Tracing Produce from Farm to Retail and Food Safety in the U.S.

Marziyeh Bahalou Horeh

Graduate student, Division of Resource Economics and Management, West Virginia University

Levan Elbakidze

Associate professor, Division of Resource Economics and Management, West Virginia University

Selected Paper prepared for presentation at the 2022 Agricultural & Applied Economics Association Annual Meeting, Anaheim, CA; July 31-August 2

Tracing Produce from Farm to Retail and Food Safety in the U.S.

Abstract

Food traceability is an important instrument in the toolbox for managing foodborne outbreaks. Rapid tracing systems help remove the contaminated produce from the supply chain by identifying the origins of contamination. We examine the economic benefits of traceability in the lettuce supply chain using a stochastic price endogenous partial equilibrium model. Using irrigation water as a potential pathogen source we show that if the average cost per foodborne illness is $8,500, then the annual benefits of traceability in the lettuce industry vary between $48 and $139 million, depending on the length of produce shelf life. Also, the number of avoided exposures to foodborne pathogens due to traceability is between 973 and 2,864 cases. We also observe that the benefits of traceability depend on microbial die-off rate, monetary value of foodborne illness damages, pathogen transmission from source water to crop, and pathogenicity of water per unit of *Generic E. coli*. The results shed light on the economic merits of investments in tracing from farm to retail.

Key words: Traceability, Foodborne illness, Partial equilibrium.

JEL classification: D18, D61, Q18.

1. Introduction

Recent large-scale food safety incidents have heightened the significance of traceability in the food supply chain (Collart and Canales, 2022; Shew et al., 2021). The U.S. Food and Drug Administration (FDA) defines food traceability as the capacity to trace the movement of food products and their ingredients through the supply chain (FDA, 2020). Food traceability involves documenting and tracking the production, storage, processing, distribution, and consumption of food products and their ingredients (FDA, 2020). Although traceability technology has been greatly improved, full traceability in food products has not been achieved. For example, in March 2021, a multistate outbreak of *E. coli* O157:H7 caused 22 illnesses, 11 hospitalizations, and 1 death. Health officials, including the FDA, Centers for Disease Control and Prevention (CDC), and the U.S. Department of Agriculture (USDA)'s Food Safety and Inspection Service (USDA-FSIS) were not able to identify the source of this outbreak. On the other hand, in April 2021, the FDA and the CDC successfully investigated a multistate outbreak of *Salmonella* that sickened 20 people. By matching the genetic fingerprints of the *Salmonella* strain to the *Salmonella* strain found in cashew-based products, the investigators linked this outbreak to a producer in Carlsbad, California. Following this investigation, the responsible producer recalled all susceptible products (FDA, 2021).

Firms' incentives to engage in product tracing include protection or restoration of reputation and improvement of supply management (Pouliot and Sumner, 2008). In the absence of traceability, it is difficult for retailers to pass liability to the responsible suppliers. Traceability can also reduce costs of distribution systems, decrease recall costs, and support high-value sales (Golan et al., 2004). The major benefit of traceability systems is improved food safety (Lusk and McCluskey, 2018). In the case of a foodborne illness outbreak or contamination event, efficient product tracing helps government agencies and those involved in producing and selling food to rapidly identify the source of contamination. This enables faster removal of the contaminated product from the marketplace, which reduces foodborne illnesses.

Food traceability has gained significant attention from government officials and policy makers. For instance, the "Requirements for Additional Traceability Records for Certain Foods" is proposed by the FDA to establish traceability recordkeeping requirements (additional to what is currently required by existing regulations). This proposal requires manufacturers, processors, packers, or holders of foods to create and maintain records involving Key Data Elements (FDA, 2020). The proposed requirements help the FDA effectively identify producers of contaminated foods and prevent or mitigate foodborne illness outbreaks.

Food traceability is costly. Costs of traceability include recordkeeping and product differentiation costs (Golan et al., 2004). Recordkeeping costs involve collecting and documenting information on food products as they move across the supply chain (Golan et al., 2004). Product differentiation costs refer to differentiating a product from other products for tracing purposes. The cost of traceability depends on various factors such as the precision, depth and breadth of the tracing system,[1] and the technological implementation arrangements (Golan et al., 2004). The greater the precision, breadth, depth, number of transactions, and technological sophistication, the larger the costs. For example, the FDA estimates that the annual cost of compliance with the "Requirements for Additional Traceability Records for Certain Foods" rule will be between $411 and $535 million under full and partial exemption for small food retail entities, respectively (FDA, 2020). The estimated costs are calculated for products on the Food Traceability List (FTL)[2] and include initial capital investment, recordkeeping, and training costs.[3]

The literature on the economics of traceability in the food supply chain is sparse. Some studies estimate the costs and benefits of food traceability. For instance, Brown et al. (2021) use a Monte Carole simulation method to analyze the benefits and costs of Whole Genome source tracking implemented by the FDA. The estimated annual benefits and costs are approximately $500 and $22 million, respectively. Mejia et al. (2010) use estimations from other studies (i.e.,

[1] The breadth of traceability indicates the amount of recorded information by a traceability system. The depth is how far in the supply chain traceability system extends. The depth of traceability depends on product attributes that regulators or firms find worthy of tracing. For instance, one traceability system may only need a product tracing to be extended to the processing stage while another system may need a product to be traced back to the first production stage (Golan et al., 2004).

[2] The FDA uses the results of a risk-ranking model to identify the products on the Food Traceability List that require additional traceability. These products include cheese, shell eggs, nut butter, cucumbers, herbs, leafy greens, melons, peppers, sprouts, tomatoes, tropical tree fruits, fruits and vegetables, finfish, crustaceans, mollusks, bivalves, and ready to eat salads (FDA, 2020).

Nganje et al., 2009) to approximate the costs and benefits of implementing produce tracing for farms covered by the California Leafy Green Marketing Agreement (LGMA). The estimated costs in their study are between $3 and $110 million, depending on the type of technology adopted for tracing. The estimated benefits of traceability from reducing foodborne illnesses are between $10 and $95 million.

Some empirical studies estimate the willingness to pay (WTP) for traceability in the food industry. For example, using a choice experiment, Lin et al. (2022) find that Chinese consumers are willing to pay a premium of $0.63 per pound of American beef for blockchain-based traceability[4] relative to conventional digital traceability. In a blockchain-based traceability transaction data are shared by all the relevant stakeholders through the supply chain, rather than being controlled by a single authority. This creates more data transparency and trust in blockchain-based traceability relative to the conventional traceability system. Shew et al. (2021) use a choice experiment method to study WTP for food traceability in the U.S. beef industry. They show that beef consumers are willing to pay a premium for the USDA certification labels. Liu et al. (2019) conduct a choice experiment to investigate the consumers' trust and willingness to pay for food traceability in China. They show that WTP depends on the degree of consumers' trust in the government's supervision of food safety and food labels. Pouliot (2011) estimates the Canadian slaughterhouses' WTP for traceability of steers using a hedonic framework. The results suggest that WTP for traceability is between C$0.02 and C$0.06/lb of carcass.

Previous literature also explores perceptions and preferences for traceability. For instance, Schulz and Tonsor (2010a) use a survey of 1998 cow-calf producers and a choice experiment model to determine perceptions regarding animal traceability. They find that most important issues in developing an animal traceability system include monitoring/managing disease and increasing consumers' confidence. Schulz and Tonsor (2010b) use a similar survey and choice experiment model to identify the preferences of cow-calf producers for the U.S. voluntary traceability systems. The results show that producers are sensitive to information requirements and price of system.

[4] Blockchain-based traceability uses a digital technology in documenting and maintaining information on food products. Rather than being approved by a single authority, this technology is approved by every party involved in a supply chain.

Some studies analyze traceability adoption behavior. For example, Zhou et al. (2022) deploy a theoretical and an empirical linear probability model to investigate the impact of policies such as food safety information disclosure on adoption of traceability by the Chinese vendors. Results indicate that vendors' traceability adoption is positively correlated with the information disclosure. Using multi-variate probit equation system and a sample of Taiwanese fruit and vegetable producers, Liao et al. (2011) study the farmers' participation in the Taiwan Agriculture and Food Traceability Program. The results suggest that farmers' awareness of the program is the significant determinant of the participation decision. Monteiro and Caswell (2009) use a discrete choice model to identify the farm-scale traceability adoption decisions in the Portuguese pear industry. The results show that farm-level adoption of traceability is best explained by farm productivity and farmer's age.

We contribute to the existing literature with an estimation of lettuce tracing value from farm to retail and an examination of the factors that influence the estimated economic value. We focus on foodborne pathogen introduction via irrigation as the first on-farm contamination stage. This study contributes to the existing literature on the benefits of traceability in two ways. First, instead of using simulations model with a dataset that includes only a portion of foodborne outbreaks (e.g., Brown et al., 2021) or back-of-envelop estimates (e.g., Mejia et al., 2010), this paper employs a more detailed analysis of foodborne outbreaks using a stochastic partial equilibrium model that considers the heterogeneity of crop production across districts. This allows the identification of the source of an outbreak and removal of contaminated product from the corresponding district. Second, our stochastic model accounts for risk via the incorporation of random pathogen content in irrigation water. Also, to the best of our knowledge, there is no prior study on the characterization of the pathogen-associated factors (e.g., pathogenicity of irrigation water) affecting the benefits of produce tracing.

Fresh or minimally processed vegetables and fruits exposed to pathogenic contamination can cause foodborne illnesses. While foodborne pathogens can be introduced all along the supply chain (e.g., during growing, harvesting, packing, processing, and storing), in this study we use irrigation water as the stochastic contaminant origin. Contaminated irrigation water is one of the major causes of foodborne illnesses and has been singled out for additional regulation in Food Safety Modernization Act (FDA, 2015).

Following numerous foodborne disease outbreaks, FSMA was introduced in 2011 (FDA, 2015) and went into effect in 2016. FSMA aims to ensure food safety in the U.S. supply chain by preventing rather than responding to food contaminations. The regulation includes standards to prevent produce contamination in various stages of production such as irrigation, harvesting, and packing. FSMA explicitly identifies irrigation water as a culprit and uses *Generic E. coli* as the microbial water quality indicator (FDA, 2015). Over 90% of animal- and human-based fecal and non-fecal sources contain *Generic E. coli* (FDA, 2014). Following the FDA, we use *Generic E. coli* as an indicator microorganism to assess the microbial contamination in irrigation water.

We focus on the lettuce industry as a case study for two reasons. First, lettuce is consumed fresh, without cooking or processing that can eliminate foodborne pathogens. Second, lettuce has frequently been associated with foodborne illness outbreaks. Between 1998 and 2019, approximately 36 outbreaks were linked to lettuce produced in California and Arizona (USDA, 2022). For instance, in 2019, an outbreak of *E. coli* that infected 62 people and hospitalized 25 individuals was traced back to romaine lettuce irrigated with contaminated water in California (CDC, 2019). Similarly, another *E. coli* outbreak in 2018 that caused 210 illnesses, 36 hospitalizations, and 5 deaths was traced to romaine lettuce irrigated with pathogenic water (CDC, 2018).

