

The Role of Income in Explaining the Shift from Preserved to Fresh Vegetable Purchases

Peyton M. Ferrier and Chen Zhen

Between 1991 and 2013, the share of fresh vegetables consumed rose from 47% to 56%. While median incomes grew 7.9%, the growth was uneven, with negligible growth occurring in the bottom two quintiles. Estimating an EASI demand system for fresh and preserved vegetables that accounts for corner solutions, we find that income elasticities are larger for fresh vegetables than for preserved vegetables for consumers in the top three but not the bottom two income quintiles. Simulations accounting for uneven income growth indicate that income growth accounts for 0.7 percentage points of the 9 percentage point increase in fresh share.

Key words: corner solutions, EASI demand, Engel curve, vegetable demand

Introduction

After falling during the 1970s and 1980s, the share of vegetables purchased fresh rather than preserved in the United States rose from 47% to 56% between 1991 and 2013. Many factors have been offered to explain this shift, including changes in availability and variety (Pollack, 2001; Just, Lund, and Price, 2012), greater availability of out-of-season imports (Nzaku, Houston, and Fonseh, 2010; Ferrier and Zhen, 2014), product promotion (Rickard et al., 2011), changes in regulatory restrictions to facilitate fresh produce imports (Peterson et al., 2013), innovations in shipping and storage (Pollack, 2001; Regmi and Gehlhar, 2001), and changing preferences associated with larger immigrant communities and greater desire for convenience items (Lucier et al., 2006). Income growth—the focus of this paper—may also have increased the fresh share of consumption if income elasticities of fresh vegetables are larger than those of preserved.

Understanding the increase in fresh consumption informs policy concerns involving nutrition and obesity, household response to time constraints, and food waste. Public health interest groups sometimes promote vegetable consumption by emphasizing fresh-state consumption. For instance, Pollan (2008) and the Union of Concerned Scientists (2013) advocate for increasing the number of farmers' markets, providing more fresh vegetables with school lunch programs, and procuring fresh produce locally as ways to raise overall vegetable consumption. While the health benefit of fresh versus preserved vegetables may be debated by nutritionists, fresh vegetables are often thought to simply taste better than preserved vegetables, making their consumption more pleasant, which might alter preferences through habit formation (Ferrier and Zhen, 2014). As poor nutrition and obesity are associated with massive medical costs (Finkelstein et al., 2009; Cawley and Meyerhoefer, 2012), public health authorities routinely advise increased consumption of fruits and vegetables as a way of

Peyton M. Ferrier is an economist at the USDA Economic Research Service. Chen Zhen is an associate professor in the Department of Agricultural and Applied Economics at the University of Georgia. The authors wish to thank Hayden Stewart and Jeffery Hyman for providing an unpublished version of their commodity concordance, which we used to develop our commodity aggregates. All errors and omissions are our own. This research was made possible through Cooperative Agreement #59-3000-2-0077 from the USDA's Economic Research Service. The thoughts and views presented here are the views of the authors and do not reflect the opinions or views of the Economic Research Service or the USDA as a whole.

Review coordinated by Hikaru Hanawa Peterson.

improving health outcomes.¹ Higher incidence of nutrition-related, non-communicable diseases—such as diabetes and coronary heart disease—is associated with lower-income households (Darmon and Drewnowski, 2008), which may lack resources for non-dietary offsetting health activities (Strazdins et al., 2011). As income growth has largely stagnated for the bottom two income quintiles since 1991, researchers and policy-makers wonder whether a policy of income provision to the poor might be effective in abating nutrition problem among low-income households (Deaton, 2002).

