Examining Price and Weather Events as Market Indicators of Outbreaks of Shiga Toxin-Producing *Escherichia Coli* Infections Linked to Romaine Lettuce Consumption: A Daily Panel Data Analysis

Minor, Travis, FDA/CFSAN; Brown, Christian, FDA/OC; Biemiller, Nathan, FDA/OC; Nolte, Kurt, FDA/CFSAN; Bazaco, Michael, FDA/CFSAN; Viazis, Stelios, FDA/CFSAN

Abstract

Background: Romaine lettuce has become a repeated vehicle of foodborne illness outbreaks. Because much of the lettuce consumed in the U.S. originates from California and Arizona, there is a large burden of illnesses associated with these growing regions. Investigators face difficulty tracing outbreaks and determining their causes due to various challenges, including comingling of product and record limitations within the supply chain. **Purpose:** We investigated the effect of price and shipment volumes and growing region temperature fluctuations on the probability of identifying outbreaks of STEC infections associated with romaine consumption to better understand when conditions conducive to an outbreak may be signaled in the marketplace. **Methodology:** We constructed a daily panel (2009-18) of market and weather data for seven romaine outbreaks in CA and AZ growing regions and conducted an econometric analysis of the relative importance of these indicators on the probability of identifying an outbreak. **Results:** Price spikes of 1 standard deviation from mean price were associated with a 5-25 percentage point increase in the probability of identifying an outbreak. Downward trends in shipments were associated with a 2-7 ppt increase in the probability of identifying an outbreak. Large fluctuations in daily temperature in growing regions were associated with a 2-12 ppt increase in the probability of identifying an outbreak. **Conclusion:** While weather events may affect crop quality and safety, price or shipment changes may be stronger signals of an outbreak. Economic indicators may help to identify market disturbances that could be indicative of an outbreak. Coupled with epidemiologic information, this may help identify conditions leading to outbreaks and either target regions with increased need for grower education, outreach, and potential regulatory activities.

Introduction

Shiga Toxin-producing *Escherichia co*li (STEC) cause over 265,000 infections in the United States every year, with 3,600 hospitalizations and 30 deaths. From 2003 to 2012, beef and leafy greens were determined to be the source of more than 25 percent of all reported *E. coli* O157 outbreaks and more than 40 percent of associated illnesses. More recently, in the years between 2009 and 2018, the Centers for Disease Control and Prevention (CDC) identified 40 foodborne illness outbreaks of STEC infections linked to leafy greens in the U.S. and Canada. These outbreaks resulted in 1,212 reported illnesses, 77 cases of hemolytic uremic syndrome (HUS), and eight deaths.

Some of the challenges to outbreak investigations include: 1) obtaining sufficient exposure information to link ill persons and the specific contaminated leafy green; 2) conducting a traceback investigation of the suspected leafy green that can identify a common source of contamination at the farm level; and 3) on-farm investigations rely on the development of actionable epidemiologic and traceback evidence, and when they do occur, they may take place on fields that are fallow or have different crops planted. As a result, the conditions and practices observed are temporally removed from the contamination event. Therefore, it is important to invest in prevention efforts to avoid contamination that may lead to outbreak events in the first place. These ongoing efforts include education and outreach on good agricultural practices and compliance with the Food Safety Modernization Act Produce Safety Rule.

Materials and Methods

We construct a daily longitudinal data set of economic, weather, and STEC outbreak data for 7 romaine growing regions (6 CA, 1 AZ), yielding 20,346 region-day observations.

We then estimate the following econometric model of the effect of market and weather effects on outbreak illnesses:

$$Outbreak\ Illness_{r,t} = \beta_0 + \beta_1 Price\ Spike_t + \beta_j \sum Shipments_{r,t} + \beta_k \sum Weather_{r,t} + \beta_7 \Delta\text{Temp}_{r,t} + \mu_r + \nu_t + \upsilon_{r,t}$$

where

- $Outbreak\ Illness_{r,t}$: indicator variable =1 if an illness associated with an outbreak traced to region r is recorded on day t;
- $Price\ Spike_t$: indicator =1 if the national price was ≥1 standard deviation above the mean at least once in the previous 7-21 days;
- $\sum Shipments_{r,t}$: two indicators =1 if shipments from region r are (1) rising or (2) falling substantially (relative to region r mean), indicating the beginning/end of harvest season for region r;
- $\sum Weather_{r,t}$: three indicators =1 if: (1) daily max. temp. rose above 104°F in a single day or stayed above 99°F for 3 consecutive days at least once in the prior 14-35 days for region r; (2) daily min. temp. fell below 31°F in a single day at least once in the prior 14-35 days; (3) daily rainfall measured 1 in. or more in a single day or measured above 0.2 in. for 3 consecutive days at least once in the prior 14-35 days;
- $\Delta\text{Temp}_{r,t}$: =1 if the difference between the daily max. and min. temp. >48° in any single day in the prior 14-35 days for region r;
- μ_r region-specific fixed-effects;
- ν_t season-specific fixed-effects; and
- $\upsilon_{r,t}$ well-behaved error term.