We use a stochastic partial equilibrium model (McCarl and Spreen, 1980) combined with an illness dose-response specification (Lichtenberg, 2010). The economic model is a short-run analysis of the benefits of traceability in which farmers cannot respond to foodborne outbreaks by changing planting and harvesting decisions. Therefore, the results provided here may be an overestimation of the traceability benefits. The model maximizes consumer and producer surplus with demand response to foodborne outbreaks minus cost of *E. coli*-related foodborne illnesses over a 60-day timeframe. The model incorporates several constraints such as daily national demand and supply balance, land use, yield and irrigation equations, stochastic water quality relations, harvest and storage constraints, traceability and produce removal specifications, and illness dose-response equations.

2. Methods and Materials

A stochastic price endogenous mathematical programming model (McCarl and Spreen, 1980) is used to examine the tradeoffs between avoided costs of food contamination and the combined impacts on consumers and producers from implementing traceability in the lettuce industry. On the one hand, the removal of produce from the supply chain hurts consumer and producer surplus. Depending on the magnitude of the recalls, consumers may experience higher prices while suppliers lose revenue. On the other hand, removal of contaminated produce from the supply chain reduces the number of foodborne illness infections and exposures. The model combines a price endogenous partial equilibrium specification with illness dose response relations adopted from Pang et al. (2017). *E. coli* concentration in irrigation water is a random water quality indicator.

California and Arizona are major lettuce producing regions in the U.S. (USDA, 2018). In 2017, nearly 10 billion pounds of lettuce were produced on 67 thousand planted acres in forty-three counties in California and Arizona (USDA, 2018). Lettuce (like other fresh produce) is produced by vertically integrated producers and shippers who grow, harvest, distribute, and sell lettuce year-round to supermarkets, bagged salad firms, and food-service providers (Buzby, 2003). Monterey, Imperial, and Fresno counties in California and Yuma County in Arizona are the largest lettuce producing areas.

2.1. Stochastic Partial Equilibrium Model

We focus on estimating the benefits of traceability, which are compared to the implementation costs obtained from the FDA (2020). The benefits are estimated as the difference in social welfare with and without traceability. The objective function maximizes social welfare including the sum of consumer and producer surplus less the social monetary cost of foodborne illnesses:

$$\text{SW}=\frac{1}{N}*\sum_{n,t}\left[\int p_{n,t}^{d}\left(x_{n,t}^{d}*\delta_{n,t}\right)dx_{n,t}^{d}-\int p_{n,t}^{s}\left(x_{n,t}^{s}\right)dx_{n,t}^{s}\right]-\beta\ *\frac{1}{N}*\sum_{n,t,k,w,d} i_{n,t-k,t,w,d} \quad (1)$$

Where, SW is social welfare and $p_{n,t}^{d}\left(x_{n,t}^{d}*\delta_{n,t}\right)$ and $p_{n,t}^{s}\left(x_{n,t}^{s}\right)$ represent inverse demand and supply functions on day *t*, respectively. $x_{n,t}^{d}$ and $x_{n,t}^{s}$ are daily aggregate demand and supply of lettuce in state of nature *n*.[5] $\delta_{n,t}$ is demand shifter for the negative demand response to foodborne outbreaks (Toledo and Villas-Boas, 2019; Bovay and Sumner, 2017; Arnade et al., 2009).[6] The

[5] The empirical analysis includes 500 random draws (N=500).

[6] Bovay and Sumner (2017) and Arnade et al. (2009) document that foodborne illness incidents decrease demand. Toledo and Villas-Boas (2019) show that foodborne outbreaks decrease demand even in the presence of traceability

expression in the square bracket is the area between demand and supply curves and represents consumer and producer surplus. β is the annual social cost per case of foodborne illness and $i_{n,t-k,t,w,d}$ is the number of individuals who consumed contaminated lettuce on day t that was harvested on day t-k in each state of nature, irrigated with water source w (ground or surface), and in production district[7] d. k is average number of lettuce storage days between harvest and consumption. The model is solved for a time period of t=60 days. Definition of parameters, variables, and their units are provided in appendix Tables (A.1) and (A.2).

The objective function is maximized subject to the following constraints:

$$x_{n,t}^{d} \leq x_{n,t}^{s} \qquad \forall n,t \qquad (2)$$

$$x_{n,t}^{s} = \sum_{d} xc_{n,t,d}^{s} + IS_{t} - \frac{NX}{365} \qquad \forall n,t \qquad (3)$$

$$xc_{n,t,d}^{s} = \sum_{k,w} s_{n,t-k,t,w,d} \qquad \forall n,t,d \qquad (4)$$

$$s_{n,t,t+1,w,d} = \frac{1}{2} * \left(a_{t,w,d} * y_{d}\right) \qquad \forall n,t,w,d \qquad (5)$$

$$s_{n,t,t+k,w,d} + \alpha * r_{n,t,t+k,w,d} = \frac{1}{2(k-1)} * \left(a_{t,w,d} * y_{d}\right) \qquad \forall n,t,w,d,k > 1 \qquad (6)$$

$$a_{t,w,d} = \frac{1}{365} * sh_{w,d} * A_{d} \qquad \forall t,w,d \qquad (7)$$

Equation (2) shows the daily lettuce supply and demand balance, where daily national demand for lettuce cannot exceed its daily national supply. Equation (3) restricts the daily national supply

---

using a scanner dataset from a large grocery chain and three California Salmonella related egg recalls in 2010 as a case study. They show that recalls reduce egg sales even when all contaminated eggs can be traced and differentiated from safe eggs.

[7] In order to reduce dimensionality of the model, we aggregate the smallest lettuce producing counties in California and Arizona into one production district. To allow for heterogeneity of *E. coli* CFUs in irrigation water and take advantage of USGS and EPA water quality data, we disaggregate Yuma and Fresno counties into 9 and 7 production districts, respectively.

of lettuce to be no greater than the aggregated district-scale supply ($xc^{s}_{n,t,d}$) plus initial storage (IS_t)[8] minus net export ($NX/365$).[9] Supply of produce comes from district d on day t from lettuce stored on day t-k ($s_{n,t-k,t,w,d}$) aggregated over k (equation 4).

Following Cai et al. (2013), we assume that demand for lettuce is positively correlated with freshness and hence negatively correlated with storage (Wang and Li, 2012). Therefore, we assume that demand for lettuce is highest during the earlier days of storage. Specifically, we assume that 50% of lettuce that enters storage on day t is consumed on the subsequent day ($t+1$) (equation 5). The remaining 50% of lettuce from day t is consumed in equal shares ($1/[k-1]$) over the produce shelf life, k.[10] The lettuce harvested on day t is the product of yield (y_d) and acreage ($a_{t,w,d}$). In equation (6), α is an indicator parameter for traceability, with $\alpha = 0$ meaning no traceability and $\alpha = 1$ referring to traceability. $r_{n,t,t+k,w,d}$ is the amount of stored lettuce that will be discarded due to foodborne outbreaks if the produce is contaminated and traceability is in place. The formulation in (6) ensures that produce can be removed from the supply chain only when tracing is implemented ($\alpha = 1$). We assume that lettuce cannot be removed from the supply chain on the first day of storage. This means that consumption of produce on the first day of storage is fixed and not removed even if tracing is implemented. We suppose that it takes at least one day for consumers to get sick from contaminated lettuce. After the first day, the exposed individuals will start to show signs of illnesses and traced contaminated produce will be removed from supply chain to prevent exposure on days $t+k$ (k>1).

On each day, an equal proportion (1/365) of annually planted acreage is assumed to be planted and harvested (equation 7). In this sense, the current model represents a short run analysis of traceability in the lettuce supply because planted acreage (A_d) is fixed. $sh_{w,d}$ is the proportion of acres irrigated with surface or ground water.

---

[8] Initial storage is the amount of lettuce that was stored in the days prior to day t and is non-zero only for t$\leq |k|$. For example, if lettuce is stored for 2 days, only IS_1 and IS_2 are nonzero.

[9] Daily next export is expressed as 1/365 times the annual net export.

[10] In sensitivity analysis we assume that an equal proportion of lettuce harvested on day t is consumed on each of the subsequent k days. In other words, equations 5 and 6 are replaced with $s_{n,t,t+1,w,d} = \frac{1}{k} * (a_{t,w,d} * y_d)$ and $s_{n,t,t+k,w,d} + \alpha * r_{n,t,t+k,w,d} = \frac{1}{k} * (a_{t,w,d} * y_d) \;\; \forall n,t,k,w,d,k > 1$.

2.2. Illness Dose-response and Traceability Specifications

In this study, we focus on irrigation water quality as a pathway for produce contamination. We assume that lettuce is irrigated every 6 days (Smith et al., 2011). We focus on 5 last irrigation events prior to harvest on day *t*. Earlier irrigation events are not included because pathogens die off before harvest (Weller et al., 2017) and do not cause foodborne illness.

The concentration of *Generic E. coli* Colony Forming Units (CFUs) per 100 ml of irrigation water ($C_{n,t,w,g,d}$) is obtained using Lognormal or Weibull distribution functions (8).[11]

$$C_{n,t,w,g,d} = Max\{(0, \Omega(k_1, k_2))\} \qquad \forall n, t, w, g, d \tag{8}$$

where Ω is *E. coli* distribution function (Lognormal or Weibull). k_1 is the scale and k_2 is the shape parameter, respectively. g refers to the last five irrigation events.

Equation 9 estimates crop pathogen content as a function of *Generic E. coli* in irrigation water. CFUs from all irrigation events are adjusted by microbial die-off between irrigation events and harvest time. We assume that 70% of *Generic E. coli* in source water remains in irrigation water when it reaches the agricultural field ($\eta = 0.7$). Transmission of *E. coli* to crop is also a function of the volume of water consumed by a crop or the irrigation efficiency (λ) (Solomon et al., 2002).[12] Irrigation efficiency depends on the type of irrigation technology. Most lettuce producers in California and Arizona use pressurized furrow irrigation as the primary irrigation technology (FDA, 2016).

$$CN_{n,t,w,d} = \sum_g \left(\frac{C_{n,t,w,g,d}}{100}\right) * \lambda * \pi * \theta * e^{-\frac{l(g)^{\zeta}}{\epsilon}} * \eta \qquad \forall n, t, w, d \tag{9}$$

where $CN_{n,t,w,d}$ is the CFU of pathogens, such as *E. coli* O157:H7, per gram of lettuce. *E. coli* O157:H7 is a strain of *E. coli* that causes most of the severe foodborne illnesses (CDC, 2020). π

---

[11] Lognormal or Weibull distribution functions are selected using water quality data obtained from the EPA STOrage and RETrieval (STORET) and the USGS National Water Information System (NWIS) Database and Akaike Information Criteria (AIC), density plots, Q-Q plots, and P-P plots.