Income growth is likely to foster a shift toward greater fresh-state vegetable consumption to the extent that consumers view fresh products as more convenient or of better quality than preserved vegetables.² Depending on the good and its preparation, preservation can cause subtle, undesirable effects on sensory and nutrition quality characteristics.^{3,4} Fresh products can be more convenient when they can be consumed raw or in more versatile ways than preserved products. Lucier et al. (2006) note that until the 1980s, convenience in produce consumption often referred to the storability aspects of preserved items (canned, frozen, and dried products) or the convenience of readily consumable fruits (bananas, apples) but now often refers to presale processing of fresh vegetables, including bagged salads and spinach and pre-cut carrots, melons, celery, and broccoli florets. Increased fresh product consumption may also have increased rates of post-harvest loss (popularly termed “food waste,” Stuart, 2009).⁵

Despite this concern, previous research (e.g., Bertail and Caillavet, 2008; Dong and Lin, 2009) has not considered how preference differences across income groups or uneven income growth affect fresh versus preserved vegetable demand at the commodity level. For example, using Consumer Expenditure Survey data, Niu and Wohlgenant (2012) find that expenditure elasticities are higher for fresh fruits and vegetables compared to their preserved counterparts as a composite aggregate and show differences in income elasticities between high- and low-income consumers.

In this study, we discuss the striking unevenness of income growth across consumers between 1991 and 2013 and show that estimated market effects under the assumption that income growth is uniform across consumers (at the median or mean rate) are substantially different from those obtained based on actual income growth rates, which vary across income quintile. To quantify income’s role in changing purchasing behavior, we use the Nielsen Homescan panel of household food purchases for at-home consumption to estimate a demand system for ten vegetables (artichokes, asparagus, broccoli, corn, green beans, mushrooms, peppers, onions, tomatoes, and yams) purchased in fresh and preserved states. Our use of the Exact Affine Stone Index (EASI) demand system (Lewbel and Pendakur, 2009) allows us to incorporate a substantially greater degree of flexibility into the shape of Engel curves than previously available demand systems, such as the Translog (TL,

¹ Specific organizations include the U.S. Departments of Health and Human Services and Agriculture (2015), the American Medical Association (Stack, 2016), the World Health Organization (2014), and the Centers for Disease Control and Prevention (Moore and Thompson, 2015).

² Three preservation methods exist to prevent spoilage: dehydration, canning (including pickling), and freezing. Dehydration—as with spices, tomatoes, onions, and beans—removes moisture from food. Canning uses either natural alcohol, the acid byproducts of fermentation (cabbage and melons), salt (pickled cucumbers), or heat (corn or beans) to inhibit microbial growth. Freezing requires vegetables to freeze quickly at very low temperatures in industrial freezers before they may be stored in home freezers, a process that halts cellular activity.

³ Loss of taste, color, and nutrients (vitamins A, C, and thiamine) occurs with both canning (Awuah, Ramaswamy, and Economides, 2007) and dehydration (Sagar and Kumar, 2010). Flash-frozen vegetables (carrots, peas, and broccoli) show reduced levels of vitamin C and loss of firmness (Creed, 2005).

⁴ The development of flash freezing, mechanical refrigeration, and home electrification completed the modern infrastructure for vegetable preservation in the first half of the twentieth century, fostering the frozen food giants Birdseye and Green Giant after World War II. Specifically, Clarence Birdseye’s flash-freezing process was developed in the 1930s and eliminated much of the quality damage sustained using previous freezing methods. Its widespread adoption, however, was slowed by the need for both stores and consumers to have electricity and mechanical refrigeration (Kurlansky, 2009).

⁵ Buzby, Wells, and Hyman (2014) estimate that postharvest losses of fresh vegetables are 10 percentage points higher than for processed vegetables (34% vs. 24% in total, with losses 4 points higher at the retail level and 6 points higher at the consumer level for fresh vegetables).