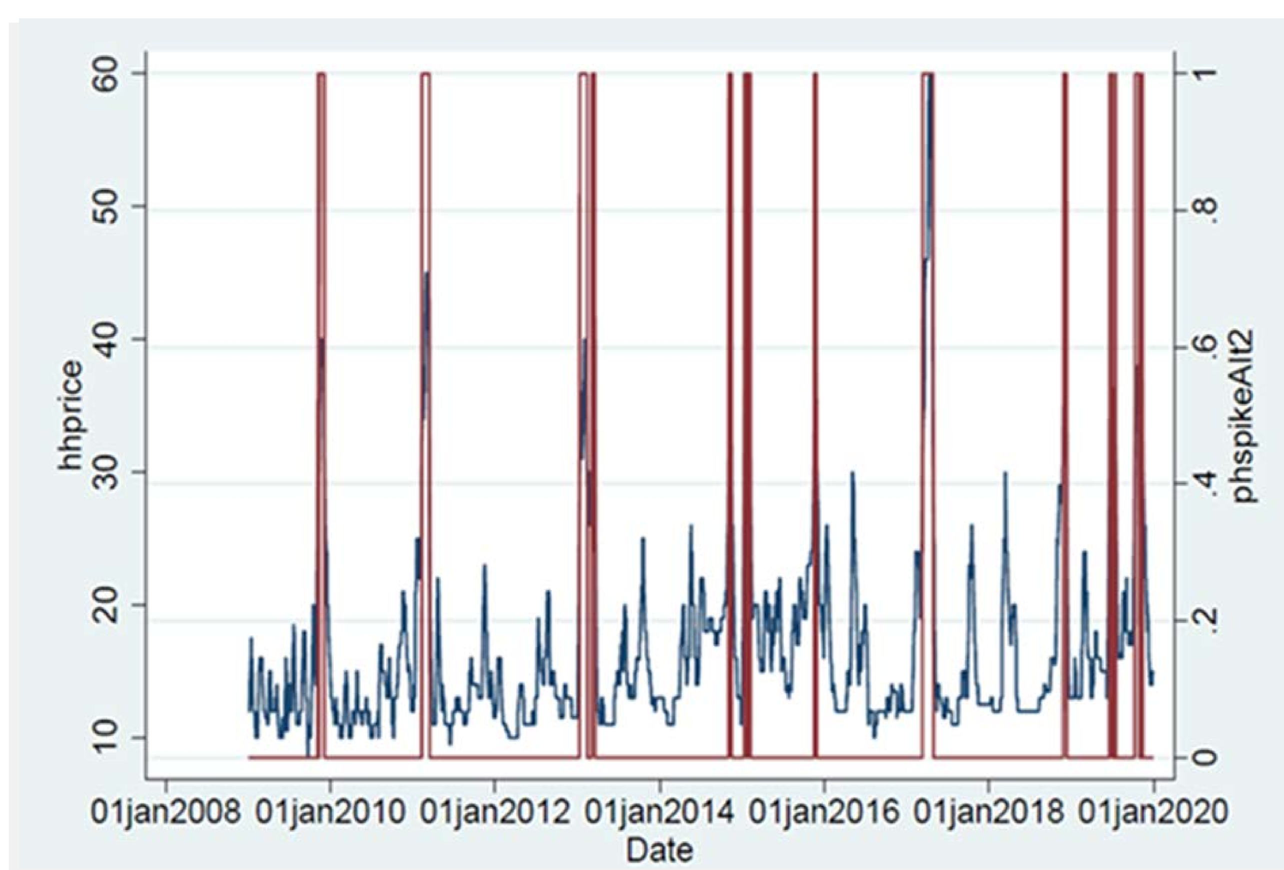

Figure 1. Price of Romaine Lettuce and Dates of Price Spikes in Growing Regions. N = 28,119. Data are drawn from 7 growing regions (6 CA, 1 AZ).