[12] Irrigation efficiency is defined as the ratio of evapotranspiration and the amount of per acre applied irrigation water (Solomon et al., 2002).

is harmful pathogens such as *E. coli* O157:H7 per CFU of *Generic E. coli* in irrigation water (Pang et al., 2017; Ottoson et al., 2011; Muniesa et al., 2006). In the sensitivity analysis we consider alternative scenarios with greater pathogenicity of irrigation water per unit of *Generic E. coli*. ($e^{-\frac{l(g)^{\zeta}}{\epsilon}}$) is the microbial decay function (Brouwer et al., 2017) and quantifies the die-off of *E. coli* from each irrigation event to harvest. $l(g)$ is time interval between irrigation event g and harvest. θ is water holding capacity of lettuce (Ottoson et al., 2011).

Equations (10), (11), and (12) establish the illness dose-response specifications adopted from Pang et al. (2017). In equation (10) serving size (ϱ) is multiplied by $CN_{n,t,w,d}$ to estimate pathogen dose ($DO_{n,t,w,d}$) per *E. coli*-contaminated lettuce serving. Dose-response relation in equation (11) estimates the probability of illness ($PR_{n,t,w,d}$) per contaminated lettuce serving. The probability of illness per *E. coli*-contaminated serving is used to estimate the expected number of illnesses $i_{n,t-k,t,w,d}$ (equation 12).[13]

$$DO_{n,t,w,d} = CN_{n,t,w,d} * \varrho \qquad \forall n,t,w,d \qquad (10)$$

$$PR_{n,t,w,d} = 1 - \left(1 + \frac{DO_{n,t,w,d}}{\omega}\right)^{-\rho} \qquad \forall n,t,w,d \qquad (11)$$

Following Pang et al. (2017), we assume that each day of storage (k) reduces contaminants in lettuce by a microbial die-off rate equal to $k * \mho$, where $\mho$ is the microbial die-off rate (CFU per day). *E. coli* in lettuce decays if stored at temperature below 5ºC but grows at higher temperatures (Pang et al., 2017; McKellar and Delaquis, 2011). The FDA recommends a storage temperature of 5ºC or lower for leafy greens (FDA, 2010). Following the FDA's recommendation, we assume that storage temperature does not exceed 5ºC, hence, *E. coli* CFUs decline during storage.

$$i_{n,t-k,t,w,d} = PR_{n,t-k,w,d} * s_{n,t-k,t,w,d} * e^{-(\mho * k)} * \frac{50{,}802.3}{\varrho} \qquad \forall n,t,k,w,d \qquad (12)$$

[13] The unit of lettuce yield is hundredweights (CWT), while the serving size is expressed in grams. Hence, we multiply $s_{n,t-k,t,w,d}$ by 50,802.3 (grams/CWT) to obtain the number of lettuce servings.

In equations (13) and (14), we use an indicator variable ($\gamma_{n,t,t+1,w,d}$) to identify an outbreak. The FDA defines a foodborne outbreak as an event that involves two or more people getting sick from consuming the same contaminated product (FDA, 2022). Following this definition, we assume that if lettuce supplied on day *t* causes more than 2 cases of illnesses in the subsequent day (*t+1*), then lettuce is contaminated and can be removed from the supply chain if traceability is implemented. Equations (13) and (14) force γ to be one if there are more than two illnesses, and zero otherwise.

Equation (15) indicates that contaminated lettuce cannot be discarded from the supply chain unless there is a foodborne outbreak. The removal of affected lettuce depends on the marginal costs (consumer and producer surplus loss) and marginal benefits (avoided foodborne illnesses) of removing the contaminated lettuce. However, removal of affected lettuce is only possible if traceability system is implemented.[14]

$$\alpha * i_{n,t,t+1,w,d} \leq \alpha * (2 + M * \gamma_{n,t,t+1,w,d}) \qquad \forall n, t, w, d \qquad (13)$$

$$i_{n,t,t+1,w,d} \geq \alpha * 2 * \gamma_{n,t,t+1,w,d} \qquad \forall n, t, w, d \qquad (14)$$

$$r_{n,t,t+k,w,d} \leq \mathrm{U} * \alpha * \gamma_{n,t,t+k,w,d} \qquad \forall n, w, d, and\ k > 1 \qquad (15)$$

where M and U are sufficiently large numbers such that constraints 13 and 15 are not binding when γ is one.

[14] We call this tracing and removal system a "smart" traceability where the removal of contaminated lettuce is an endogenous decision variable and depends on the marginal costs and benefits of removing affected lettuce. We also ran the model with a "myopic" tracing where we assume that all traced lettuce must be removed from the supply chain following a foodborne outbreak. In other words, we add the following equation to the model setup:

$$s_{n,t,t+k,w,d} \leq U * (1 - \alpha * \gamma_{n,t,t+k,w,d}) \qquad \forall n, t, w, d, and\ k > 1$$

The results under "myopic" tracing and removal assumptions are similar to the "smart" tracing and available upon request.

2.3. Demand Response Relations

Foodborne outbreaks lead to negative demand shocks (Toledo and Villas-Boas, 2019; Bovay and Sumner, 2017; Arnade et al., 2009). The demand response is proportional to the number of foodborne illnesses caused by consumption of contaminated lettuce on the first day of storage (equation 16). We assume that contaminated lettuce supplied on day t and consumed on day $t+1$ will reduce national demand on 5 succeeding days ($t+j$). $\aleph$ is the median reduction in demand per case of foodborne illness in each state of nature and day t and is obtained from Bovay and Sumner (2017) and Arnade et al. (2009).

Bovay and Sumner (2017) and Arnade et al. (2009) estimate that a foodborne outbreak with 204 illnesses causes a 6.9% decline in demand. Using this estimate the reduction in demand is 0.034% per case of foodborne illness. Shuval et al. (1999) estimate that the ratio of *E. coli*-related clinical illnesses to the number of infected individuals is 1 to 100. In addition, Scharff et al. (2016) and Scallan et al. (2011) estimate that the ratio of reported cases to total number of symptomatic people is 1 to 26. We divide the demand shift parameter (0.034% per case of illness) by 26*100 and approximate the demand response ($\aleph$). We then multiply $\aleph$ by the daily number of illnesses to obtain a demand response ($\delta_{n,t+j}$) (equation 16).

$$\delta_{n,t+j} = \aleph * \sum_{w,d} i_{n,t,t+1,w,d} \qquad \forall n, t, j \in \{1,2,3,4,5\} \qquad (16)$$

3. Data

The baseline year for the partial equilibrium model is 2016. We use observed price-quantity points in 2016 and own-price elasticities of demand and supply to derive inverse demand and supply as linear functions of lettuce quantity. Lettuce own-price elasticity of demand and supply come from Okrent and Alston (2012) and Lohr and Park (1992), respectively. We use USDA Economic Research Service (ERS) and National Agricultural Statistics Service (NASS) data for consumption, production, and price of lettuce. Other data including planted acreage, net export, and water used in irrigation for lettuce come from USDA ERS and NASS, as well.

USDA (2019) estimates that cost of per case foodborne illness treatment due to *E. coli* O157:*H7* is $8,500. Hammitt and Haninger (2007) show that WTP to decrease risk of foodborne diseases is between ten and twenty thousand dollars per case of foodborne illness. They use a survey of respondents selected from the Knowledge Networks panel. The respondents were asked to choose between a safer and conventional food product. Hammitt and Haninger (2007) estimates include WTP to reduce the risk of all microbial pathogens. We use USDA (2019) estimates of per case foodborne illness damages as the baseline value and include estimates from Hammitt and Haninger (2007) in sensitivity analysis to investigate the influence of illness severity.

Foodborne pathogens can contact produce all along the supply chain (e.g., during growing, harvesting, packing, processing, and storing). However, using contaminated irrigation water during growing season is one of the major culprits of foodborne outbreaks (FDA, 2021). Hence, in this study we focus on pathogen introduction during irrigation as the first contamination stage in the stochastic model simulation.

County-scale past spatial data for *Generic E. coli* content in ground and surface water are obtained from the EPA STOrage and RETrieval (STORET) and the USGS National Water Information System (NWIS) Database (USGS-EPA, 2020). The maximum and minimum number of observations are 5,931 and 32 in Ventura and San Joaquin counties, respectively. The county-specific distribution functions are selected using Akaike Information Criteria (AIC), density plots, Q-Q plots, and P-P plots. The selected functions are Weibull or Log-normal distributions.

STORET and NWIS do not report data for *Generic E. coli* in 17 counties in California. We use the values from adjacent counties to obtain linearly interpolated values for the counties with no data. Also, STORET and NWIS databases contain data for *E. coli* CFUs in groundwater in only three counties. For these counties Weibull distributions are selected using AIC, density plots, Q-Q plots, and P-P plots. The density functions for counties with missing *E. coli* data in groundwater are estimated by shifting their surface water distributions to the left. The shift parameter is obtained as the average ratio of mean number of *E. coli* in ground and surface water distributions in the counties with available data for both water sources.

All production districts in a county are assumed to have the same density functions. The model is solved with five hundred random draw simulations for microbial water quality data. Mean values of *E. coli* CFU/100 ml across counties are provided in Figure (A.1).

4. Results

We validate the model using calibrated inverse demand and supply functions and observed acreage data in 2016. Model solution produces daily lettuce quantities and prices that are close to the observed data in 2016 (appendix Table A.3). These solutions provide a solid benchmark for the subsequent scenario analysis.

We first provide baseline annual results. Annual values are obtained by extrapolating the results of the 60-day time framework model. In the baseline case, monetary value of foodborne illness damages (β) is $8,500. Pathogenicity per unit of *Generic E. coli* (π) is 1%. The pathogens transmission rate from water to crop through irrigation (η) is 70%. The daily microbial die-off rate (Ʊ) is 31%. We then present the results for sensitivity analysis where we allow β, π, Ʊ, and η to vary from their baseline values to examine the influence of these parameters on the value of traceability. We also alter the assumption about the amount of lettuce produced on day t and consumed on day $t+k$. All results are obtained using planted acreage, prices, and quantities observed in 2016.[15] Hence, the model solutions represent a short-term market equilibrium with fixed acreage and production.

Following Cai et al. (2019), we assume that lettuce has a shelf life of up to 16 days. Therefore, we provide alternative scenarios where shelf life varies between 2 and 16 days. We do not present the results for one day of shelf life because traceability is pointless if all lettuce is consumed within a day of harvest.

In all scenarios, the benefit of produce tracing is calculated by comparing the value of social welfare with and without traceability. Similarly, the number of avoided exposures is obtained by comparing the average number of exposures with and without traceability.

