FOOD SAFETY AND THE DEMAND FOR LEAFY GREENS

ELINA TSELEPIDAKIS
ECONOMIC RESEARCH SERVICE
U.S. DEPARTMENT OF AGRICULTURE
elina.tselepidakis@ers.usda.gov

May 27, 2015

PRELIMINARY PAPER. PLEASE DO NOT CITE.


Any opinions, findings, recommendations, or conclusions are those of the authors and do not necessarily reflect the views of the Economic Research Service or the U.S. Department of Agriculture. The analysis, findings, and conclusions expressed in this paper also should not be attributed to either Nielsen or Information Resources, Inc. (IRI).
FOOD SAFETY AND THE DEMAND FOR LEAFY GREENS

ELINA TSELEPIDAKIS

Unsafe contaminated foods are responsible for millions of illnesses and lead to significant losses in life and productivity. The Centers for Disease Control and Prevention (CDC) estimate that foodborne disease is the cause of approximately 48 million illnesses, 128,000 hospitalizations, and 3,000 deaths annually within the United States (Scallan et al. 2011a and 2011b). Only 20 percent of illnesses (9.4 million illnesses) can be attributed to a specific pathogen. In fiscal terms, these 9.4 million illnesses impose an estimated annual cost of 15.5 billion dollars covering medical expenditures, lost productivity, and quality of life losses (Hoffman et al. 2015). Leafy green safety has been of particular concern for consumers, producers, and regulators, especially following an unprecedentedly large, multi-state E. coli O157:H7 outbreak linked to contaminated bagged spinach. In all, the September 2006 outbreak resulted in 205 known illnesses, 104 hospitalizations, and three deaths. While the 2006 spinach outbreak is a dramatic example of a food safety incident involving leafy greens, bacterial contamination cases of lettuce and spinach continue to occur on a regular basis. Using data from outbreak-associated illnesses from 1998 to 2008, the CDC recently determined that more foodborne illnesses were attributed to leafy vegetables (22 percent) than to any other commodity, including meat, poultry, dairy, and eggs (Painter et al. 2013).

The increasing prevalence and prominence of incidents of foodborne illness linked to leafy greens has the strong potential to undermine consumer confidence in the national supply, especially the supply of packaged and bagged leafy green products. Inevitably, concerns of bacterial contamination and foodborne disease can significantly influence demand for leafy green products. Perceived risk of bacterial contamination is particularly heightened following recalls of leafy green products as consumers often perceive recall events to be an unbiased proxy...
for low quality (Marsh et al. 2004). Using a Bayesian learning model to motivate how recalls impact a consumer’s perception of risk (and in turn consumer demand), the primary objective of this research is to investigate the effect of food safety recalls on the demand for leafy green products.

There is already an extensive and growing body of literature investigating the impact of food safety information on the demand for food products. Many of these studies have measured the impact of food safety information by analyzing media indices (Smith et al. 1988; Brown and Schrader 1990; Burton and Young 1996; Kinnucan et al. 1997; Dahlgran and Fairchild 2002; Piggott and Marsh 2004; Coffey et al. 2011), singular events (Foster and Just 1989; Shimshack et al. 2007; Schlenker and Villa-Boas 2009), and/or aggregate data. Generally, these studies have found statistically significant evidence of own-effects on the demand for the contaminated product involved, and some have even found evidence of cross-effects on the demand for other products. However, analyses of media indices and singular events do not necessarily capture the impact of actualized recurrent food safety events, and analyses of data that have been aggregated across households and across time often ignore household heterogeneity, localized impacts of regional events, and any immediate short-run effects.

Specifically concerned with the spinach contamination event of 2006, several authors have investigated the impact of the outbreak on leafy green demand. Arnade et al. (2009), using aggregate expenditure data for spinach, bagged salads, and other leafy greens, showed that the 2006 spinach recall led to a substantial decrease in purchases of bagged spinach and a marginal decrease in purchases of bulk spinach, with impacts persisting for over a year. Similarly, Arnade et al. (2011), using an error correction model to estimate the rate of adjustment from
disequilibrium after the shock to equilibrium, found that it took consumers 8.5 weeks to return to the equilibrium for bagged spinach demand.

