "Competition" is a measure of the degree of competition among certification bodies in the vicinity of the manufacturing site. It is defined as the number of certification bodies that have an office that is within a certain distance *d* from the manufacturing site. We take *d* to be the average distance between sites and the closest office of their chosen certification bodies.

Table 1 provides descriptive statistics of the variables in our dataset. On a four-point scale, the average audit grade is 3.84. Of all site×date observations, 69% were multisites, 29% were certified for the first time, only about 1% had a food recall recorded at the FDA during the period of study. On average a site has 6.7 certification bodies that have an office at a site's vicinity. The distance *d*, defined in the previous paragraph to capture the average distance between a site and its chosen certification body, ranges from 559 miles in 2011 to 616 miles in 2015. The sites in our data were certified by eighteen different certification bodies with offices located in 24 states.

[8] See https://www.census.gov/geo/reference/gtc/gtc_cbsa.html.

[9] The way we defined multisites, a site could be a multisite in a certain year, but not in another year. Alternatively, we can define the multisites indicator as equal to 1 if a site shared an owner with another in any of the years in our sample. Our results are robust to this alternative definition.

Table 1. Summary Statistics for U.S. BRC Certified Sites (2011–2015)

Variable	Description	Mean	Standard Deviation	Min.	Max.
Grade	Audit grade (A = 4, B = 3, C = 2)	3.84	0.38	2	4
FirstTime	Equals 1 if certified for the first time	0.29	0.45	0	1
SecondTime	Equals 1 if certified for the second time	0.24	0.43	0	1
ThirdTime	Equals 1 if certified for the third time	0.21	0.40	0	1
FourthTime	Equals 1 if certified for the fourth time	0.16	0.37	0	1
FifthTime	Equals 1 if certified for the fifth time	0.10	0.30	0	1
Multisites	Equals 1 if more than 1 site under the same owner is certified	0.69	0.46	0	1
Recall	Equals 1 if site's parent company has recall	0.01	0.11	0	1
Competition	Number of certification bodies in the vicinity of a certified site	6.69	2.64	0	12
Category1	Equals 1 if products are raw meat, poultry, or fish products	0.26	0.44	0	1
Category2	Equals 1 if products are fruits, vegetables, or nuts	0.11	0.32	0	1
Category3	Equals 1 if products are dairy or liquid egg	0.06	0.24	0	1
Category4	Equals 1 if products are cooked meat and fish, ready-to-eat meals	0.13	0.34	0	1
Category5	Equals 1 if products are cans and jars	0.07	0.26	0	1
Category6	Equals 1 if products are ambient stable products (beverages etc.)	0.51	0.50	0	1
Public	Equals 1 if parent company is a publicly traded company	0.16	0.37	0	1
Employees	Number of employees of the parent company (1,000)	13.63	37.46	0.01	124
Sales	Sales revenue of the parent company ($ billion)	4.30	11.60	0.00	81.20
Metropolitan	Population >=50,000	0.712	0.45	0	1

Note: The numbers of observations are 3,183, 3,181, and 3,151 for the variables, public, employees, and sales respectively, 7048 for *Metropolitan* and are 7,061 for the rest. Minimum sales is $20,000 rounded to zero billions.

Table 2 describes the average grades obtained by manufacturing sites in various segments of our data in each year, and in the combined data. Our last row, which includes all observations, shows that the average grade among multisites was 3.89 and among single sites, 3.74. Sites that certified for the first time had an average grade of 3.71 while those that repeated certification had an average grade 3.89. The average grade of sites that had a recall was 3.76 compared with 3.84 for those that did not. Splitting the sample to those whose measure of the degree of competition was below the median (7 certification bodies nearby) and those for whom it was above the median, we observe average grades of 3.82 and 3.87 respectively. These patterns are consistent with our hypothesis and later empirical findings, in all cases, equality of means in the all years row is rejected.