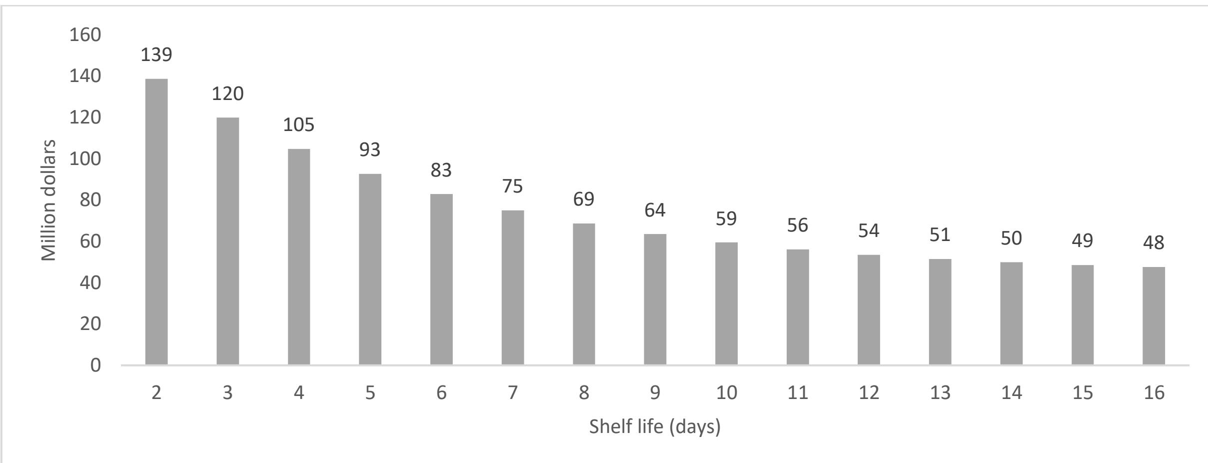
4.1. Baseline Results

We present the benefits of traceability for various scenarios of lettuce storage in Figure 1. The value of produce tracing is a function of number of avoided exposed individuals to foodborne illnesses and consumer and producer surplus loss. Removal of all produce originating from the affected district can decrease consumer and producer surplus. Consumers may experience higher prices while suppliers lose revenue depending on the magnitude of the recall. On the other hand, removal of contaminated produce from the supply chain reduces damages associated with foodborne illness infections.

[15] 2016 is the most recent year with the highest observed lettuce production.

The baseline value of produce tracing in the lettuce industry is the highest when lettuce is stored for two days ($139 million). These benefits include consumer and producer surplus and are due to avoided exposures to foodborne illnesses including symptomatic and asymptomatic cases. An increase in storage of lettuce results in a lower traceability value because of greater decay of pathogens with longer storage prior to consumption. The longer the storage, the higher the die-off and the lower the number of foodborne illness exposures. For instance, if average lettuce storage is 16 days, the benefit of tracing is $48 million.

Figure 1: Annual benefits of produce tracing in the lettuce industry

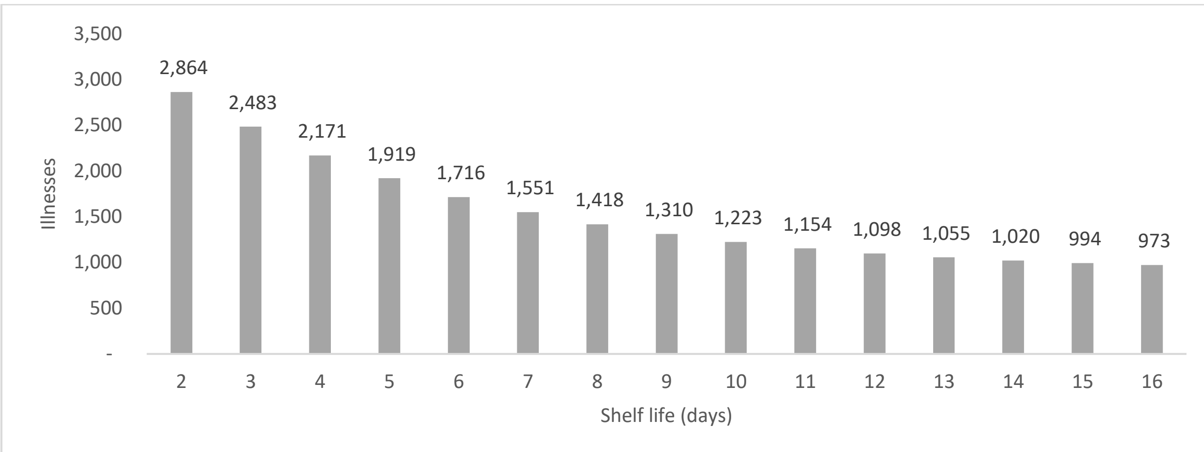


The estimated benefits of produce tracing can be compared with the existing estimated costs. The FDA estimates that the annual cost of compliance with the "Requirements for Additional Traceability Records for Certain Foods" proposed rule for products on the Food Traceability List will be between $411 and $535 million under full and partial exemption for small retail food entities scenarios, respectively (FDA, 2020). Adjusting this estimate using the ratio of lettuce value of production relative to total production value of food products included on the FTL, the traceability cost for the lettuce industry is between $29 and $37 million annually.[16] The estimated baseline benefits of traceability in this study are greater than the estimated costs by the FDA (2020).

[16] If we adjust this estimate using the ratio of harvested acreage of lettuce relative to total harvested area of fresh fruits and vegetables, the cost for the lettuce industry is between $29 and $37 million annually.

The number of prevented exposures to foodborne diseases (including symptomatic and asymptomatic cases) due to traceability in the lettuce supply chain are presented in Figure 2. The number of avoided exposures to foodborne pathogens includes asymptomatic, mild discomfort, and life-threatening cases. The number of avoided exposed cases is the highest when lettuce is stored for 2 days (2,864 cases). If storage of lettuce is 16 days, the number of avoided exposed cases due to traceability decreases significantly. As the shelf life increases, the number of prevented exposures to foodborne illnesses decreases.

Figure 2: Annual number of avoided exposures to foodborne illnesses with traceability



4.2. Sensitivity Analysis

This section provides the results for sensitivity analysis. We vary monetary value of foodborne illness damages (β), the pathogenicity of water per unit of *Generic E. coli* (π), the pathogens transmission rate from water to crop through irrigation (η), and the daily microbial die-off rate (Ʊ). We examine the effect of these parameters on the benefits of traceability in the lettuce supply chain. We also provide the results of sensitivity analysis where we vary the assumption about the amount of lettuce produced on day t and consumed on day $t+k$.

4.2.1. Monetary Value of Foodborne Illness Damages

We vary the per case cost of foodborne illness from its baseline value ($8,500) to examine the impact of foodborne illness severity on traceability value. We consider two scenarios where the cost of illness is lower than the baseline ($5,000 and $7,000) and two scenarios where the cost of illness is greater than the baseline ($10,000 and $12,000).

Figure 3: Annual benefits of tracing for alternative foodborne illness severity scenarios

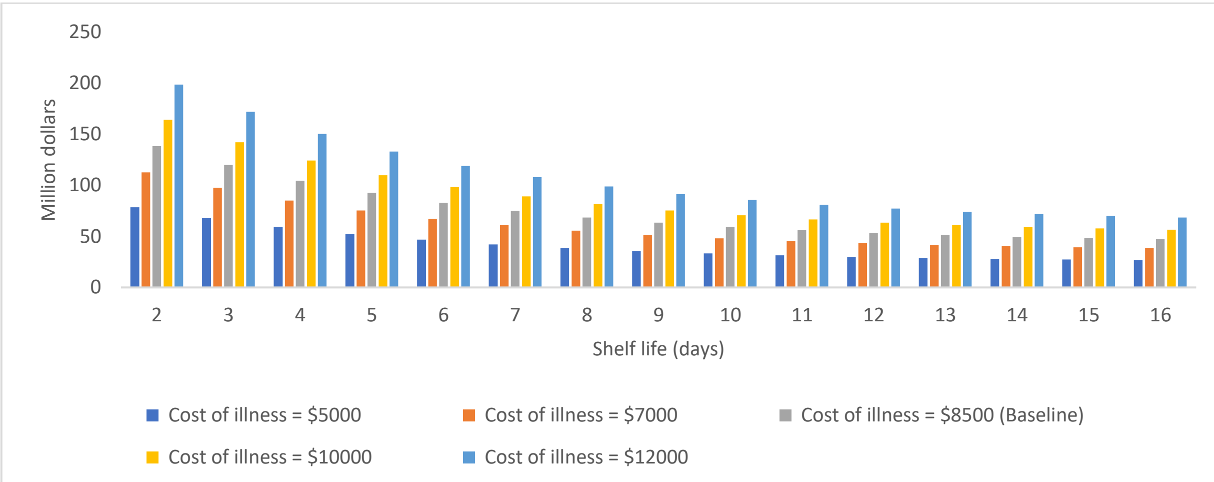
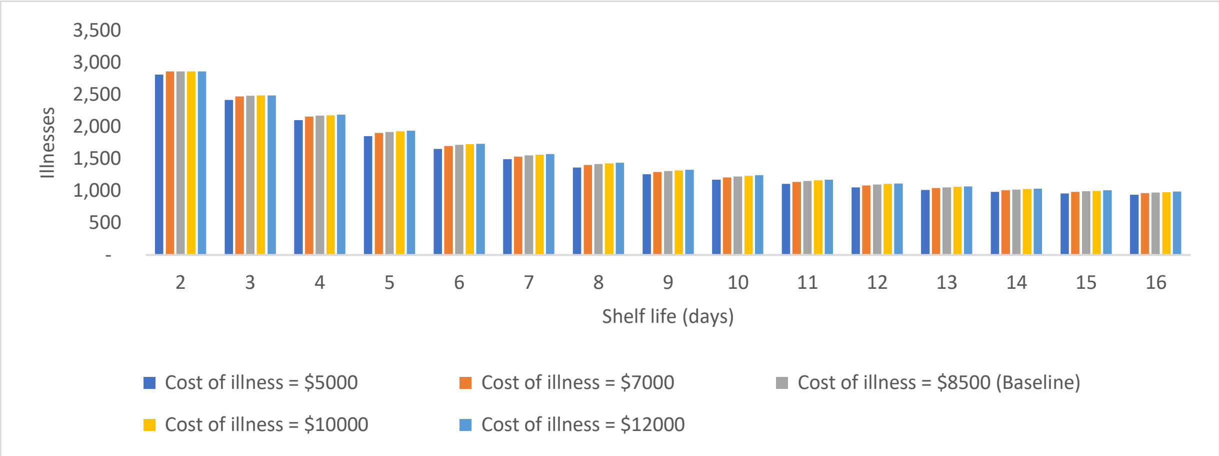


Figure 3 shows the benefits of tracing in the lettuce supply chain under alternative scenarios of foodborne illness severity. Results indicate that benefits of tracing vary between $27 to $199 depending on illnesses severity and storage. For each scenario of storage, the value of traceability increases as the cost of illness treatment rises. For instance, with 15 days of storage, the benefit of tracing rises from $27 (with cost of illness equal to $5,000) to $70 million (with cost of illness equal to $12,000). An increase in the cost of illness is translated into greater economic damages due to the foodborne illness exposures. Therefore, preventing more costly cases via traceability results in greater value of traceability. The value of tracing is not very sensitive to the severity of illnesses. This is because we use a short-run model in which farmers cannot alter their planting and harvesting decisions in response to the outbreaks.