The proposed present study contributes to the leafy green demand literature, and the food safety literature overall, by analyzing disaggregated household-level data to estimate the effect of multiple leafy green recall events that vary over time and space. By using disaggregated household demographic and purchasing data, this study effectively takes advantage of the geographic and temporal variability of leafy green product recalls in order to accurately measure the impact of multiple food safety signals on the demand for leafy green products. That is, temporal variability allows for the analysis of multiple recall events over time, and geographic variability allows for the analysis of a regional recall on the impacted region as compared to the rest of the nation. Additionally, the present study considers the impact of individual household-level characteristics on purchasing behavior and accounts for heterogeneity amongst households. And lastly, assuming recall events serve as a signal for product safety, the study considers varying levels of signal strength; for example, recalls prompted by consumer illness investigations are considered stronger signals than recalls prompted by microbial testing or inspection.

**BACKGROUND: FOOD PRODUCT RECALLS**

Within the United States, the two federal authorities responsible for food safety are the Department of Agriculture’s (USDA) Food Safety and Inspection Service (FSIS) and the Department of Health and Human Services’ Food and Drug Administration (FDA). The FSIS inspects and regulates meat, poultry, catfish, and processed egg products, while the FDA inspects and regulates all other food products, including leafy greens. In order to ensure that the nation’s food products are safe, wholesome, and accurately labeled, both the FSIS and FDA coordinate
and oversee the recalls of products that may cause increased health risks. Examples of possible health risks include pathogen contamination, foreign object contamination, undeclared allergens, and undeclared sulfites.

Health risks are usually discovered one of four ways: the manufacturer or distributor discovers the presence of a health risk through testing or inspection and contacts the FDA or FSIS; a USDA or FDA inspector discovers the presence of a health risk through testing or inspection; a state agency discovers the presence of a health risk through testing or inspection; or a consumer illness prompts an investigation and the source of illness is traced back to a specific product and manufacturer. As soon as the threat is discovered and the manufacturer decides (or is mandated) to recall the contaminated product, the FSIS or FDA determines the severity of the threat posed by the marketed product and assigns the recall one of three classifications: Class I, II, or III. Class I represents a health hazard situation in which there is reasonable probability that consuming the product will cause health problems or death; Class II represents a potential health hazard situation in which there is a remote probability of adverse health consequences from the consumption of the product; and Class III represents a situation in which consuming the product will not cause adverse health consequences. The same classification system is used by both the FDA and FSIS.

Once the recall is assigned a severity classification, the FSIS, the FDA, or the manufacturer issues a press release to vendors and media outlets in the areas where the product was distributed. For recalls overseen by the FDA, the press release includes the date of the FDA recall announcement, a description of the product recalled, the reason for the recall and the health risk involved, the distribution of the contaminated product, and information on how the

---

1 Under the Food Safety Modernization Act of 2011, the FDA, for the first time, has the authority to impose a mandatory recall and shut down operations at food production facilities if it deems that there is a significant threat to public health.
health risk was discovered. In addition, the press release usually includes information as to whether the contaminated product was available for retail purchase, or distributed to restaurants and institutional facilities (schools, prisons, nursing homes, etc.). However, the FDA does not issue a press release for every recall under its authority; they will only seek media publicity when the situation warrants widespread and public awareness, for example, the nationwide distribution of a Class I recalled product.

Following a press release or notice from the manufacturer, vendors of the contaminated product are instructed to remove the product from the market so that it is no longer available for purchase or consumption. Likewise, consumers are instructed to check any products they may have purchased before the recall announcement and determine whether their products match the description of the contaminated product. If the description is a match, consumers are strongly encouraged to discard the product or return the product for a refund. If a consumer has already consumed the product, the consumer is instructed to closely monitor his or her health and seek any necessary medical attention.