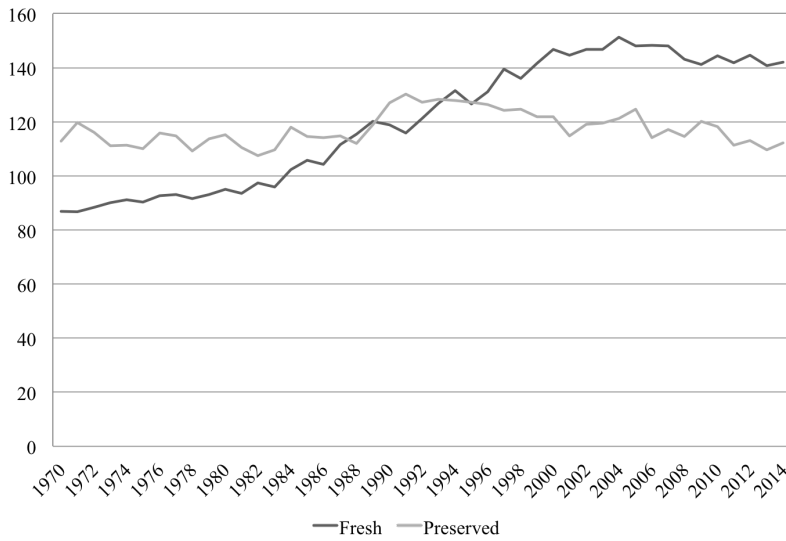


Figure 1. Per Capita Consumption of Fresh and Processed Vegetables

Source: U.S. Department of Agriculture, Economic Research Service (2016).

Christensen, Jorgenson, and Lau, 1973) and Almost Ideal (AI, Deaton and Muellbauer, 1980).⁶ We then use the estimated income elasticities for fresh and preserved vegetables that vary by income quintile to calculate the change in the fresh share of vegetable consumption based on historical income growth by quintile.

Per capita consumption of vegetables has increased generally since the 1970s (figure 1), a shift primarily emerging from increased fresh consumption. While many causal factors are likely at play, this study focuses on the role of income growth and its unevenness. We find that, while vegetable demand is much more income elastic in the bottom two income quintiles, income growth plays little role in determining consumption changes because it has been negligible for these income segments of consumers. For the other three income quintiles, however, its effect is more substantial. In general, we find that neglecting the heterogeneity in income growth rates across income quintiles by only considering overall median income growth to predict consumption results in substantive bias in the predicted effect on consumption growth. When only overall median income growth is considered, income growth explains only 0.1 percentage point of the 9.0 percentage point overall increase in the fresh consumption share. In contrast, when incorporating heterogeneity in income growth in our estimation, it explains 0.7 percentage points of the change.

Demand Modeling Approach

Three main objectives guide our demand estimation approach are (i) to estimate disaggregated price and income effects by including both fresh and preserved vegetables at the commodity level, (ii) to allow for flexibility in the estimation of income effects, and (iii) to impose as much consistency with economic theory as feasible for purchase data with corner solutions. By including both types of vegetable preparations, separate income elasticities for fresh and preserved vegetables can be used to consider the sources of fresh-share growth. Most households, however, only purchase a subset of the available set of vegetables, creating corner solutions where consumers purchase zero units of the good when the price rises above a given level. In this framework, a price decrease can increase consumption by either increasing purchase quantity among consuming households or

⁶ The EASI demand system is also more flexible than the Quadratic AI Demand System as it can incorporate any number of higher-order income effects.

inducing previously non-consuming households to make a purchase. While discrete choice models can address corner solutions directly (Wooldridge, 2002), they are not suited to evaluating the positive region of a continuous demand.

Econometrically, however, zero purchases present two challenges when applied in conjunction with continuous flexible demand systems. First, they necessitate the use of a demand system that accounts for the zero lower bound on purchases in the consumer optimization problem. Second, they require imputation of unobserved prices faced by consumers not making a purchase. Following Perali and Chavas (2000) and Meyerhoefer, Ranney, and Sahn (2005), we use the Tobit model to account for zero values in purchases. Prices for goods not purchased are unobserved, so we impute them based on observed prices—a standard practice in similar demand studies (Kasteridis and Yen, 2012; Zhen et al., 2014).⁷

To allow for sufficient flexibility in the estimation of income effects, we use the EASI demand system developed by Lewbel and Pendakur (2009). The basic AI and TL demand systems limit Engel curves to be linear in log income. While extensions of these models—such as the generalized forms of the AI and TL demand systems (Piggott, 2003) and the Quadratic AI demand system (Banks, Blundell, and Lewbel, 1997)—introduce additional flexibility to Engel curves, the EASI model permits them to take on any arbitrary shape through the inclusion of higher-order income terms. In other words, the EASI model is restricted neither by the rank-three limit for nominal income exactly aggregable demand systems (Gorman, 1981) nor by the rank-four limit for deflated income exactly aggregable demand systems (Lewbel, 1989). Like the AI and TL demand systems, the EASI demand system allows for homogeneity, symmetry, and adding-up conditions to be imposed through parametric restrictions and has a linear approximate form suitable for estimation.