Figure 4 shows the number of avoided foodborne exposures for alternative monetary values of illness. The number of prevented exposures depends mainly on storage and to smaller extent on the severity of illness. For example, if per case cost of illness is $8,500, the number of avoided illness exposures with 2- and 16-days storage is 2,864 and 973, respectively. Similar to the baseline results, an increase in storage of lettuce reduces the number of avoided exposures to foodborne outbreaks due to greater microbial decay. The number of avoided exposures increases slightly as the monetary value of foodborne damages increases. This is so, because the higher the cost per case of illnesses, the more substantial the economic damages imposed by these illnesses. Therefore, with severe illnesses more cases are prevented to reduce the damages.

Figure 4: Annual number of avoided exposures to foodborne illnesses for alternative foodborne illness severity scenarios



4.2.2. Microbial Die-off Rate

We vary the daily microbial die-off rate (U) from its baseline value (31%) to assess the impact of pathogen die-off on the value of traceability. We consider two scenarios with lower die-off rates (0 % and 12%) and two scenarios with higher rates (36 % and 48%) than the baseline. Figure 5 shows that the greatest benefit of tracing in the lettuce industry is $270 million (with 0% die-off rate). The lowest benefit of tracing ($27 million) corresponds to a die-off rate equal to 48% and 16 days of storage. A higher die-off rate implies a lower survival of foodborne pathogens and fewer exposures to illnesses. Die-off rate affects the benefits of tracing in each of the storage length scenarios. For instance, with 15 days of storage, an increase in die-off rate from 0% to 48% causes a ten-fold decrease in the value of traceability. If the die-off rate is zero, the value of traceability does not change across different scenarios of storage.

Figure 5: Impact of microbial die-off rate (Ʊ) on the annual benefits of traceability

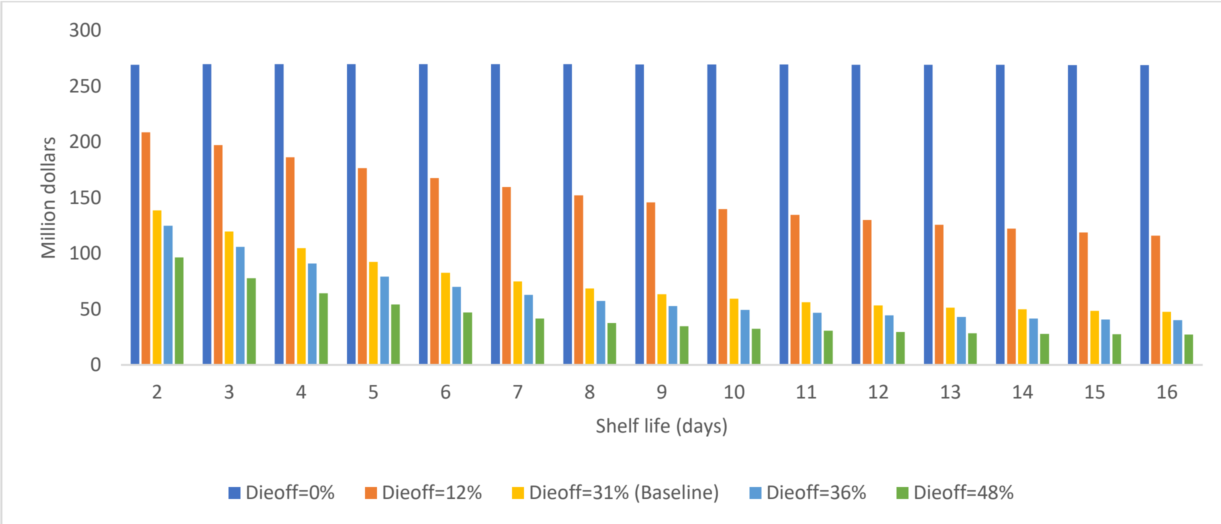
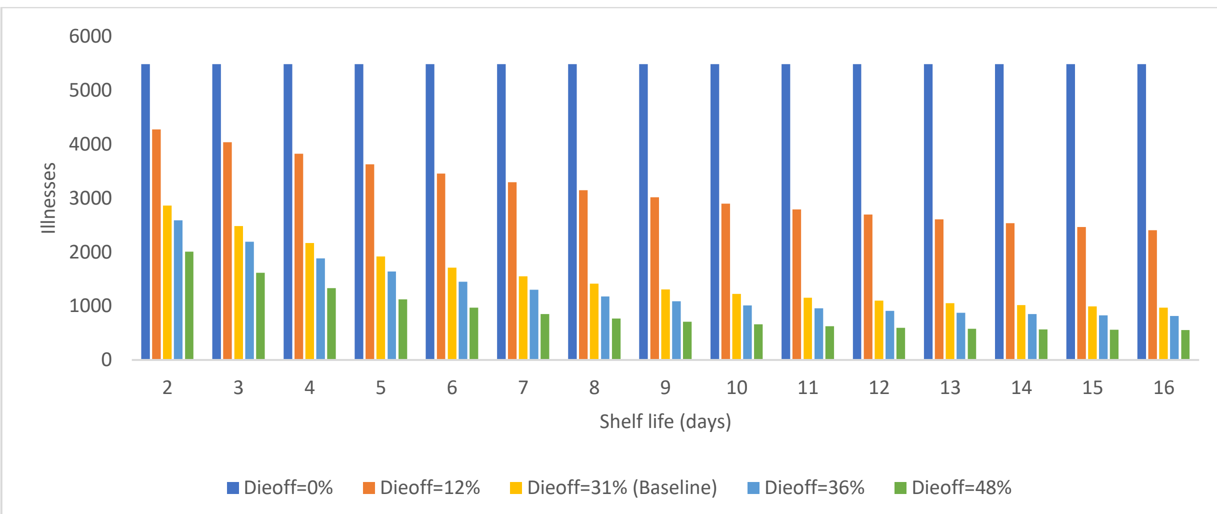


Figure 6: Annual number of avoided exposures for various scenarios of microbial die-off rate (Ʊ)



The number of avoided foodborne illness exposures for scenarios of pathogen die-off rate are provided in Figure 6. As expected, the number of prevented foodborne related exposures decreases with higher die-off rate. For instance, in a scenario with 16 days of storage and microbial decay rate equal to 48%, the number of avoided exposed people due to traceability is 553. If the die-off rate decreases to 12%, the number of avoided exposures to foodborne pathogens increases to 2,407 individuals for the same storage scenario.

4.2.3. Pathogenicity per Unit of *Generic E. coli* (π)

We vary the pathogenicity per unit of *Generic E. coli* from its baseline value (1.0%) to assess the impact on the value of traceability. We consider two scenarios with lower ratio (0.4% and 0.7%) and two scenarios with higher ratio (3% and 5%) than the baseline. Figure 7 shows that, as expected, the greatest economic benefit of tracing in the lettuce industry ($770 million) is when π is the highest and average storage is the lowest at 5% and 2 days, respectively. This value corresponds to 15,577 avoided exposed people to foodborne illnesses (Figure 8). The lowest benefit of tracing ($9 million) is with π and storage equal to 0.4% and 16 days, respectively. A greater pathogenicity per unit of *Generic E. coli* results in larger probability of foodborne illnesses which in turn leads to a greater number of avoided exposures due to traceability. The results show that the benefits of tracing are highly sensitive to pathogenicity of produce. For instance, with 15 days of shelf life, the benefits vary between $10 ($\pi$ = 0.4%) and $181 million ($\pi$ = 5%).

Figure 7: Impact of pathogenicity per unit of *Generic E. coli* ($\boldsymbol{\pi}$) on the annual benefits of traceability

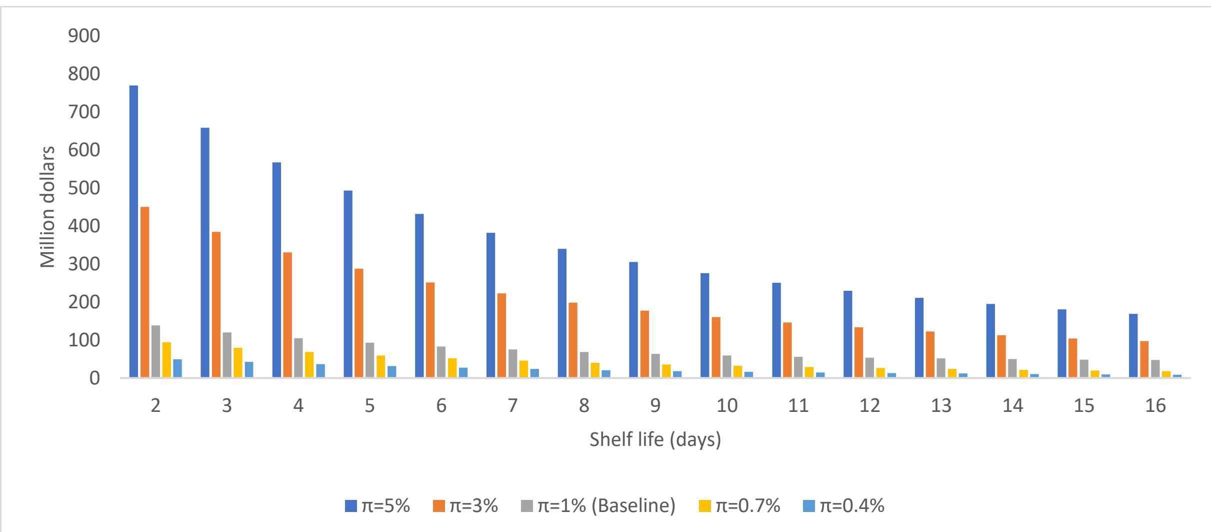
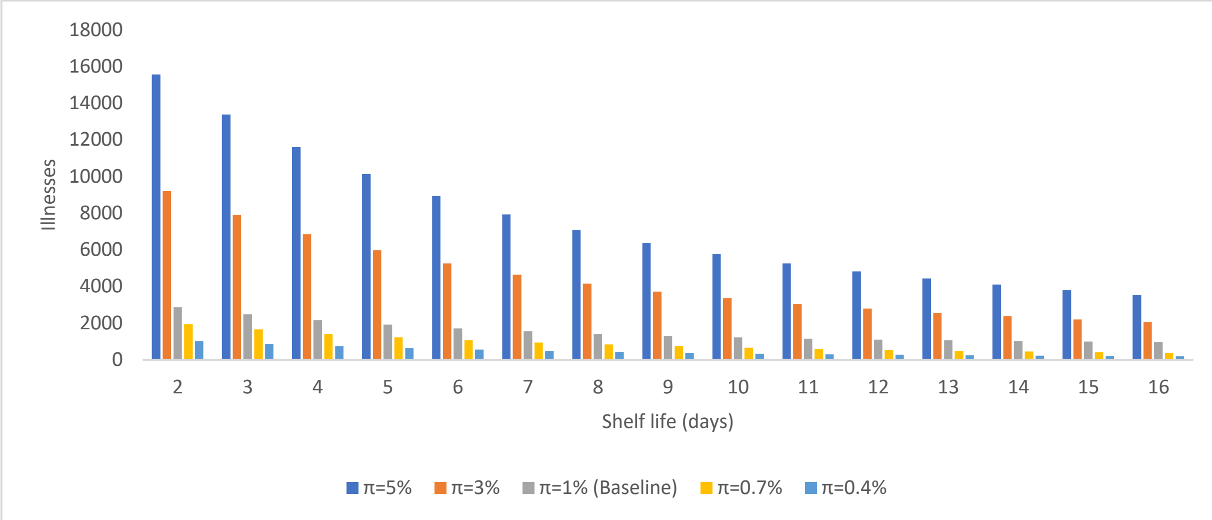