While not all FDA recalls are issued press releases, they all are included in FDA's weekly Enforcement Report once they are classified according to the level of hazard involved. Press releases often include more information about the discovery of the health risk and the distribution of the product (retail, foodservice, institutions, etc.), while Enforcement Reports include more information about the quantity of the product recalled and the severity classification. As this study is primarily concerned with household purchases of retail goods in response to recalls, only publicized food recalls will be considered (i.e., recalls with press releases), and Enforcement Reports will be used to verify the information contained in the corresponding press releases.
The present study focuses on the years between 2008 and 2012, and the products chosen for analysis are packaged leafy greens, specifically lettuce and spinach. Between 2008 and 2012, there were over 2,500 food product recalls overseen by the FDA,\(^2\) 41 of which were leafy green (lettuce or spinach) product recalls due to microbial contamination. Details of these recalls are summarized in Table 1. The most common microbial pathogen associated with these recalls was *Salmonella*; however, other pathogens include *E. coli O157:H7*, *E. coli O145*, and *Listeria monocytogenes*.

### Table 1. Number of Leafy Green Recalls, 2008-2012

<table>
<thead>
<tr>
<th></th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leafy Green Recalls</td>
<td>0</td>
<td>4</td>
<td>14</td>
<td>16</td>
<td>17</td>
<td>51</td>
</tr>
<tr>
<td>Publicized Leafy Green Recalls</td>
<td>0</td>
<td>3</td>
<td>8</td>
<td>13</td>
<td>17</td>
<td>41</td>
</tr>
<tr>
<td>Publicized Lettuce Recalls</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>7</td>
<td>14</td>
<td>27</td>
</tr>
<tr>
<td>Publicized Spinach Recalls</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>Publicized Nationwide Leafy Green Recalls</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Source: FDA Press Releases and Enforcement Reports. Computed by author.

As mentioned in the previous section, the greatest advantage of using multiple recall events to measure the impact of food safety information is the temporal and geographic variability. Of the 41 publicized leafy green products recalled between 2008 and 2012, only two were distributed nationwide. The remaining recalled products were distributed to regions identified by the FDA, and the size of the affected regions ranged from a single state to several dozen states. Specifically, the average regional recall impacted 11 states, while the most expansive regional recall impacted 26 states. Additionally, recalls with varying discovery origins may impact consumers differently; that is, recalls prompted by a consumer illness investigation may have a

---

\(^2\) While the FDA also oversees the recalls of drugs, medical devices, pet food, and dietary supplements, the data collected for this project only covers food and beverages.
stronger impact on purchasing behavior than a recall prompted by sample testing. Of the 41 publicized leafy green products recalled, two were prompted by a consumer illness investigation.

**ECONOMIC FRAMEWORK**

*Risk Perception*

Fundamental to the present analysis is the assumption that consumers derive value from food safety because it signals a lower degree of health risk. Yet consumers often face imperfect information; that is, they are mostly uncertain regarding the safety of available food products. If a consumer had perfect information, the safety of the food product would be no different than other quality attributes, such as taste, appearance, source, etc. And the consumer would make a purchase decision based on his or her preferences, income, and price of the product. Without perfect information, however, consumers must assess a food product’s safety based on their personal experience and communications. Often, when purchases are supported by satisfactory experiences, a consumer will form routine shopping behaviors (Hoyer 1984). Established routine purchases will continue until the consumer receives a signal strong enough to revise prior risk perceptions and decision rules.

The learning process by which consumers process information and update their risk perceptions can be expressed with a Bayesian revision process (see Viscusi and O’Conner 1984; Viscusi 1989; Liu et al. 1998; Böcker and Hanf, 2000). Following Viscusi (1989) and Liu et al. (1998), let \( r_t \) denote perceived risk, i.e., the perceived probability that a given good is unsafe at time \( t \), and let \( s_t \) denote a negative information signal on the safety of a good, e.g., a recall. The updated perceived risk (posterior risk) can then be expressed as the weighted average of the prior perceived risk and the sample risk:

\[
r_{t+1} = \omega_{t+1} r_t + (1 - \omega_{t+1}) f(s_{t+1}),
\] (1)
where $f(\cdot)$ is a function that converts the signal into a sample risk and $\omega_t$ is a weight for combining the prior perceived risk and the sample risk. For simplicity, assume no news is good news so that $s_t \geq 0$, $f(s_t = 0) = 0$, and $f(s_t > 0) > 0$. That is, when no signal is observed, the consumer perceives the food product to be safe and the sample risk is zero.\(^3\) Under this assumption, were the consumer never to receive a signal, the posterior risk would eventually converge to zero. However, given the ubiquity of information signals such as food product recalls, convergence to zero is an unlikely scenario. This Bayesian framework illustrates frequently observed behavior following the release of negative information: an immediate change in behavior, followed by a gradual return to previous, routine behavior. Through a dynamic adjustment process, perceived risk may eventually return to initial levels.

How long the return to baseline behavior takes depends on the strength of the negative signal. Stronger signals will inevitably require a longer recovery period than weaker signals, for example, recalls prompted by a consumer illness investigation will likely require more recovery time than recalls prompted by product testing. To demonstrate this, assume the weight $\omega$ is constant throughout time. Suppose the baseline perceived risk is $r_0$ and there is a negative shock in period 1. Assuming no more shocks, the perceived risk at time $T$ is

$$r_T = \omega^T(r_0 + \omega^{-1}(1 - \omega)f(s_1)).$$  \(2\)

To determine an expression for the length of time to recovery, set $r_0 = r_T$ and solve for $T$.

$$T = \frac{\ln(r_0) - \ln(r_0 + \omega^{-1}(1 - \omega)f(s_1))}{\ln(\omega)}.$$  \(3\)

Taking the derivative with respect to the signal, $s_1$, the expression becomes

---

\(^3\) The assumption that consumers may perceive no news as good news is not unfounded. A study of consumer attitudes towards food safety showed that individuals often display ‘optimistic bias’ and hold an illusion of personal invulnerability with regard to food safety hazards (Redmond and Griffith 2004).
\[ \frac{\partial T}{\partial s_1} = -\frac{(1-\omega)f'(s_1)}{\ln(\omega)(\omega r_0 + (1-\omega)f(s_1))}. \]  

(4)

Assuming \( f'(\cdot) > 0 \) (stronger signals translate to greater risk) and because \( 0 < \omega < 1 \), then
\[ \frac{\partial T}{\partial s_1} > 0. \]
Therefore, stronger negative signals require greater recovery times.

Now consider the possibility of multiple negative signals. Suppose a shock occurs every period between 1 and \( \tau \), followed by periods where no shocks are observed between \( \tau + 1 \) and \( T \) (\( 1 < \tau < T \)). The posterior risk perception in time \( T \) would then be
\[ r_T = \omega^T (r_0 + \sum_{t=1}^{\tau} \omega^{-t} (1 - \omega) f(s_t)) \]  

(5)

Again, we set \( r_0 = r_T \) and solve for \( T \).
\[ T = \frac{\ln(r_0) - \ln(r_0 + \sum_{t=1}^{\tau} \omega^{-t} (1-\omega)f(s_t))}{\ln(\omega)} \]  

(6)

Now define \( \tau_2 \) as period of time greater than \( \tau_1 \), and again shocks occur every period between 1 and \( \tau_1 \) or \( \tau_2 \). Because \( \tau_2 > \tau_1 \), then the length of time to recovery following \( \tau_2 \) shocks, \( T_2 \), will be greater the length of time to recovery following \( \tau_1 \) shocks, \( T_1 \); that is, \( T_2 - T_1 > 0 \). To prove this, assume otherwise: \( T_2 - T_1 \leq 0 \).