Results and Discussion

Our results suggest that a price change of 1-s.d. from the average price is associated with a 5-25 ppt increase in the probability of identifying an outbreak illness, although there is some evidence that lagged prices may provide an even more salient signal than price spikes as defined in our analysis. Similarly, a downward trend in shipments, which signifies the end of a region's harvest season, is associated with a 2-7 ppt increase in the probability of identifying an outbreak illness. In addition, large fluctuations in the daily temperature are associated with a 2-12 ppt increase in the probability of identifying an outbreak illnesses. Finally, other weather events, such as significant heat or freeze events, are negatively associated with the probability of identifying an outbreak.

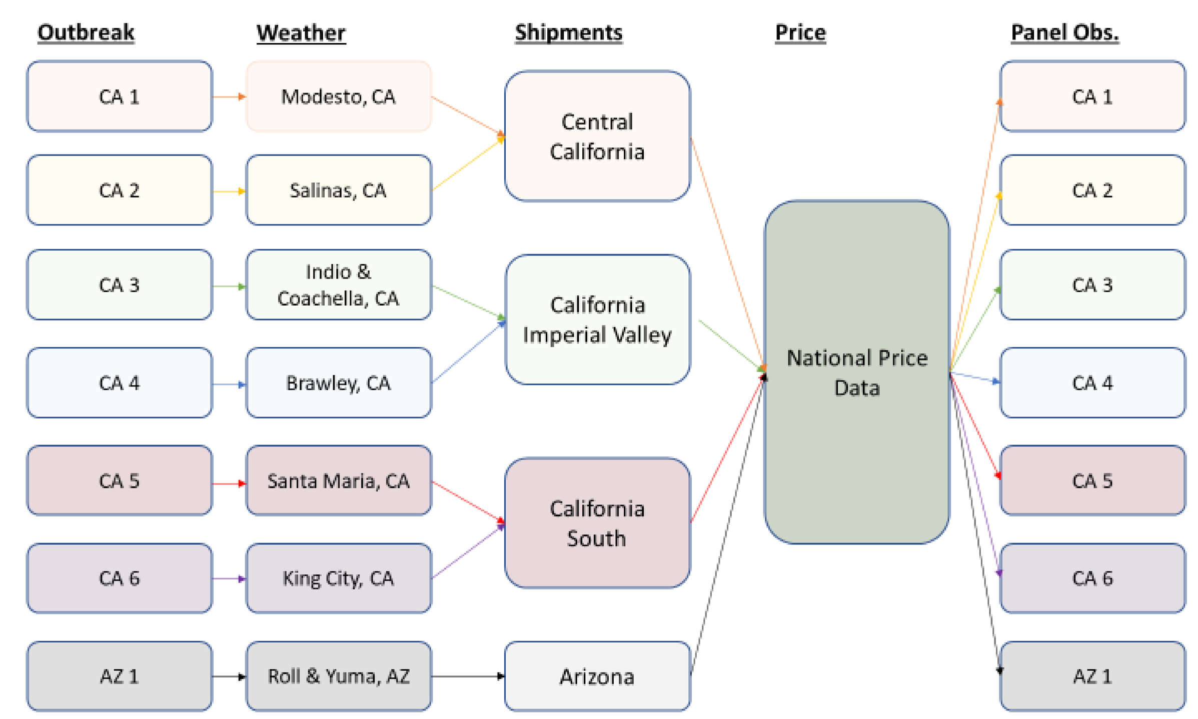

Figure 2. Structure of Daily Growing Region Data. Product supply chain data to identify the source of food served or sold at a specific locations are traced back to field of origin. Weather data is collected from regional weather stations. Shipment data is collected at the regional level by USDA/AMS. Romaine prices are collected at the national level by USDA/AMS.