Other approaches to identifying the role of income in fruit and vegetable consumption have been applied as well. Dividing consumers into high- and low-income groups for separate estimations of an AI model while accounting for corner solutions, Niu and Wohlgenant (2012) found that high-income consumers have larger expenditure elasticities for fresh vegetables than low-income consumers but that this relationship is reversed for preserved fruits and vegetables. Moreover, fresh vegetable expenditure elasticity is greater than preserved vegetable elasticity for both high- and low-income consumers. Gustavsen and Rickertsen (2006) estimated a quantile regression for vegetable demand in a single-equation model and found that income effects vary significantly across quantiles and that income is closely correlated with consumption. Bertail and Caillavet (2008) used a finite-mixture AI demand system for aggregated groups of fruits and vegetables to evaluate the responsiveness of consumer demand to price and income variables as they differ across groups identified by socioeconomic variables, including income. They found that consumers with both the lowest income and the lowest fruit and vegetable consumption showed the smallest demand response to price and income changes. Note that their result confounds the longstanding empirical finding (i.e., Engel's Law, Zimmerman, 1932) that income elasticities decrease as household income grows. In contrast to Bertail and Caillavet's results, we find that income elasticities are largest among the lowest income quintiles.

The EASI Demand Model

Suppressing the time and household subscripts to simplify notation, the unconditional linear approximate EASI demand model, which includes zero purchases as corner solutions, is specified

⁷ Using Tobit models to address zero-value purchases is less structural, in terms of recovering parameters of consumer preferences, than the virtual price approach (Lee and Pitt, 1986). However, the latter approach is computationally infeasible in a large demand system and appears to be limited to the TL functional form, which is less flexible than EASI with respect to income effects.

as a system of Tobit models:

$$(1) \quad w_j^* = \sum_{r=1}^R b_{rj}y^r + \sum_{l=1}^L c_{lj}z_l + \sum_{k=1}^J a_{kj}p_k + \varepsilon_j,$$

where w_j^* is the latent budget share for good j and is related to observed budget share w_j according to $w_j^* = \max\{0, w_j\}$; x is total income; w_j is equal to expenditure on good j divided by income (x); y is real income calculated as $y = \ln x - \sum_k w_k \ln p_k$; R is the highest degree of polynomial on real income; z_l is the l th of L demand shifters; p_k is the price index for good k ; J is the number of goods including the numéraire; ε_j is the residual; and b , c , and a are coefficients.

The demand equation (1) is unconditional because it is not conditioned on total food expenditures, which are endogenous with food prices. The coefficients b_{rj} and c_{lj} capture the income and demographic effects, respectively. By increasing the number of higher-order polynomial terms on real income, the Engel curves can be made more flexible and take any curvature shape as determined by data. The a_{kj} coefficient captures the price effect. Heteroskedasticity may be present as consumers with lower incomes may exhibit greater error terms when estimating budget shares.⁹ To control for this possibility, the residual variance is specified as a function of income and its polynomials:

$$(2) \quad \sigma_{jht} = \sigma_j \left(1 + \sum_{s=1}^S \gamma_{sj} (\ln x_{ht})^s \right),$$

where s indexes the degree of income polynomial in the heteroskedastic link function, h indexes household, t indexes time, and s and γ_{sj} are parameters.

If prices are endogenous, the EASI model can be estimated using an instrumental variables method (e.g., Zhen et al., 2014). The diffuse production, lags, and yield variability associated with vegetable production suggest that aggregate supply is exogenous at the commodity level, but individual households can typically purchase unlimited amounts at a fixed and exogenous market price. However, at the household level, prices paid may vary in response to unobserved (to the econometrician) demand factors that cause simultaneity bias in the price coefficients. For instance, households that particularly like a commodity may actively search for lower prices through coupons or other discounts. Heckman shows that household-specific price differences can be separated from the effects of general price movements by the inclusion of a demand shifter variable representing the average price faced by each household for a commodity (Heckman and Vytlacil, 1998; Meyerhoefer, Ranney, and Sahn, 2005). We take this approach, also known as correlated random effects, to correct for potential simultaneity bias in the price coefficients.