\[-\ln \left( r_0 + \sum_{t=1}^{\tau_2} \omega^{-t} (1 - \omega) f(s_t) \right) + \ln \left( r_0 + \sum_{t=1}^{\tau_1} \omega^{-t} (1 - \omega) f(s_t) \right) \geq 0 \]
\[ \ln \left( r_0 + \sum_{t=1}^{\tau_2} \omega^{-t} (1 - \omega) f(s_t) \right) - \ln \left( r_0 + \sum_{t=1}^{\tau_1} \omega^{-t} (1 - \omega) f(s_t) \right) \leq 0 \]
\[ \sum_{t=1}^{\tau_2} \omega^{-t} (1 - \omega) f(s_t) - \sum_{t=1}^{\tau_1} \omega^{-t} (1 - \omega) f(s_t) \leq 0 \]
\[ \sum_{t=\tau_1+1}^{\tau_2} \omega^{-t} (1 - \omega) f(s_t) \leq 0 \]
This leads to a contradiction because each term in the summation expression is positive:

\[ 0 < \omega < 1 \text{ and } f(s_t > 0) > 0. \]

Therefore, \( T_2 - T_1 > 0 \), which proves that multiple signals only lengthen the time necessary for recovery. Alternatively, this can be demonstrated by redefining the length of period \( t \) to be inclusive of multiple shocks. Multiple shocks in a single period can be interpreted as stronger signals than a single or fewer shocks in a period, and because we’ve already shown that stronger signals lead to greater recovery times, we can also conclude that, similarly, multiple shocks also lead to greater recovery times.

\textit{Demand}

Now consider a consumer who derives utility directly from the consumption of good, \( y \), and the quality or safety of that good, \( q \).\textsuperscript{4} Assume that the quality of the potentially risky good, \( q \), has a binary distribution. That is, the product is either contaminated with a harmful pathogen, \( q^C \), or not, \( q^{NC} \). However, as previously stated, although quality enters a consumer’s utility function, the exact quality or safety of a particular good is not known to the consumer prior to purchase. The consumer only has formed a perception of risk (the probability that a good is unsafe), previously defined as \( r \). At this stage, several additional plausible assumptions are necessary regarding utility. Namely, the quantity of goods and the consumer’s utility are positively correlated: \( U_y > 0 \); quality and utility are positively correlated: \( U_q > 0 \); and lastly, the consumer’s utility function is concave with respect to \( y \). Ultimately, the consumer’s utility can be expressed as

\[ U(y, q) = (1 - r)U(y, q^{NC}) + rU(y, q^C) \quad (7) \]

\textsuperscript{4} Several authors have modeled demand for food safety by including a health function in the theoretical framework, where health in turn is a function of quality and other factors (see van Ravenswaay and Howehn, 1996; Antle, 2001). Alternatively, others have modeled demand for food safety by incorporating quality directly into the utility framework (see Piggott and Marsh, 2004; Coffey et al., 2011). For simplicity, the latter method is applied here.
To determine the comparative statistic $\frac{dy}{dr}$, the change in demand in response to a change in perceived risk, we use the implicit function theorem upon calculating the first-order and second-order conditions.

$$\frac{dy}{dr} = \frac{U_y(y,q^{NC}) - U_y(y,q^C)}{(1-r)U_{yy}(y,q^{NC}) + rU_{yy}(y,q^C)} < 0$$  \(8\)

Assuming that the marginal utility from a non-contaminated good is greater than the marginal utility of a contaminated good (numerator) and knowing that $U_{yy}(\cdot) < 0$ because utility is concave (denominator), then as perceived risk for a good increases, demand for the good decreases, $\frac{dy}{dr} < 0$. Note that the comparative statistic $\frac{dy}{dr}$ is derived here from a simple one-good utility maximization problem without an income constraint. This derivation was chosen because leafy green expenditures presumably represent a very small fraction of a household’s income and therefore any income effect would be very small or non-existent. However, it is also possible to demonstrate a similar result from a two-good (one no-risk good and one risky good) utility maximization problem with an income constraint.