Table 2. Average Audit Grades

	Multisites		**First Time**		**Had Recall**		**Competition < Median**	
	Yes	**No**	**Yes**	**No**	**Yes**	**No**	**Yes**	**No**
2011	3.80	3.51	3.73	NA	No site	3.73	3.70	3.75
2012	3.88	3.71	3.76	3.87	3.67	3.84	3.81	3.87
2013	3.90	3.74	3.67	3.89	3.76	3.85	3.84	3.87
2014	3.92	3.80	3.72	3.91	3.88	3.88	3.86	3.90
2015	3.93	3.77	3.63	3.90	3.65	3.87	3.87	3.88
All Years	3.89	3.74	3.71	3.89	3.76	3.84	3.82	3.87
H0: equal mean	$p < 0.01$		$p < 0.01$		$p = 0.05$		$p < 0.01$	

Note: For the last column, the median is seven certification bodies.

BRC classifies manufacturing sites into six broad categories according to the types of products they produce. These product categories are: raw products of animal or vegetable origin that require cooking prior to consumption; fruit, vegetables, and nuts; processed foods and liquids with pasteurization or heat treatment; ready-to-eat or heat and eat processed foods; ambient stable products with pasteurization or sterilization; and ambient stable products not involving sterilizations. The shares of the sites classified into each of the six categories are 26%, 11%, 6%, 13%, 7%, and 51%, respectively. The shares add to more than one because a site can be classified into more than one category, if it produces multiple products. Broad product categories are further divided into 18 detailed categories. Our Appendix table A1 presents a breakdown of the BRC broad categories to the detailed product categories.

3. Empirical Evidence

3.1. Determinants of Audit Grades

To identify factors that determine food manufacturers' audit grades in the BRC food safety certification, we use an ordered probit model, with the dependent variable being the three grade categories. Manufacturing sites obtain a higher audit grade when the audit reveals fewer non-conformities. Food manufacturers' ability to adhere to the requirements of the standard likely depends on resources available to them. As such, larger firms are expected to be better able to maintain safe manufacturing sites. While we do not have a direct measure of firms' resources, we have attempted to proxy for firms' resources using sales and the number of employees (which we only have for about half of the observations), an indicator for publicly traded company. Additionally, we can identify manufacturers that own multiple BRC certified sites. We classify a manufacturing site that shares the same owner as at least one other manufacturing site as a multisite. We expect multisites to obtain higher audit grades since these are usually more established and less financially constrained firms which can make it easier for them to comply with the requirements of the standard. An additional reason for multisites' higher grades might be that certification bodies have a stronger incentive to grade multisites more leniently as they are likely particularly eager to attract these larger producers.

If there is an element of learning in food safety certification, we would expect more experienced firms to be better able to comply with the standard, to avoid or fix non-conformities that arose in a previous audit. If this is the case, manufacturers that are audited for the first time should have lower grades compared with those audited for at least the second time. Some of the firms in our data set experienced food recalls. Recalls plausibly indicate less safe manufacturing, and thus we expect firms that had a recall to be associated with lower audit grades on average. Because only a small portion of the sites had recalls, we did not distinguish between the more severe type of recalls, class I recalls, from the other two types of recalls (classes II and III). Manufacturing sites in our data are classified into six broad (and 18 detailed) product categories.

The degree of food safety risk (e.g., raw versus processed foods), the requirements of the standard, as well as the difficulty complying with these requirements could vary by product category, and thus we control for product categories in our regressions.

Audit grades depend on the intensity of audits, and the certification bodies' tendency for leniency in grading. Certification body fixed effects can control for unobserved certification body's grading practices. We include year effects to capture changes in grading practices over time, in particular these dummies will control for grade inflation if it is present.

With these hypotheses in mind, we estimate variations of the following model:

$$(1)\quad y^*_{i,j,t} = \beta_1 Multisites_{i,t} + \beta_2 FirstTime_{i,t} + \beta_3 Recall_{i,t} + \boldsymbol{\beta_4 X_i} + \phi_j + \delta_t + \varepsilon_{i,j,t}$$