Figure 8: Annual number of avoided exposures to foodborne illnesses for alternative scenarios of pathogenicity per unit of *Generic E. coli* ($\boldsymbol{\pi}$)



4.2.4. Transmission of *E. coli* from Source Water to Crop Through Irrigation (θ)

We vary the transmission of *E. coli* from water to crop through irrigation (θ) to assess its impact on the value of traceability. Two scenarios with lower transmission rate (50% and 60%) and two scenarios with higher transmission rate (80% and 90%) than the baseline are considered. The value of tracing in the lettuce supply chain and the corresponding number of avoided illness exposures are represented in Figures 9 and 10, respectively. Figure 9 shows that an increase in θ results in larger benefits of tracing as expected. For instance, with 15 days of shelf life, the value of tracing is \$33 million when θ is 50%. This value increases to \$64 million as θ increases to 90%. A greater transmission of *E. coli* from source water to crop leads to a larger probability and number of foodborne illness exposures. This in turn is causes a higher value of traceability due to avoiding these exposures. Figure 9 also shows that the value of produce tracing varies between \$33 (16 days of storage and θ equal to 50%) and \$182 (2 days of storage and θ equal to 90%) million. These values for tracing correspond to 3,762 and 666 avoided exposures, respectively (Figure 10).

Figure 9: Impact of transmission of *E. coli* from source water to crop (θ) on the annual benefits of traceability

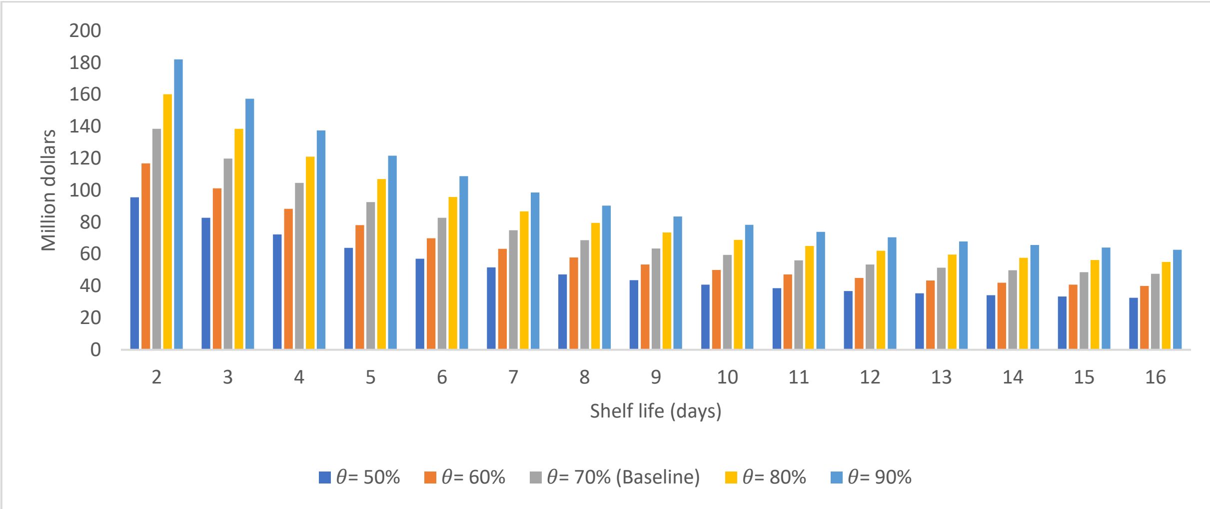
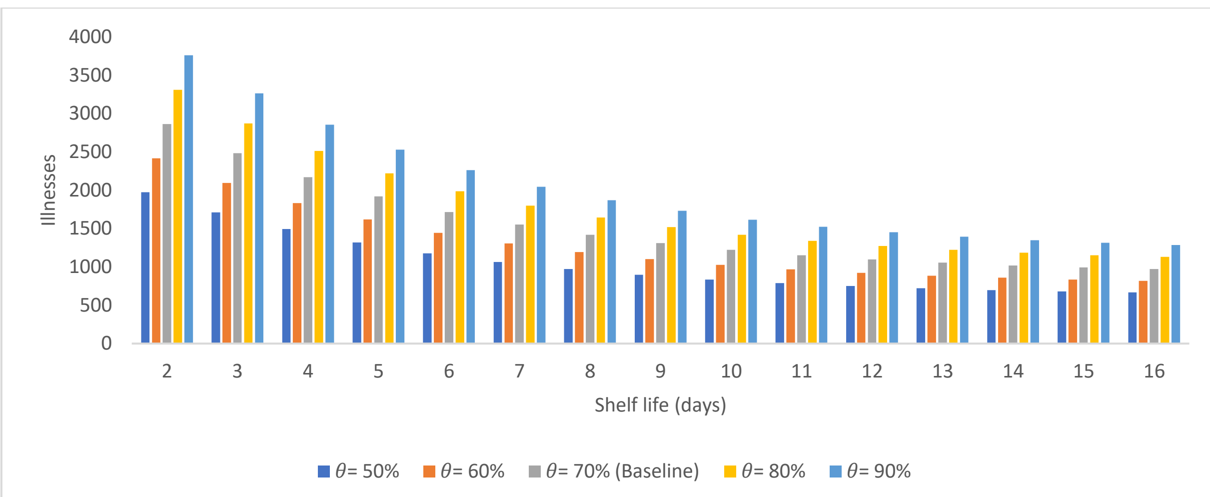


Figure 10: Annual number of avoided exposures to foodborne illnesses in different scenarios of amount of remaining *E. coli* in applied water (θ)



4.2.5. Consumption During Storage

In this section, we alter the consumption distribution for lettuce produced on day t and consumed on day t+k. Recall that in equations 5 and 6 we assumed 50% of lettuce produced on day t is consumed on day t+1 and the remaining 50% is consumed in equal proportions during the rest of storage (*baseline* scenario). In this section, we change the consumption assumption so that lettuce produced on day t is consumed in equal shares on days $t+k$ (*equal-share* scenario). For instance, under the 4-day storage scenario, 25% of lettuce produced on day t is consumed equally on days $t+k$, where k=1,2,3,4.

The results corresponding to this assumption with baseline value for all parameters are provided in Figure 11. This figure shows that the economic benefits of traceability in the *equal-share* scenario is roughly between 30% and 50% higher than its value under the baseline assumption, depending on the shelf life. The *equal-share* scenario implies that less lettuce is consumed on day t+1 and more is stored. This implies that fewer people are exposed on day t+1 and more illness exposures are avoided on days $t+k$ with a tracing system in place. This results in higher value of tracing in *equal-share* scenario relative to the baseline scenario.

Figure 11: Annual benefits of traceability for *equal-share* scenario

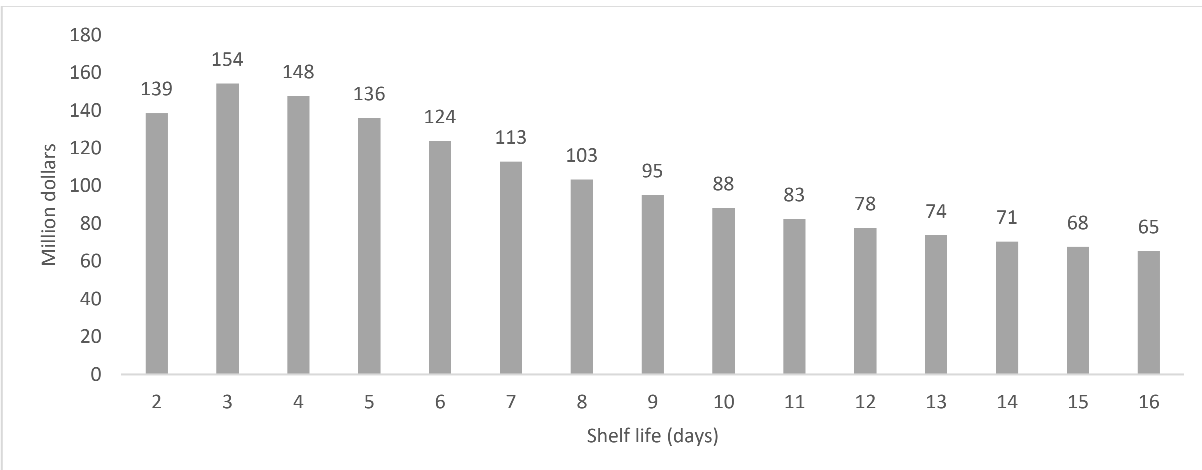


Figure 11 also shows that benefits of traceability increase if storage of lettuce rises from 2 to 3 days and decrease afterwards. Two opposing factors are responsible for the shape of the benefit estimates. On the one hand, the amount of contaminated produce that is consumed versus removed depends on the length of storage. The longer the storage, the less is consumed on the first day of retail and more affected produce is removed from supply chain when traceability is in place and there is an outbreak. This implies that with longer storage traceability enables prevention of more illnesses when storage is longer. As a result, the value of traceability is greater when storage is longer. On the other hand, longer storage implies greater microbial die-off and less foodborne illnesses even without food traceability. Hence, traceability prevents less foodborne illnesses with longer storage than with shorter storage scenarios. As a result, the value of traceability decreases as storage length increases. The results show that for shorter storage scenarios the former effect dominates the latter. However, after 3days of storage the latter effect is more significant then then former.

5. Conclusions

Food traceability is an important instrument for managing foodborne outbreaks. Rapid tracing systems help remove the contaminated produce from the supply chain by identifying the source of contamination. We estimate the economic value of tracing lettuce in the U.S. in terms of preventing foodborne illnesses and examine the sensitivity of the results with respect to key modeling parameters. Our baseline results show that if the average per case foodborne illness is $8,500, then the annual value of traceability in the lettuce industry is between $48 and $139 million, depending on the average storage length of lettuce post-harvest prior to retail. Overall, the value of traceability in the lettuce supply chain is between $9 and $770 million across all considered scenarios. The number of avoided foodborne illnesses due to traceability ranges between 973 and 2,864 cases in the baseline and 184 and 15,577 in all scenarios.