In the EASI demand system (without corner solutions), the Hicksian elasticity of demand (h_{ij}) is

$$(3) \quad h_{ij} = \frac{a_{ij}}{w_j} + w_j - \delta_{ij},$$

where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. The $J \times 1$ vector of income elasticities is

$$(4) \quad \mathbf{E} = [\text{diag}(\mathbf{W})]^{-1} [(\mathbf{I}_J + \mathbf{B}\mathbf{P}')\mathbf{B}] + \mathbf{1}_J,$$

where \mathbf{W} is the $J \times 1$ vector of observed budget shares, \mathbf{B} is a $J \times 1$ vector whose j th element equals $\sum_{r=1}^R r b_{rj} y^{r-1}$, \mathbf{P} is the $J \times 1$ vector of log prices, and $\mathbf{1}_J$ is a $J \times 1$ vector of ones. The Marshallian

⁸ Lewbel and Pendakur (2009) found this term to be a good approximation of the real income term in a fully nonlinear EASI demand system.

⁹ Specifically, since our budget shares are unconditional, low-income consumers have larger budget shares for any given level of purchase for a good not included in the numéraire. Idiosyncratic fluctuation in purchases of the same dollar amount (e.g., unexpectedly observing \$10 less frozen corn purchased in a given month because weather delayed a shopping trip) will make the shares vary more for low-income households compared with high-income households.

Table 1. Household Demographic Statistics, Fresh Foods Panel

Variables	Stat	Income Quintile				
		1st	2nd	3rd	4th	5th
Monthly household income	Mean	1,402	2,945	4,725	7,526	10,668
	Median	1,465	3,120	4,590	6,959	9,246
	Minimum	209	2,280	3,942	6,329	9,148
	Std. Dev.	479	458	561	1,027	2,335
Household size	Mean	1.89	2.26	2.59	2.81	3.07
	Std. Dev.	1.25	1.31	1.34	1.30	1.27
Census regions ^a						
Midwest	Mean	0.15	0.15	0.15	0.13	0.10
South	Mean	0.43	0.42	0.40	0.38	0.32
West	Mean	0.19	0.20	0.22	0.25	0.34
Household head race ^a						
Black	Mean	0.154	0.147	0.141	0.120	0.090
Asian	Mean	0.011	0.026	0.032	0.051	0.143
Other race	Mean	0.074	0.058	0.062	0.069	0.045
Hispanic household	Mean	0.080	0.069	0.084	0.092	0.075
Female head, aged 40 and under	Mean	0.181	0.183	0.179	0.168	0.088
Female head, college educated	Mean	0.167	0.249	0.360	0.488	0.812
Observation Counts						
Unique Household Observations	N	2,599	4,251	4,508	4,062	801
Total Observations	N	61,064	94,500	98,382	92,379	11,833

Notes: There are 12,039 unique households and 358,158 household-month observations in our sample. The region, race, Hispanic, and female head variables are dummy variables. Hence, their averages are proportion measures. Statistics are not weighted by Homescan projection factor. ^aNortheast is the reference region. ^bWhite is the reference race.

price elasticity (e_{ij}) is recovered from the Slutsky equation as

$$(5) \quad e_{ij} = h_{ij} - w_j e_j,$$

where e_i is income elasticity for good i from equation (4).

Data

We use Nielsen Homescan data for 2002 to 2006 to estimate demand. Households in this dataset log their purchases of food items from retail stores using handheld scanners. Einav, Leibtag, and Nevo (2010) provide a detailed description of the price-recording process and the potential for recording error and find that such errors are comparable to those found in social economic datasets. A subset of the Homescan panel, called the Fresh Foods panel, reports purchases of foods with barcodes and random-weight foods. Because many fresh vegetables are random weight, we analyze purchase data reported by the Fresh Foods panel, which had about 9,000 unique households per year from 52 Nielsen markets and 9 remaining areas of the contiguous United States. Unfortunately, Nielsen stopped detailed reporting of random-weight purchases after 2006, making it impossible to estimate disaggregated fruit and vegetable demand using more recent data.