where *i* indexes the manufacturing site, *j* indexes the certification body, *t* indexes the period *t* (*t* = 2011 to 2015), and $y^*_{i,j,t}$ captures a latent food safety performance measure of site *i* if audited by certification body *j* in year *t*. The $\boldsymbol{X_i}$ term is a vector of product categories, $\beta's$ are parameters to be estimated, and ϕ_j and δ_t are certification body and time fixed effects. We conjecture that manufacturers with a performance measure that exceeds a certain threshold will obtain the grade A (assigned the numeric value 4), those whose performance measure is below a threshold would obtain the grade C (assigned the numeric value 2), and intermediate performance measure sites will be given a grade B (assigned the numeric value 3). Specifically,

$$(2)\quad y_{i,j,t} = 2 \text{ if } y^*_{i,j,t} \leq \mu_1$$

$$(3)\quad y_{i,j,t} = 3 \text{ if } \mu_1 < y^*_{i,j,t} \leq \mu_2$$

$$(4)\quad y_{i,j,t} = 4 \text{ if } \mu_2 < y^*_{i,j,t},$$

where μ_1 and μ_2 are the threshold parameters to be estimated.

Table 3 shows the estimates of our first set of models. In column (1) we only include 3 core variables of interest: Multisites, FirstTime and Recall. We add to these certification body and time fixed effects in column (2) and further add six broad product category controls in column (3). Throughout the paper, we cluster standard errors at the site level. In all specifications, we find that multisites obtain significantly higher grades, manufacturers that are audited for the first time obtains a significantly lower grade. These effects are significant at the 1% level in the three models. The coefficient on recalls is negative, but loses its significance in the full model described in column (3). Only 1% of our data had recalls.

In column (3) where we added the broad categories, only two of the product categories were significant and only at the 10% level. “Raw products” had a positive effect on grades, while “Fruits, vegetables, and nuts” category had a negative effect on grades. Raw product category includes raw red meat, raw poultry, raw prepared products, and raw fish. Because U.S. Department of Agriculture’s Food Safety and Inspection branch inspects all meat and poultry plants, such governance oversight likely leads to a higher BRC audit grade for this category.

Certain certification bodies have significant positive or negative fixed effects, after controlling for site specific characteristics.[1] This can be viewed as an indication that certification bodies have heterogeneous grading practices, exhibiting varying degrees of leniency in audit grading. The base time period in the table is 2011. Time effects in column (3)-(5) do not seem to suggest a clear grade inflation trend. Except for significant higher grades in 2014, time effects were insignificant. Thus, during the period we study, average grades are relatively stable, with an increase in average grades in 2014. The estimated threshold values are $\mu_1 = (-2.40)$ and $\mu_2 = (-0.90)$ and are highly statistically significant and significantly different from each other.[2]

We have shown that when manufacturers certify for the first time they obtain lower grades. To see if every additional year of experience continues to contribute to improvements in grades, we include in column (4) a dummy for each year of experience (excluding first time certifications this time). The coefficients of two, three, four, and five years of experience are monotonically increasing as one would expect, but the difference between the second and third year of experience is only significant at the 10% level, and the other differences between coefficients of consecutive years of experience are not significant. This likely suggests that grade improvement due to experience with the standard and with the certification process is limited to the first two years.

---

[1] The reference certification body for the certification body fixed effects is CB1 which certified the largest number of sites. We dropped a certification body out of our sample if it did not have ten or more certified sites in any of the five years. By this criterion, we dropped three certification bodies and kept fifteen.

[2] If we were to adopt Greene (2012) normalization we would have $\mu_1 = 0$ and $\mu_2 = 1.5$.