	Description	Mean	Std. Dev.	Min.	Max.
Outcome					
Outbreak Illness	=1 if in the illness onset window for an identified outbreak by region	0.01	0.11	0	1
Market Signals					
Price Spike	=1 if price rose more than 2 standard deviations above the mean at least once in the harvest window	0.11	0.31	0	1
End of Season	=1 if shipments fell by more than 2 standard deviations over 2 weeks within the harvest window	0.06	0.24	0	1
Weather Events					
Heat Event	=1 if temp. rose above 100 in a single day in the growing window	0.19	0.40	0	1
Freeze Event	=1 if temp. fell below 32 in the growing window	0.18	0.38	0	1
Rain Event	=1 if 0.5 in. of rain or more fell in a single day o in the growing window	0.10	0.30	0	1
Temp. Swing	=1 if the difference between the min. and max. temperature >= 48 degrees in the growing window	0.09	0.29	0	1
Lag Variables					
L. Price	14-day lag of market price	16.82	7.20	8.5	60
L. Shipments	14-day lag of regional shipments (in 1,000s)	31.52	49.68	0	168.37
L. Shipping Change	14-day lag of the change in regional shipments from two weeks prior	0.00	0.11	-0.52	0.57
L. Temp Max.	28-day lag of maximum daily temperate	76.95	14.88	39	121
L. Temp Min.	28-day lag of minimum daily temperate	49.55	11.59	15	92
L. Precipitation	28-day lag of daily rainfall total	0.02	0.10	0	3.64

Table 1. Summary Statistics. N = 28,119. Harvest window is defined as 7-21 days prior. Growing window is defined as 14-35 days prior.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Key Variables								
In Season	0.018***	0.018***	0.018***	0.017***				
	(0.003)	(0.003)	(0.003)	(0.003)				
Price Spike		-0.001		-0.010**		-0.001		-0.017***
		(0.003)		(0.004)		(0.004)		(0.005)
Seasonality								
End of Season	0.028***	0.028***	0.020***	0.019***	0.039***	0.039***	0.035***	0.033***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Heat Event	-0.002	-0.002	-0.008*	-0.007*	0.012**	0.011**	0.002	0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)
Rain Event	-0.025***	-0.025***	-0.019***	-0.018***	-0.034***	-0.034***	-0.025***	-0.024***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)
Freeze Event	-0.027***	-0.026***	-0.022***	-0.022***	-0.034***	-0.034***	-0.022***	-0.023***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)
Temp. Change	0.003	0.003	0.001	0.002	0.000	0.000	-0.003	-0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Lags								
L. Price			0.000*	0.001***			0.001***	0.001***
			(0.000)	(0.000)			(0.000)	(0.000)
L. Shipments			0.000***	0.000***			0.000***	0.000***
			(0.000)	(0.000)			(0.000)	(0.000)
L. Shipping Change			-0.000***	-0.000***			-0.000***	-0.000***
			(0.000)	(0.000)			(0.000)	(0.000)
L. Temp Max.			0.001***	0.001***			0.001***	0.001***
			(0.000)	(0.000)			(0.000)	(0.000)
L. Temp Min.			0.000	-0.000			0.001***	0.001***
			(0.000)	(0.000)			(0.000)	(0.000)
L. Precipitation			-0.018	-0.019			-0.025	-0.026
			(0.023)	(0.023)			(0.030)	(0.030)
Observations	20,346	20,346	20,318	20,318	13,244	13,244	13,216	13,216
Controls								
Price Spike	No	Yes	No	Yes	No	Yes	No	Yes
In Season Only	No	No	No	No	Yes	Yes	Yes	Yes
Lagged Vars	No	No	Yes	Yes	No	No	Yes	Yes

Table 2. Marginal Effects of Price and Weather Events on Identified Outbreak Illnesses. Significance levels: ***1%, **5%, *10%.

Conclusion

Our findings suggest that market signals may help investigators earlier identify conditions leading to contamination events linked to outbreaks of STEC infections associated with the consumption of romaine lettuce, which could facilitate targeted education and outreach efforts or inspectional activities to avoid food safety lapses leading to future outbreaks. Specifically, our analysis shows that the end of a region's growing season presents a higher likelihood of an associated foodborne outbreak event. Abnormal weather events may also influence the probability of a foodborne outbreak event. The signs of estimated effects, however, are somewhat puzzling: extreme heat is shown to increase the outbreak probability in some specifications and decrease outbreak probability in others. Freeze and rain events are estimated to reduce the probability of a foodborne outbreak in romaine. Similarly, prices and shipments seem to matter, but effects may be better captured by lagged measures than specific cutoff values. This is shown in both the measured effects on price spikes (sometimes reducing the probability of an outbreak event) and heat events (sometimes negative) while lagged price and maximum daily temperature remain strongly, positively associated with the probability of detecting an outbreak event. This may suggest it is better to monitor growing conditions and market fluctuations in areas of concern, rather than always rely on product recalls that happen after consumers become ill.