Next, linking food safety signals, $s$, to demand is straightforward. Applying the chain rule, the relationship becomes

$$\frac{dy_t}{ds_t} = \frac{dy_t}{dr_t} \frac{dr_t}{ds_t}$$  \(9\)

and because $\frac{dy_t}{dr_t} < 0$ and $\frac{dr_t}{ds_t} > 0$, then $\frac{dy_t}{ds_t} < 0$. That is, as a consumer receives stronger negative signals thereby increasing perceived risk, the likelihood of purchasing potentially contaminated food products declines. Additionally, we can further deduce that

$$\frac{dy_t}{ds_t} < \frac{dy_{t+1}}{ds_t} < 0.$$  \(10\)
To prove this, assume otherwise: \( \frac{dy_t}{ds_t} \geq \frac{dy_{t+1}}{ds_t} \). 

\[
\frac{dy_t}{dr_{lt}} \frac{\partial r_{lt}}{\partial s_{lt}} \geq \frac{dy_{t+1}}{dr_{lt+1}} \frac{\partial r_{lt+1}}{\partial r_{lt}} \frac{\partial r_{lt}}{\partial s_{lt}} 
\]

\[
\frac{dy_{lt}}{dr_{lt}} \geq \frac{dy_{lt+1}}{dr_{lt+1}} \omega 
\]

Assuming that the change in demand in response to a change in risk perception does not vary from period to period, that is, \( \frac{dy_{lt}}{dr_{lt}} = \frac{dy_{lt+1}}{dr_{lt+1}} \), then

\( \omega \leq 0 \).

This leads to a clear contradiction because \( \omega \) is already defined as \( 0 < \omega < 1 \). Therefore, we can conclude that the impact of a negative information signal on demand diminishes over time.

Similarly, given the direct link between information signals and demand, we can posit that stronger signals and multiple signals will lead to longer recovery times.

Empirically estimating simple demand functions for leafy greens will be an informative exercise in order to determine whether these relationships hold in reality. Of particular interest is the impact of negative signals on the quantity demanded of leafy greens and the estimated time to recovery.

**DATA**

The primary dataset used in this analysis is the Information Resources, Inc. (IRI) Consumer Network™ - a nationwide panel of households that provide a detailed account of their retail food purchases. The panel is selected to be geographically and demographically representative of the United States population. Households participating in the panel were provided with a handheld scanner to scan the Universal Product Code (UPC) on all their purchases and upload all
information through the Internet or a landline telephone. The data of household leafy green purchases include a detailed product description, product brand, leafy green type, date of purchase, total quantity, and total expenditure for every item purchased. Households also provide demographic data on an annual basis including county of residence, household composition, household size, income, education, age, and race.

As previously stated, the years of interested are 2008 through 2012 and a monthly periodicity was selected. The products chosen for analysis are packaged iceberg and iceberg-based products (iceberg with shredded cabbage and/or carrot), romaine and romaine-based products (romaine with shredded cabbage and/or carrot), spinach products, other leafy green products (other lettuces, arugula, kale, chard, cabbage, etc.), and mixed green products (products containing more than one type of leafy green). Leafy green products containing dressing, toppings (croutons, nuts, berries, etc.), and other vegetables were not considered for this analysis.

Preliminary summary statistics of household purchases of leafy greens are presented in table 2.

<table>
<thead>
<tr>
<th>Table 2. Household Leafy Green Purchases, 2008-2012</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>2008          2009        2010        2011        2012</td>
</tr>
<tr>
<td>No. of Leafy Green Purchases   411,746  503,088  514,846  532,846  523,666</td>
</tr>
<tr>
<td>No. of Households Purchasing Leafy Greens  42,528   50,070   51,076   51,767   50,452</td>
</tr>
<tr>
<td>Average No. of Leafy Green Purchases per Household  9.68    10.05    10.08    10.29    10.38</td>
</tr>
<tr>
<td>Total No. of Static5 Households Participating in IRI Panel  53,610  62,674  63,593  64,330  62,503</td>
</tr>
</tbody>
</table>


5 The static panel only includes households that reported purchases at least once every four weeks for 80 percent of the year (11 of 13 four-week periods) and reported average weekly expenditures of 25 dollars for one member households, 35 dollars for two member households, and 45 dollars for three or more member households.
**Empirical Methodology**

The popular Almost Ideal (AI) demand system, first proposed by Deaton and Muellbauer (1980), is employed for this analysis. Consistent with microeconomic theory, the AI system is derived from a generalized consumer expenditure function with given prices and a predetermined level of utility. Moreover, the linear AI system is favored for empirical estimation because it satisfies the axioms of choice, allows for a flexible functional form of the indirect utility and expenditure function, and is generally simple to estimate.