Table 3. Effects of First Time Sites, Multisites, and Certification Bodies

	(1) Core	(2) CB and Year	(3) Product Categories	(4) Second–Fifth Time	(5) Parent Info.
FirstTime	-0.608***	-0.561***	-0.570***		
	(0.04)	(0.05)	(0.05)		
Multisites	0.527***	0.524***	0.513***	0.497***	0.561***
	(0.04)	(0.05)	(0.05)	(0.05)	(0.07)
Recall	-0.277*	-0.277*	-0.228	-0.232	-0.349*
	(0.15)	(0.14)	(0.14)	(0.15)	(0.18)
Year_2012		0.054	0.047	0.107	0.077
		(0.07)	(0.07)	(0.07)	(0.10)
Year_2013		0.058	0.047	0.083	0.067
		(0.07)	(0.07)	(0.07)	(0.11)
Year_2014		0.214***	0.204***	0.177**	0.268**
		(0.07)	(0.07)	(0.08)	(0.12)
Year_2015		0.143*	0.132*	0.038	-0.043
		(0.08)	(0.08)	(0.09)	(0.13)
CB_2		0.218***	0.154*	0.161*	0.191
		(0.08)	(0.09)	(0.09)	(0.12)
CB_3		-0.482	-0.392	-0.201	-0.816**
		(0.31)	(0.32)	(0.33)	(0.41)
CB_4		-0.274**	-0.294**	-0.266**	-0.451**
		(0.12)	(0.12)	(0.12)	(0.18)
CB_5		-0.256***	-0.358***	-0.358***	-0.349***
		(0.07)	(0.08)	(0.08)	(0.12)
CB_6		0.484***	0.512***	0.520***	0.705**
		(0.14)	(0.14)	(0.14)	(0.35)
CB_7		-0.702***	-0.794***	-0.521*	-0.625**
		(0.26)	(0.27)	(0.28)	(0.29)
CB_8		-0.051	-0.096	-0.093	-0.083
		(0.07)	(0.07)	(0.07)	(0.12)
CB_9		-0.712***	-0.698***	-0.506**	0.074
		(0.22)	(0.22)	(0.23)	(0.29)
CB_10		0.009	0.044	0.259	0.102
		(0.31)	(0.31)	(0.33)	(0.45)
CB_11		0.015	0.02	0.027	0.048
		(0.08)	(0.09)	(0.09)	(0.13)
CB_12		-$0.091	-$0.121	-0.1	-0.175
		(0.13)	(0.13)	(0.13)	(0.19)
CB_13		-0.266**	-0.283**	-0.264**	-0.161
		(0.11)	(0.12)	(0.12)	(0.17)
CB_14		0.287	0.322	0.685	0.046

		(0.54)	(0.55)	(0.54)	(0.62)
CB_15		-0.195	-0.242	-0.107	-0.105
		(0.15)	(0.15)	(0.15)	(0.27)
Category1			0.154*	0.155*	0.105
			(0.09)	(0.09)	(0.14)
Category2			-0.161*	-0.163*	-0.236**
			(0.08)	(0.08)	(0.12)
Category3			0.006	0.001	0.007
			(0.11)	(0.11)	(0.17)
Category4			-0.038	-0.031	-0.191*
			(0.07)	(0.07)	(0.11)
Category5			-0.119	-0.128	-0.263*
			(0.10)	(0.10)	(0.15)
Category6			0.002	0.005	-0.074
			(0.08)	(0.08)	(0.12)
SecondTime				0.464***	0.561***
				(0.05)	(0.08)
ThirdTime				0.565***	0.655***
				(0.07)	(0.10)
FourthTime				0.664***	0.759***
				(0.09)	(0.13)
FifthTime				0.795***	0.846***
				(0.11)	(0.17)
Public parent					0.395**
					(0.20)
No of employees					-0.007
					(0.01)
Sales					0.016
					(0.02)
μ_1	-2.406***	-2.350***	-2.404***	-1.838***	-1.874***
	(0.06)	(0.10)	(0.12)	(0.11)	(0.17)
μ_2	-0.930***	-0.852***	-0.900***	-0.336***	-0.372**
	(0.04)	(0.09)	(0.11)	(0.10)	(0.15)
No. of observations (N)	7,061	7,061	7,061	7,061	3,151
No. of clusters (unique sites)	2,027	2,027	2,027	2,027	942
Pseudo R^2	0.076	0.094	0.097	0.097	0.133
Log pseudo likelihood (LPL)	-2,945	-2,888	-2,877	-2,877	-1,204

Notes: ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses.

In column (5) of table 3 we add site specific characteristics related to the parent company including: number of employees, sales and an indicator for whether the manufacturing site is owned by a publicly traded company. Since these data were not available for many of the manufacturers in our data, the sample size drops by more than a half. The estimates reported in column (5) suggest that publically traded firms have significantly higher grades. The numbers of employees and sales did not have a significant effect, possibly because these are not measured at the site level.