Comparing the estimated traceability benefits in this study against the existing estimated costs would offer some insight into the importance of investment in produce tracing. The FDA estimates that the "Requirements for Additional Traceability Records for Certain Foods" proposed rule will cost between $411 and $535 million annually (FDA, 2020). These estimates are calculated for foods on the Food Traceability List and correspond to full and partial exemption for small retail food establishment scenarios, respectively. Adjusting this estimate using the ratio of lettuce value of production relative to total production value of food products included on the Food Traceability List, the annual traceability cost for the lettuce industry is between $29 and $37 million. Our estimated benefits of traceability exceed the estimated costs by the FDA (2020) in most of the scenarios.

The sensitivity analysis evaluates the benefits of traceability in the U.S. lettuce industry for alternative assumptions of microbial die-off rate, monetary value of foodborne illness damages, transmission of pathogens from source water to crop, and pathogenicity of water per unit of *Generic E. coli*. A 40 percent increase in the severity of foodborne illness damages results in a 44 percent increase in benefits of traceability. Similarly, a 30 percent increase in transmission of pathogens from source water to crop relative to the baseline scenario leads to a 30 percent increase in benefits of traceability. The sensitivity of the results also depends on the average length of storage prior to retail. With 16 days of storage and relative to the baseline, a 5-fold increase in pathogenicity of water per unit of *Generic E. coli* and a 3-fold decrease in microbial die-off generate a 4- and 3-fold increase in the value of traceability, respectively.

We focused on lettuce as one of the common vehicles for foodborne illnesses. However, other leafy greens and fresh vegetables can also cause food contamination. On the one hand, a narrow focus on the lettuce industry is a limitation of this study because it does not account for potential benefits in other markets. On the other hand, by focusing on the lettuce market we are able to examine the tracing system with a greater detail combining the economic framework with a pathogen dose-response specification. The empirical model in this study can be adapted to other fresh vegetables and fruits. The results will vary depending on different production, processing, storage, and handling practices.

Contaminated irrigation water is one of the major culprits of foodborne outbreaks (FDA, 2021). Hence, in this study we focus on contamination from irrigation water. Foodborne pathogens can be introduced to produce all along the supply chain including growing, harvesting, packing, and handling. In this sense, our results are a lower bound for traceability benefits in the lettuce supply chain.

This study has several limitations. First, due to the dimensionality of the model, the analysis focuses on estimating short-run benefits of traceability with fixed production acreage. A long run analysis would consider endogenous production decisions so that additional marginal costs of traceability can affect endogenous acreage decisions. Second, following the FDA's recommendations, this study assumes that lettuce is stored at temperatures below 5ºC (FDA, 2010) and that *E. coli* CFUs decrease with storage (Pang et al., 2017; McKellar and Delaquis, 2011). If lettuce is stored at temperatures exceeding 5ºC, *E. coli* and other pathogens can grow (Pang et al., 2017; McKellar and Delaquis, 2011). Our estimates do not account for the growth of pathogens due to improper storage. The growth of *E. coli* in lettuce can change our estimates for benefits of tracing. Third, we do not account for a potential decrease in the value of produce during storage. We assume that the quality of lettuce does not decline within the recommended storage window and that all lettuce is sold within the recommended storage period. Fourth, we focus on pathogen introduction through contaminated irrigation water to reflect the stochastic nature of lettuce contamination. Although irrigation water has been identified as the origin of some of the recent foodborne illness outbreaks, other pathways for pathogen introduction are possible including shipping and handling. In this sense, the estimates in this study are a lower bound of traceability benefits.

Despite these caveats, the results of this study have an important policy implication. Our estimates indicate that the benefits of traceability are likely to outweigh the costs. The limitations of our estimates imply that the Benefit/Cost ratio is likely to be even greater than what is suggested in our results. This study emphasizes the importance of investment in produce tracing from farm to retail. Better traceability can help the authorities identify foodborne outbreaks in the early stages and prevent foodborne illnesses. The sooner the information about a foodborne outbreak is available, the sooner the appropriate response actions can be taken to minimize economic damages.

References

Arnade, C., Calvin, L., & Kuchler, F. (2009). Consumer response to a food safety shock: The 2006 foodborne illness outbreak of *E. coli O157: H7* linked to spinach. *Applied Economic Perspectives and Policy*, 31(4), 734-750.

Bovay, J., & Sumner, D. A. (2017). Economic effects of the U.S. Food Safety Modernization Act. *Applied Economic Perspectives and Policy*, 40(3), 402-420.

Brouwer, A. F., Eisenberg, M. C., Remais, J. V., Collender, P. A., Meza, R., & Eisenberg, J. N. (2017). Modeling biphasic environmental decay of pathogens and implications for risk analysis. *Environmental Science & Technology*, 51(4), 2186-2196.

Brown, B., Allard, M., Bazaco, M. C., Blankenship, J., & Minor, T. (2021). An economic evaluation of the Whole Genome Sequencing source tracking program in the U.S. *PloS One*, *16*(10), e0258262.

Buzby, J. C. (2003). *International trade and food safety: economic theory and case studies*. United States Department of Agriculture, Economic Research Service. Available at: https://www.ers.usda.gov/publications/pub-details/?pubid=41618, last accessed April 2022.

Cai, S., Worobo, R. W., & Snyder, A. B. (2019). Combined effect of storage condition, surface integrity, and length of shelf life on the growth of Listeria monocytogenes and spoilage microbiota on refrigerated ready-to-eat products. *Journal of Food Protection*, 82(8), 1423-1432.

Cai, X., Chen, J., Xiao, Y., Xu, X., & Yu, G. (2013). Fresh-product supply chain management with logistics outsourcing. *Omega*, 41(4), 752-765.

California Leafy-Greens Marketing Agreements. (2022). https://lgma.ca.gov/, last accessed April 2022.

Centers for Disease Control and Prevention. (2020). *Questions and answers/E. coli*. Available at: https://www.cdc.gov/ecoli/general/index.html, last accessed April 2022.

Centers for Disease Control and Prevention. (2019). *Outbreak of E. coli infections linked to romaine lettuce*. Available at: https://www.cdc.gov/ecoli/2018/o157h7-11-18/index.html, last accessed April 2022.

Centers for Disease Control and Prevention. (2018). *Multistate Outbreak of E. coli O157:H7 Infections Linked to Romaine Lettuce (Final Update)*. Available at: https://www.cdc.gov/ecoli/2018/o157h7-04-18/index.html, last accessed April 2022.

Collart, A. J., & Canales, E. (2022). How might broad adoption of blockchain-based traceability impact the US fresh produce supply chain?. *Applied Economic Perspectives and Policy*, 44(1), 219-236.

Golan, E. H., Krissoff, B., Kuchler, F., Calvin, L., Nelson, K. E., & Price, G. K. (2004). *Traceability in the US food supply: economic theory and industry studies*, the U.S. Washington DC: U.S. Department of Agriculture, Agricultural Economic Report No. 830.

Hammitt, J. K., & Haninger, K. (2007). Willingness to pay for food safety: sensitivity to duration and severity of illness. *American Journal of Agricultural Economics*, 89(5), 1170-1175.

Liao, P. A., Chang, H. H., & Chang, C. Y. (2011). Why is the food traceability system unsuccessful in Taiwan? Empirical evidence from a national survey of fruit and vegetable farmers. *Food Policy*, 36(5), 686-693.

Lichtenberg, E. (2010). Economics of health risk assessment. *Annual Review of Resource Economics*, 2(1), 53-75.

Lin, W., Ortega, D. L., Ufer, D., Caputo, V., & Awokuse, T. (2022). Blockchain-based traceability and demand for US beef in China. *Applied Economic Perspectives and Policy*, 44(1), 253-272.

Liu, R., Gao, Z., Nayga Jr, R. M., Snell, H. A., & Ma, H. (2019). Consumers' valuation for food traceability in China: Does trust matter?. *Food Policy*, 88, 101768.

Lohr, L., & Park, T. (1992). Certification and supply response in the organic lettuce market. *Journal of Agricultural and Resource Economics*, 253-265.

Lusk, J. L., & McCluskey, J. (2018). Understanding the impacts of food consumer choice and food policy outcomes. *Applied Economic Perspectives and Policy*, 40(1), 5-21.

McCarl, B. A., & Spreen, T. H. (1980). Price endogenous mathematical programming as a tool for sector analysis. *American Journal of Agricultural Economics*, 62(1), 87-102.

McKellar, R. C., & Delaquis, P. (2011). Development of a dynamic growth–death model for *Escherichia coli O157: H7* in minimally processed leafy green vegetables. *International Journal of Food Microbiology*, 151(1), 7-14.

Mejia, C., McEntire, J., Keener, K., Muth, M. K., Nganje, W., Stinson, T., & Jensen, H. (2010). Traceability (product tracing) in food systems: an IFT report submitted to the FDA, volume 2: cost considerations and implications. *Comprehensive Reviews in Food Science and Food Safety*, 9(1), 159-175.

Monteiro, D. M. S., & Caswell, J. A. (2009). Traceability adoption at the farm level: An empirical analysis of the Portuguese pear industry. *Food Policy*, 34(1), 94-101.

Muniesa, M., Jofre, J., García-Aljaro, C., & Blanch, A. R. (2006). Occurrence *of Escherichia coli O157: H7* and other enterohemorrhagic *Escherichia coli* in the environment. *Environmental Science & Technology*, 40(23), 7141-7149.

Nganje, W. E., Skilton, P. F., Jensen, H. H., & Onyeaghala, R. (2011). Traceability in food systems: an economic analysis of LGMA and the 2006 spinach outbreak. 85th Annual Conference of the Agricultural Economics Society, Warwick University, 18-20 April 2011.

Okrent, A., & Alston, J. (2012). *The demand for disaggregated food-away-from-home and food-at-home products in the United States*. USDA-ERS Economic Research Report No. 139. Available at SSRN: https://ssrn.com/abstract=2171315 or http://dx.doi.org/10.2139/ssrn.2171315, last accessed April 2022.

Ottoson, J. R., Nyberg, K., Lindqvist, R., & Albihn, A. (2011). Quantitative microbial risk assessment for *Escherichia coli O157* on lettuce, based on survival data from controlled studies in a climate chamber. *Journal of Food Protection*, 74(12), 2000-2007.

Pang, H., Lambertini, E., Buchanan, R. L., Schaffner, D. W., & Pradhan, A. K. (2017). Quantitative microbial risk assessment for *Escherichia coli O157: H7* in fresh-cut lettuce. *Journal of Food Protection*, 80(2), 302-311.

Pouliot, S. (2011). Market evidence of packer willingness to pay for traceability. *American Journal of Agricultural Economics*, *93*(3), 739-755.

Pouliot, S., & Sumner, D. A. (2008). Traceability, liability, and incentives for food safety and quality. *American Journal of Agricultural Economics*, 90(1), 15-27.

Scallan, E., Griffin, P. M., Angulo, F. J., Tauxe, R. V., & Hoekstra, R. M. (2011). Foodborne illness acquired in the United States—unspecified agents. *Emerging Infectious Diseases*, *17*(1), 16.