The dataset contains demographic information, including household annual income (in income brackets), household size, characteristics of male and female heads of household, presence and age groups of children, market area, and race. We applied the method developed by Stewart

Table 2. Income Growth by Quintile

	Period	Quintile					Average
		1st (bottom)	2nd	3rd	4th	5th (top)	
Growth Rate (annual)	1991–2013	-0.02%	0.00%	0.17%	0.41%	0.76%	0.27%
	1967–2013	0.29%	0.27%	0.44%	0.73%	1.03%	0.55%
	1967–1990	0.72%	0.64%	0.79%	1.10%	1.40%	0.93%
Growth rate (cumulative)	1991–2013	-0.48%	0.09%	3.76%	9.52%	18.18%	6.21%
	1967–2013	14.25%	13.19%	22.46%	39.55%	60.54%	30.00%
	1967–1990	18.06%	15.77%	19.84%	28.57%	37.76%	24.00%

Source: Historical Income Inequality Tables, U.S. Bureau of Labor Statistics (2017).

(1983) to impute a continuous household income based on reported income brackets and household demographics.¹⁰ By design, the imputed continuous income falls within the household’s reported income bracket but may be superior to alternative measures such as using the midpoint of the reported income range (Stewart, 1983). We divide households into income quintiles using data from the historical income inequality tables of the U.S. Bureau of Labor Statistics (2017) indexed to 2013 prices. Table 1 provides unweighted summary statistics, divided by quintile, on the eleven demographic variables used in equation (1), which describe monthly income, household size, census region (three dummy variables), race (four dummy variables), a dummy variable for having a female head under the age of 40,¹¹ and a dummy variable for having a female head of household with a college degree. Table 1 also shows that number of households and number of observations in each income quintile. Although Nielsen aims to make the full static and Fresh Foods panels nationally representative, table 1 shows that households in the 1st and 5th quintiles are generally underrepresented.¹² In all, 12,039 unique households are observed, but variation in reported income over the sample period in the panel allows some households to appear in multiple income quintile groups.¹³

As reported in table 2, the difference in income growth across income quintiles has been significant in recent decades. Between 1991 and 2013, incomes in the bottom two quintiles registered negligible growth. At the same time, income in the top 3rd, 4th, and 5th quintiles grew by approximately 0.17%, 0.41%, and 0.76% annually and 3.76%, 9.52%, and 18.18% in total. In contrast, between 1967 and 1990, the bottom three income quintiles showed annual growth rates (0.72%, 0.64%, and 0.79%) approximately half that of the top two quintiles (1.1% and 1.4%), although growth rates were generally higher for all groups in that period.

¹⁰ Specifically, we assume the latent unobserved continuous income is a linear function of household demographics including ages of household heads, age groups of children, dummy variables for household head percentage time employed and occupation types, race, census region, Hispanic origin, residence type, and internet access. We then assign each household the conditional expectation of the dependent variable as the imputed household income while restricting the imputed value to lie within its reported income bracket.

¹¹ While we feel that the information contained in the female head of the household under the age of forty is informative to the household’s decision-making process (by perhaps being correlated with the household’s constraints on time or its skill and experience at finding bargains) we acknowledge that the specific age and gender delineation is somewhat arbitrary. Other studies have used similar but slightly different specifications (e.g., Zhen et al., 2014), but we don’t think that altering the specification would change our results substantively.

¹² We did not use projection factors included in the Nielsen data in our analysis. In our empirical exercise, we calculate the elasticities for each income quintile and then use income growth within the quintiles to calculate the effect of income of demand growth for each of those subsets of consumers. In this way, we are essentially weighting the simulation estimate *ex post* to account for the differential demand response within quintile before calculating an average effect rather than weighting the observations *ex ante* to generate a representative income elasticity as a direct outcome of our demand model estimation.

¹³ The Nielsen panel allows for annual updating of household income, but, for most households, this variable does not change much over their times in the panel. Of the 12,039 households, 39.7% have only one income level reported, 31.8% have two, 20.1% have 3, 7.4% have four, and 1.0% have five.