3.2. Robustness Tests

In table 4 we offer a number of alternative specifications to check the robustness of our results. Each model in this table departs from our preferred specification, the model in column (3) of Table 3, and introduces one change to that model. As can be seen in this table, results are robust to a number of changes. In column (1) of this table, we replace the two sets of fixed effects, certification body and time effects, with certification body by time fixed effects.

In column (2) we replace the six product categories with the eighteen more detailed product categories listed in Table A1. Most categories had insignificant effects, except a positive effect for raw red meat, raw poultry, dried goods, oils and fats, and a negative effect for beverages. The effect of the oils and fats category is statistically significant at the 10% level. Note that whereas sites classified to the raw meat and poultry product categories, which are also subject to government regulation, have significantly higher grades, while raw prepared products and raw fish category, which are not subject to government regulation, do not have a significant positive effect on grades. Such results provide further support for our argument that government mandatory inspection over raw red meat and poultry improves food safety.

In column (3) we estimate a random effects model to further exploit the panel nature of our data. We use the random effects ordered probit model, which has been much more popular than a fixed effects probit model (Greene, 2012, p. 795). Again, results changed little from our preferred specification. The effects of multisites and first time sites become slightly larger. In column (4) we restrict the sample to the 2012-2015 years (excluding the first year of data which was provided to us separately with a gap in observations). With about 1,000 less observations, the product category effects become insignificant. Other results remain robust.

Finally, we added an interaction term between first-time sites and multisites. The interaction term stemmed from the consideration that a multisite new to the certification has an opportunity to learn from other certified multisite(s) under the same parent company. Therefore, we expect the interaction effect to be positive. Results reported in column (5) show that the effects of first-time sites and multisites remain robust and the interaction term is positive and statistically significant. Such result provides support for our hypothesis that food safety practice/knowledge can transfer between manufacturing sites under the same parent company.

Table 4. Robustness Checks

	(1) CB*Year	(2) 18 Product Categories	(3) Random Effects	(4) 2012–2015	(5) Interaction
FirstTime	-0.585***	-0.576***	-0.686***	-0.574***	-0.675***
	(0.05)	(0.05)	(0.06)	(0.05)	(0.06)
Multisites	0.524***	0.504***	0.614***	0.478***	0.429***
	(0.05)	(0.05)	(0.06)	(0.05)	(0.06)
*Firsttime*Multisites*					0.220***
					(0.08)
Recall	-0.228	-0.234	-0.294*	-0.252*	-0.227
	(0.15)	(0.15)	(0.18)	(0.14)	(0.14)
PC1 (*Raw red meat*)		0.204**			
		(0.10)			
PC2 (Raw Poultry)		0.411***			
		(0.10)			
PC12 (Beverages)		-0.387***			
		(0.13)			
PC15 (Dried goods)		0.251***			
		(0.07)			
*PC18 (*Oils and fats)		0.227*			
		(0.13)			
μ_1	-2.330***	-2.346***	-2.927***	-2.436***	-2.419***
	(0.14)	(0.12)	(0.16)	(0.13)	(0.12)
μ_2	-0.812***	-0.827***	-1.106***	-0.935***	-0.909***
	(0.13)	(0.11)	(0.14)	(0.12)	(0.12)
sigma2			0.504***		
			(0.06)		
CB and year effects	No	Yes	Yes	Yes	Yes
CB*year effects	Yes	No	No	No	No
Broad categories	Yes	No	Yes	Yes	Yes
Detailed categories	No	Yes	No	No	No
N	7,061	7,061	7,061	6,107	6,107
No. of clusters	2,027	2,027	2,027	1,954	2,027
Pseudo R^2	0.107	0.107		0.094	0.098
LPL	-2,845	-2,845	-2,790	-2,302	-2,874

Notes: ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses. Only statistically significant product category (PC) effects are displayed but all are included in column (2) specification.

3.3. Certification Bodies Competition and Grades

The grading practices of certification bodies could be influenced by the degree of competition between them. Bar and Zheng (2016) have shown that manufacturers prefer certification bodies that are geographically closer, and that assigned a higher share of A grades in the previous year. If certification bodies compete for manufacturers, they may have an incentive to assign them higher grades. In the next set of models, we investigate whether the number of competing certification bodies in a site's vicinity affects its audit grade.