Scharff, R. L., Besser, J., Sharp, D. J., Jones, T. F., Peter, G. S., & Hedberg, C. W. (2016). An economic evaluation of PulseNet: a network for foodborne disease surveillance. *American Journal of Preventive Medicine*, *50*(5), S66-S73.

Schulz, L. L., & Tonsor, G. T. (2010a). Cow-calf producer perceptions regarding individual animal traceability. *Journal of Agricultural and Applied Economics*, 42: 659–677.

Schulz, L. L., & Tonsor, G. T. (2010b). Cow-calf producer preferences for voluntary traceability systems. *Journal of Agricultural Economics*, 61: 138–162.

Shew, A. M., Snell, H. A., Nayga Jr, R. M., & Lacity, M. C. (2021). Consumer valuation of blockchain traceability for beef in the United States. *Applied Economic Perspectives and Policy*.

Shuval, H., Lampert, Y., & Fattal, B. (1997). Development of a risk assessment approach for evaluating wastewater reuse standards for agriculture. *Water Science and Technology*, 35(11-12), 15-20.

Smith, R., Cahn, M., Daugovish, O., & Koike, S. (2011). *Leaf lettuce production in California*. UCANR Publications.

Solomon, E. B., Potenski, C. J., & Matthews, K. R. (2002). Effect of irrigation method on transmission to and persistence of *Escherichia coli O157: H7* on lettuce. *Journal of Food Protection*, *65*(4), 673-676.

Toledo, C., & Villas-Boas, S. B. (2019). Safe or not? Consumer responses to recalls with traceability. *Applied Economic Perspectives and Policy*, 41(3), 519-541.

U.S. Department of Agriculture. (2022). *Fall seasonal effects connected to E. coli outbreaks in bagged romaine*. Available at: https://www.ars.usda.gov/news-events/news/research-news/2022/fall-seasonal-effects-connected-to-e-coli-outbreaks-in-bagged-romaine/, last accessed April 2022.

U.S. Department of Agriculture, Economic Research Service. (2019). *Cost estimates of foodborne illnesses*. Available at: https://www.ers.usda.gov/data-products/cost-estimates-of-foodborne-illnesses.aspx, last accessed April 2022.

U.S. Department of Agriculture, Economic Research Service. (2018). *Vegetables and pulses yearbook*. Available at: https://www.ers.usda.gov/data-products/vegetables-and-pulses-data/vegetables-and-pulses-yearbook-tables/, last accessed April 2022.

U.S. Department of Agriculture, Economic Research Service. (2013). *Western irrigated agriculture: production value, water use, costs, and technology vary by farm size***.** Available at: https://www.ers.usda.gov/amber-waves/2013/september/western-irrigated-agriculture-production-value-water-use-costs-and-technology-vary-by-farm-size/, last accessed April 2022.

U.S. Department of Agriculture, National Agricultural Statistics Service. *Quick Stats*. Available at: https://quickstats.nass.usda.gov/, last accessed April 2022.

U.S. Food and Drug Administration. (2022). *Outbreaks of foodborne illness*. Available at: https://www.fda.gov/food/recalls-outbreaks-emergencies/outbreaks-foodborne-illness#:~:text=When%20two%20or%20more%20people,from%20happening%20in%20the%20future. last accessed April 2022.

U.S. Food and Drug Administration. (2021). *Efficacy protocol for reduction of foodborne bacteria in preharvest agricultural water*. Available at: https://www.fda.gov/food/food-safety-modernization-act-fsma/efficacy-protocol-reduction-foodborne-bacteria-preharvest-agricultural-water, last accessed April 2022.

U.S. Food and Drug Administration. (2021). *Outbreak investigation of salmonella: Jule's Cashew Brie (April 2021)*. Available at: https://www.fda.gov/food/outbreaks-foodborne-illness/outbreak-investigation-salmonella-jules-cashew-brie-april-2021. last accessed April 2022.

U.S. Food and Drug Administration. (2020). *Requirements for additional traceability records for certain foods.* Available at: https://www.federalregister.gov/documents/2020/09/23/2020-

20100/requirements-for-additional-traceability-records-for-certain-foods, last accessed April 2022.

U.S. Food and Drug Administration. (2016). *Leaf lettuce*. Available at: https://www.wifss.ucdavis.edu/wp-content/uploads/2016/10/LeafLettuce_PDF.pdf, last accessed April 2022.

U.S. Food and Drug Administration. (2015). *Standards for the growing, harvesting, packing, and holding of produce for human consumption—final regulatory impact analysis—Final regulatory flexibility analysis—Unfunded mandates reform act analysis*. Available at: https://www.fda.gov/media/132899/download, last accessed April 2022.

U.S. Food and Drug Administration. (2014). *Supplemental notice of proposed rulemaking: standards for the growing, harvesting, packing and holding of produce for human consumption, economic impact Analysis*. September 2014. Available at: https://www.federalregister.gov/documents/2014/09/29/2014-22447/standards-for-the-growing-harvesting-packing-and-holding-of-produce-for-human-consumption, last accessed April 2022.

U.S. Food and Drug Administration. (2010). *Program information manual retail food protection: recommendations for the temperature control of cut leafy greens during storage and display in retail food establishments*. Available at: https://www.fda.gov/media/78982/download, last accessed April 2022.

USGS-EPA. (2020). *Water quality portal, national water quality monitoring council*, Available at: https://www.waterqualitydata.us/, last accessed April 2022.

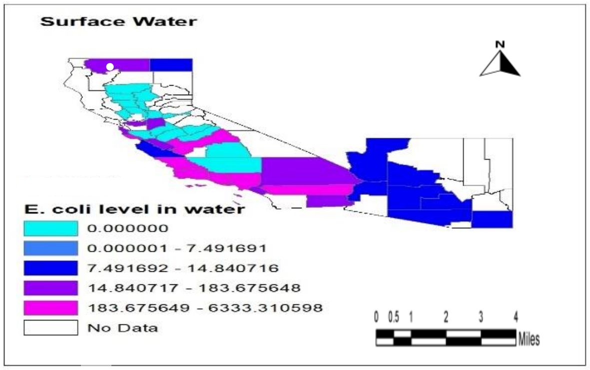
Wang, X., & Li, D. (2012). A dynamic product quality evaluation based pricing model for perishable food supply chains. *Omega*, 40(6), 906-917.

Weller, D. L., Kovac, J., Roof, S., Kent, D. J., Tokman, J. I., Kowalcyk, B., Oryang, B., Ivanek, R., Aceituno, A., Sroka, C., & Wiedmann, M. (2017). Survival of *Escherichia coli* on lettuce under field conditions encountered in the northeastern United States. *Journal of Food Protection*, 80(7), 1214-1221.

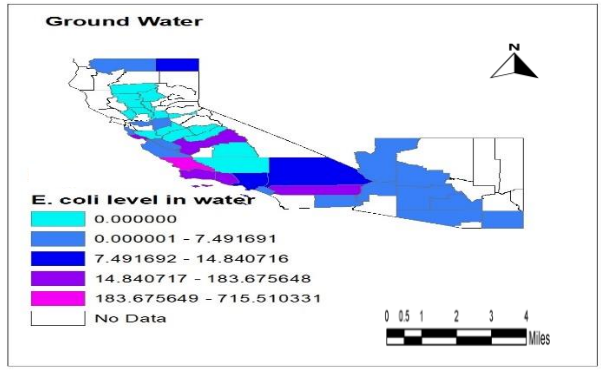
Zhou, J., Jin, Y., & Liang, Q. (2022). Effects of regulatory policy mixes on traceability adoption in wholesale markets: Food safety inspection and information disclosure. *Food Policy*, 107, 102218.

Appendix A. Additional Data and Figures

Figure A.1: Average *E. coli* CFU/100 ml of surface and ground water in Arizona and California



a) Surface water



b) Ground water

Table A.1: Summary of variables used in the model

Symbol	Variable	Unit
SW	Expected value of social welfare	$
$p_{n,t}^{d}(x_{n,t}^{d} * \delta_{n,t})$	Inverse demand function	-
$p_{n,t}^{s}(x_{n,t}^{s})$	Inverse supply function	-
$x_{n,t}^{d}$	Quantity demanded of lettuce	CWT
$x_{n,t}^{s}$	Quantity supplied of lettuce	CWT
$i_{n,t,t+k,w,d}$	Number of illness cases	-
$s_{n,t,t+k,w,d}$	Amount of stored lettuce	CWT
$r_{n,t,t+k,w,d}$	Amount of discarded lettuce	CWT
$\gamma_{n,t,t+k,w,d}$	Indicator variable for foodborne outbreaks	-

Table A.2: Summary of parameters used in the model

Symbol	Parameter	Baseline Values	Unit	Source
A_d	Observed acres of lettuce planted in 2016	Various	Acres	USDA (various years)
$C_{n,t,w,g,d}$	Concertation of *Generic E. coli* in water	-	CFU/100 ml	USGS and EPA (2020)
$CN_{n,t,w,d}$	Concentration of *E. coli* in lettuce	Various	CFU/ml	Estimated in the model
$DO_{n,t,w,d}$	Dose per contaminated serving	Various	CFU/serving	Estimated in the model
$PR_{n,t,w,d}$	Probability of illness per serving	Various	Probability/ serving	Estimated in the model
π	Pathogenicity of water per unit of *Generic E. coli*	$10^{-1.9}$	-	Pang et al. (2017); Ottoson et al. (2011)
ϱ	Serving size of lettuce	85	Gram	FDA (2015)
α	Parameter representing whether there is a traceability	-	-	Authors assumption
ω	Dose-response relation parameter	229.3	-	Pang et al. (2017)
ρ	Dose-response relation parameter	0.267	-	Pang et al. (2017)
$l(g)$	Number of days between irrigation events	6	Days	Smith et al. (2011)
ζ	Die-off function parameter	2.1	-	Brouwer et al. (2017)
ϵ	Die-off function parameter	0.59	-	Brouwer et al. (2017)
β	Economic losses per foodborne illness	8,500	$	USDA (2019)
η	Proportion of *E. coli* that remains in the applied irrigation water	0.7	-	Authors assumption
λ	Irrigation efficiency	0.7	-	USDA (2013)
θ	Volume of water on produce	0.96	%	Ottoson et al. (2011)
y_d	Average observed yield	Various	CWT	USDA (Various years)
N	Number of states of nature	500	-	Authors assumption
NX	Net export of lettuce	2147056.7	CWT	USDA (2018)
Ʊ	Microbial die-off rate during storage	31	%	Pang et al. (2017)
M	Upper bound for number of illnesses	1e+3	Cases	Authors assumption
U	Upper bound for amount of stored lettuce	1e+5	CWT	Authors assumption
א	Demand shift per number of reported and clinically diseased individuals	1.3e-7	-	Bovay and Sumner (2017); Arande et al. (2009); Shuval et al. (1999)

Table A.3: Validation and baseline results

	Validation results	Observed in 2016	% change
Daily Produced Quantity (1,000 CWT)	238,016	238,016	0
Price ($/CWT)	19.62	19.62	0