Table 3. Household Vegetable Expenditures at the Mean and by Income Quintile

	Income Quintile					Mean	Income Quintile					Mean	
	1st	2nd	3rd	4th	5th		1st	2nd	3rd	4th	5th		
	9.12	10.76	11.96	13.48	14.56		11.64	11.44%	16.0%	17.7%	20.9%		23.5%
Vegetable expenditures (\$/household-month)													
	Conditional Budget Share					Percentage of Uncensored Observations							
1. Artichoke, fresh	0.3%	0.4%	0.3%	0.4%	0.7%	0.4%	1.1%	1.4%	1.5%	2.0%	2.4%	1.6%	
2. Artichoke, preserved	0.4%	0.4%	0.5%	0.7%	0.9%	0.5%	1.1%	1.3%	1.8%	2.5%	3.1%	1.8%	
3. Asparagus, fresh	1.4%	1.8%	2.0%	2.5%	3.4%	2.1%	4.8%	6.2%	7.3%	9.5%	13.1%	7.3%	
4. Asparagus, preserved	0.7%	0.8%	0.9%	0.7%	0.7%	0.8%	2.7%	2.9%	3.3%	2.6%	2.2%	2.9%	
5. Broccoli, fresh	2.9%	2.7%	2.8%	3.2%	3.6%	3.0%	14.4%	16.0%	17.7%	20.9%	23.5%	17.7%	
6. Broccoli, preserved	1.5%	1.8%	1.7%	1.6%	1.3%	1.6%	6.5%	8.1%	8.4%	8.5%	6.9%	8.0%	
7. Corn, fresh	1.6%	1.7%	1.6%	1.6%	1.7%	1.6%	8.7%	10.0%	10.3%	11.4%	12.6%	10.3%	
8. Corn, preserved	4.7%	4.6%	4.9%	4.6%	4.1%	4.7%	22.8%	24.9%	26.2%	25.5%	21.7%	24.9%	
9. Green beans, fresh	0.9%	1.1%	1.1%	1.2%	1.4%	1.1%	5.1%	6.5%	7.4%	8.9%	11.1%	7.3%	
10. Green beans, preserved	3.9%	4.2%	4.2%	3.9%	3.1%	4.0%	19.9%	21.9%	22.6%	21.7%	17.5%	21.5%	
11. Mushrooms, fresh	0.8%	0.8%	0.8%	1.0%	1.6%	0.9%	3.7%	4.2%	4.5%	5.5%	7.8%	4.6%	
12. Mushrooms, preserved	3.8%	4.4%	4.7%	5.3%	5.7%	4.7%	15.0%	18.9%	20.8%	23.3%	24.6%	20.1%	
13. Onions, fresh	7.2%	7.0%	6.9%	6.8%	6.7%	6.9%	41.2%	43.5%	44.6%	46.3%	46.1%	44.2%	
14. Onions, preserved	0.7%	0.7%	0.7%	0.7%	0.6%	0.7%	2.4%	2.9%	3.2%	3.7%	3.0%	3.1%	
15. Peppers, fresh	1.6%	1.8%	1.9%	2.1%	2.2%	1.9%	7.6%	9.4%	10.1%	11.7%	12.1%	10.0%	
16. Peppers, preserved	1.5%	1.3%	1.3%	1.0%	0.7%	1.2%	10.3%	9.9%	9.9%	8.1%	5.4%	9.3%	
17. Tomatoes, fresh	14.4%	14.1%	14.1%	14.9%	16.0%	14.5%	43.9%	45.9%	47.8%	50.3%	51.9%	47.4%	
18. Tomatoes, preserved	0.6%	0.5%	0.6%	0.5%	0.6%	0.6%	2.9%	2.8%	3.1%	3.0%	3.2%	3.0%	
19. Yams, fresh	2.3%	2.1%	1.7%	1.6%	1.6%	1.8%	11.6%	11.9%	10.5%	10.1%	10.4%	10.9%	
20. Yams, preserved	0.7%	0.6%	0.6%	0.4%	0.4%	0.5%	3.4%	3.4%	3.1%	2.7%	2.1%	3.1%	
21. Other vegetables	48.1%	47.5%	46.7%	45.2%	43.3%	46.5%	86.4%	87.7%	88.1%	88.1%	87.6%	87.7%	

Notes: Conditional budget share is conditional on total vegetable expenditures.