To account for the effect of competition between certification bodies in the site's vicinity we define a site specific "Competition" variable as the number of certification bodies that are at a distance of at most *d* from the site. We would like *d* to be a "reasonable" travel distance between the certification body office and the audited site. We first take *d* to be the average distance in our data traveled between sites and their chosen certification bodies in a given year. For robustness, we also consider alternative definitions of *d* to be 10 percent more than this actual average travel distance or 10 percent less.

In the first three columns of Table 5 we estimate the same models as in the first three columns of Table 3 with the added competition variable. First we observe that adding the competition variable does not change our main results from Table 3. Multisites have higher grades, and first time certified sites have lower grades. The coefficient on FDA recall is marginally significant in the first two columns, but becomes insignificant in the full model. Our measure of competition—the number of competing certification bodies in each site's vicinity—has a positive and significant effect on grades in all specifications. Thus, sites obtain higher audit grades when there are more competing certification bodies around them. We control for the characteristics of a site's location by adding state fixed effects in column (4) or an indicator site being in a metropolitan area (with population larger than 50,000) in column (5). These controls alleviate concerns that the positive competition effect could be driven by sites location being simultaneously more attractive to high quality manufacturers and to certification bodies. Most state effects were insignificant (only Colorado, New Hampshire, and Utah had significant effects at 10% level, compared with the base state of Alabama). The effect of the metropolitan indicator is positive but insignificant. The positive effect of competition remains significant after the addition of location characteristics.

Table 5. Effect of Certification Body Competition

	(1) Core	**(2) CB and Year Effects**	**(3) Product Categories**	**(4) State Effects**	**(5) Metro-politan**
FirstTime	-0.581***	-0.557***	-0.565***	-0.570***	-0.567***
	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)
Multisites	0.542***	0.531***	0.520***	0.514***	0.518***
	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)
Recall	-0.276*	-0.274*	-0.229	-0.237	-0.231
	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)
Competition	0.030***	0.026***	0.024***	0.044**	0.025***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Metropolitan					0.046
					(0.05)
CB and year FE	Yes	Yes	Yes	Yes	Yes
Broad categories	No	No	Yes	Yes	Yes
State FE	No	No	No	Yes	No
μ_1	-2.194***	-2.206***	-2.272***	-2.116***	-2.239***
	(0.08)	(0.11)	(0.13)	0.26	0.14
μ_2	-0.715***	-0.707***	-0.766***	-0.589**	-0.732***
	(0.07)	(0.10)	(0.12)	0.26	0.13
N	7,061	7,061	7,061	7,061	7,048
No. of clusters	2,027	2,027	2,027	2,027	2,023
Pseudo R^2	0.079	0.096	0.099	0.110	0.099
LPL	-2,936	-2,882	-2,872	-2,836	-2,869

Notes: ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses.

In Table 6 we include robustness checks for table 5. In column (1) we add state fixed effects and measures of population density at the same time. The coefficients on first time certification and multisites remain significant at the 1% level, the negative effects of recalls becomes significant at the 10% level and the significance of the effect of competition variable is weakened in this specification, but remains positive. In columns (2) and (3) we estimate the same model as in table 5 column (3) only replace the competition variable with the two perturbed definitions of the competition variable, results remain essentially unchanged, although the magnitude of the coefficient of competition variable slightly decreases.

Table 6. Effect of Certification Body Competition

	(1) State and Population	**(2) Competition1.1**	**(3) Competition0.9**
FirstTime	-0.571***	-0.565***	-0.566***
	(0.05)	(0.05)	(0.05)
Multisites	0.515***	0.518***	0.522***
	(0.05)	(0.05)	(0.05)
Recall	-0.241*	-0.231	-0.23
	(0.15)	(0.14)	(0.15)
Competition	0.044**	0.021**	0.023***
	(0.02)	(0.01)	(0.01)
Metropolitan	0.05		
	(0.06)		
CB and year FE	Yes	Yes	Yes
Broad categories	No	Yes	Yes
State FE	Yes	No	No
μ_1	-2.080***	-2.280***	-2.289***
	(0.27)	(0.13)	(0.13)
μ_2	-0.552**	-0.774***	-0.784***
	(0.26)	(0.13)	(0.12)
N	7,048	7,061	7,061
No. of clusters	2,023	2,027	2,027
Pseudo R^2	0.111	0.098	0.098
LPL	-2,832	-2,874	-2,873

Notes: ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses.