A central issue in disaggregated demand analysis is the high proportion of zero expenditures for individual commodities in any given time period. The censored nature of the data inevitably makes it difficult to estimate large, theoretically consistent, disaggregated consumer demand models. To address this issue, there is a large empirical literature offering feasible estimation techniques (see Wales and Woodland, 1983; Lee and Pitt, 1986; Heien and Wessels, 1990; Shonkwiler and Yen, 1999; Perali and Chavas, 2000; Golan et al., 2001; Yen et al., 2003; Dong et al., 2004; Meyerhoefer et al., 2005; Yen and Lin, 2006).

This study adopts the two-step estimation approach of Meyerhoefer, Ranney, and Sahn (2005) so as to address possible corner solutions and control for heterogeneous preferences amongst household. This approach is similar to the two-stage estimator first proposed by Perali and Chavas (2000) in that the reduced form parameters are estimated in the first stage followed by the imposition of demand theory restrictions and identification of structural parameters in the second stage. Meyerhoefer et al. extended this specification to make use of panel data and control for household heterogeneity.
In the first stage, the econometric model for the household share \( w_{ith} \) of the \( i \)-th leafy green at time \( t \) is
\[
  w_{ith} = \bar{\alpha}_{ith} + \sum_{j=1}^{N} \gamma_{ij} \log(p_{jht}) + \beta_i \log\left(\frac{x_{ht}}{P_{ht}}\right) + \bar{\epsilon}_{ith}
\]
where \( p_{jht} \) denotes the price of good \( j \) at time \( t \) faced by household \( h \), \( x_{ht} \) denotes the total household expenditure on leafy green products in time period \( t \), \( P_{ht} \) denotes a leafy greens price index for household \( h \) in time period \( t \), and \( \bar{\epsilon}_{ith} \) is the error term that is heteroskedastic within the share equation for one good and correlated across the share equations for different goods:
\[
  \bar{\epsilon}_{ith} = \epsilon_{ith} - \beta_i \sum_j \log(p_{jht}\epsilon_{jht}).
\]
The price index, \( P_{ht} \), is the geometrically weighted average of prices, calculated as \( \log(P_{ht}) = \sum_i w_{ith}^0 \log(p_{ith}) \) where \( w_{ith}^0 = \frac{1}{T} \sum_{t=1}^{T} w_{ith} \). Demographic factors of interest, such as household size, and other demand shifters, including recall events and seasonality, are incorporated in the intercept term, \( \bar{\alpha}_{ith} \), where
\[
  \bar{\alpha}_{ith} = \alpha_i + \sum_k \rho_{ik} d_{kht}
\]
given \( K \) different demand shifters, \( d_{kht} \).

Following a heteroskedastic Tobit model estimation of equation 11, the first-stage reduced form parameter estimates can be denoted as row vector \( \pi^t = (\pi^t_{i1}, ..., \pi^t_{iT})' \) for demand equation \( i \) in time period \( t \) of length \((K + N + 2)\). As previously stated, the second stage of estimation consists of imposition of demand theory restriction and identification of structural parameters. As outlined by Meyerhoefer et al., define \( \psi \) as a \( Q \)-dimentional vector of structural parameters. To solve for \( \psi \), the following minimum distance estimator is solved using a generalized methods of moments (GMM) estimation procedure
\[
  \min_{\psi} \left[ \hat{\pi} - h(\psi) \right]' \hat{\Omega}^{-1} \left[ \hat{\pi} - h(\psi) \right]
\]
\[\]
\[\]
\[6\] This empirical analysis is based on the assumption that leafy greens are weakly separable from all other consumer goods.
where $h(\cdot)$ is a nonlinear function mapping $\psi$ onto $\pi$ that is used to impose demand theory restrictions on the reduced form parameters.

RESULTS & DISCUSSION

Preliminary analysis is ongoing.

REFERENCES