4. Concluding Remarks

The oversight on food safety is increasingly shifting from the government to private, third party food safety certification. On January 4, 2011 the Food Safety Modernization Act (FSMA) was signed into law. A stated aim of this legislation is to "ensure the U.S. food supply is safe by shifting the focus of federal regulators from responding to contamination to preventing it."[3] Third party certification plays an important role in the implementation of this reform. Retailers are also increasingly relying on third party certification. Our research contributes to better understanding of third party certification by examining the determinants of audit grades in certification for the BRC food safety standard.

Several results emerge from our empirical analysis. First, our paper provides evidence that larger, more established, and publicly traded firms are better able to comply with food safety standards. This suggests that food safety certification likely creates a competitive advantage for firms with more resources. Furthermore, we also found evidence that food safety practices or knowledge can transfer from certified manufacturing sites to sites without certification under the same parent company.

Second, firms obtain higher grades as they gain experience in certification. This seems to offer a positive message: certification leads manufacturing sites to improve their quality and safety practices and to continue benefiting from their investments and learning in future years. The result also provides support for the use of third party certification to effectively increase the overall level of food safety, because we see the number of non-conformities decreased for repeating manufacturing sites compared with first-time certified sites.

Third, we have seen that the degree of competition between certification bodies can affect audit grades. We also find variations in average grades between certification bodies, suggesting certification bodies are heterogeneous in their degree of audit intensity or grading leniency. Economic theory has provided reasons for concern about a conflict of interest in the certification business: the audited firms chooses certification bodies and pay for auditing costs. While pleasing existing clients and attracting new clients could provide certification bodies an incentive for leniency in audit grades, the need to adhere to the requirements of the standard program

[3] See http://www.fda.gov/Food/GuidanceRegulation/FSMA/ucm247546.htm

moderates this tendency. Certification bodies' tendency for leniency could be higher for clients that they are particularly eager to attract, such as large scale multisite producers. This could in part explain the higher grades obtained by multisites. Heterogeneous grading leniency can provide firms an incentive to shop around to obtain a higher grade. Currently, the food safety certification industry is not regulated. Standard holders are responsible for making sure the certification bodies are performing properly. Our paper raises a doubt on whether a self-regulated food safety certification industry can be fully objective.

Although we did not find any clear grade inflation trend, BRC audit grades tend to be compressed at the top with more than 85 percent of sites obtaining the grade A, and more than 73 percent of first time certified sites obtaining the grade A. BRC's newly introduced AA grades might be one attempt to address this grade compression phenomenon. With competitive pressure on grades, time will tell if AA grades will become the new A.

5. References

Albano, G.L., and A. Lizzeri. 2001. "Strategic Certification and Provision of Quality." *International Economic Review*, vol. 42(1), pp. 267–283.

Anders S., D.M. Monteiro, and E. Rouviere. 2007. "Objectiveness in the Market for Third-Party Certification: Does market structure matter?" *International Marketing and International Trade of Quality Food Products*, 105th *Seminar of the European Association of Agriculture Economics*, pp. 651–64.

Bae, K.-H., J.-K. Kang, and J. Wang. 2015. "Does Increased Competition Affect Credit Ratings? A Reexamination of the Effect of Fitch's Market Share on Credit Ratings in the Corporate Bond Market," *Journal of Financial and Quantitative Analysis*, vol. 50(5), pp. 1011–1035.

Bar T., V. Kadiyali, and A. Zussman. 2009. "Grade Information and Grade Inflation: The Cornell Experiment." *Journal of Economics Perspectives*, vol. 23(3), pp. 93–108.

Bar T., and Y. Zheng. 2016. "Strategic Selection of Certifiers: Evidence from the BRC Food Safety Standard." University of Connecticut and University of Kentucky. Working paper.

Baron, D. P., 2011. "Credence attributes, voluntary organizations, and social pressure," *Journal of Public Economics*, 95, 1331-1338.

Cohen A., and M.D. Manuszak. 2013. "Ratings Competition in the CMBS Market," *Journal of Money, Credit and Banking*, Supplement to vol. 45(1), pp. 93-119.

Dranove D., and G.Z. Jin. 2010. "Quality Disclosure and Certification: Theory and Practice." *Journal of Economic Literature*, vol. 48(4), pp. 935–63.

Fagotto, E. 2014. "Private Roles in Food Safety Provision: the Law and Economics of Private Food Safety." *European Journal of Law and Economics*, vol. 37, pp. 83–109.

Farhi, E., J. Lerner, and J. Tirole. 2013. "Fear of Rejection? Tiered Certification and Transparency." *RAND Journal of Economics*, vol. 44(4), pp. 610–31.

Fulponi, L. 2006. "Private Voluntary Standards in the Food System: The Perspective of Major Food Retailers in OECD Countries." *Food Policy*, vol. 31, pp. 1–13.

Gaigné C., and B. Larue. 2016. "Quality Standards, Industry Structure, and Welfare in a Global Economy," *American Journal of Agriculture Economics*. aaw039v1-aaw039.

Greene, W.H. 2012. Econometric Analysis, seventh edition, Prentice Hall, New Jersey.

Hatanaka, M., C. Bain, and L. Busch. 2005. "Third-party Certification in the Global Agrifood System." *Food Policy*, vol. 30(3), pp. 354–69.

Henson, S., and J. Caswell. 1999. "Food Safety Regulation: An Overview of Contemporary Issues." *Food Policy*, vol. 24, pp. 589–603.

Hubbard, T. N. 1998. "An empirical examination of moral hazard in the vehicle inspection market." *RAND Journal of Economics*, vol. 29(2), pp. 406–426.

Jahn, G., M. Schramm, and A. Spiller. 2005. "The Reliability of Certification: Quality Labels as a Consumer Policy Tool." *Journal of Consumer Policy*, vol. 28, pp. 53–73.

Johnson, V.E. 2003. *Grade Inflation: A Crisis in College Education.* New York: Springer-Verlag.

Lizzeri, A. 1999. "Information Revelation and Certification Intermediaries." *RAND Journal of Economics*, vol. 30(2), pp. 214–31.

Podhorsky, Andrea. 2013. "Certification Programs and North-South Trade," *Journal of Public Economics*, Vol. 108, p. 90-104.

Tanner, B. 2000. "Independent Assessment by Third-Party Certification Bodies." *Food Control*, vol. 11, pp. 415–17.

U.S. Food and Drug Administration, 2001. "Assessment of the Third-Party Certification Pilot for Aquacultured Shrimp," available at http://www.fda.gov/ForConsumers/ConsumerUpdates/ucm265934.htm

Appendix

Table A1. BRC Product Categories

Broad Category (BRC Field of Audit)	**Field Description**	**Detailed Category**	**Category Description**
1	Raw products of animal or vegetable origin that require cooking prior to consumption	1	Raw red meat
		2	Raw poultry
		3	Raw prepared products
		4	Raw fish
2	Fruits, vegetables, and nuts	5	Fruits, vegetables, and nuts
		6	Prepared fruit, vegetables and nuts
3	Processed foods and liquids with pasteurisation or UHT as eat treatment or similar technology	7	Dairy, liquid egg
4	Processed foods, ready-to-eat or heat	8	Cooked meat and fish
		9	Raw cured and/or fermented meat and fish
		10	Ready-to-eat meals
5	Ambient stable products with pasteurisation or sterilisation as heat treatment	11	Cans and jars
6	Ambient stable products not involving sterilisation as heat treatment	12	Beverages
		13	Alcoholic drinks
		14	Bakery
		15	Dried goods
		16	Confectionery Dried goods
		17	Cereals and nuts
		18	Oils and fats

Source: BRC Global Standard for Food Safety Issue 7.