Item Aggregation and Price Indexes

While the underlying data are recorded on a daily basis, we aggregate to the monthly level to make estimation tractable. To keep the number of goods manageable, our demand system includes twenty fresh and preserved vegetables (created from ten vegetable commodities, each divided into fresh and preserved categories), one composite good for all other vegetables, and a numéraire good for a total of twenty-two goods. The composite vegetable includes both vegetables that lack analog fresh and preserved states—including spinach, lettuce, and potatoes—and fresh vegetables consumed by only a small portion of consumers—including radishes and beets.¹⁴ Table 3 provides vegetable expenditure shares by commodity, average household total vegetable expenditures, and the percentage of uncensored household purchases at monthly frequency.¹⁵ Excluding the other vegetable composite, fresh tomatoes are the largest vegetable category in terms of per household spending, followed by fresh onions. The large percentages of zero-purchase observations suggest that it is important to account formally for corner solutions when analyzing this dataset.

We include a numéraire good to represent consumption of all other goods and services for two reasons. First, it captures the cross-price effects between vegetables and outside goods. LaFrance and Hanemann (1989) prove that such a numéraire good allows the demand system to provide correct welfare measures, in contrast to conditional demand models that assume group expenditure (i.e., total expenditures on all goods within the conditional system) to be separable from the outside good. Second, it makes vegetable demand a direct function of income—a main explanatory variable of policy interest. In an application to demand for twenty-three packaged food groups, Zhen et al. (2014) found that such an unconditional model generated more plausible cross-price elasticity estimates than a model conditional on group expenditures. The numéraire good is constructed from the remaining income not allocated to vegetable expenditures. As a practical matter, this specification makes the budget shares allocated to the other twenty-one vegetables within the demand system very small but does not otherwise complicate the estimation or interpretation of parameters.

Food item identifiers in Homescan differ across product modules. Dry foods and frozen foods typically come with a UPC, and such items with the same UPC are fairly consistent in terms of weight, level of processing, and other product characteristics. With random-weight foods, product weights are random in that they can vary widely from package to package of the same product. There are also significant quality or processing variations across items in the same vegetable category. For example, fresh pre-cut broccoli has a higher price by weight but also has a larger edible portion than fresh whole broccoli. To aggregate the various UPC codes and product identifies into our twenty-one fresh and preserved vegetable categories, we obtained and utilized the product concordance previously developed by Stewart et al. (2011) using the same Nielsen Homescan dataset.¹⁶

¹⁴ Initially, we had a larger set of goods in our demand system, but this introduced a price observation problem. As described in Appendix A, if the consumer does not purchase a good within a commodity category, then we estimate the price that consumer observes for the commodity aggregate based on purchases by other consumers in that region. If consumers in that region do not purchase, we used the prices in neighboring regions and, if necessary, national prices. Since this estimation process for prices introduces potential sources of error into the demand system, we sought to cap its use by limiting the number of infrequently purchased goods.

¹⁵ We use the term “censored” somewhat liberally to refer to observations of zero purchase levels for specific commodities. More properly, these consumer choices may be considered corner solutions to the consumer choice problem (Wooldridge, 2002, p. 518).

¹⁶ In the Nielsen Homescan data, goods are identified either by the combination of a four-digit product module ID and five-digit product number or by an eleven-digit UPC. The concordance was developed by manually sorting product identifiers by broad categories and then manually validating that the product category matched the written product description. Fresh products have fewer unique product identifiers because they are not distinguished by size and less likely to be identified by brand. The median and mean number of unique product identifiers are 63 and 592 for the ten fresh vegetables and 1,735 and 2,314 for preserved products, respectively. We sincerely thank Hayden Stewart and Jeffery Hyman for their assistance with this product concordance. For a specific listing of the UPCs and module IDs in the commodity aggregations, please contact the authors.

Because Homescan records the quantity of and expenditure for items purchased, the per unit cost consumers pay for each good purchased can be calculated directly and interpreted as the price.¹⁷ If a consumer does not purchase an item, its price is missing. On average, consumers purchase from about five vegetable categories each month. To impute missing product-level prices, we use a simple linear model of prices estimated based on observed prices (details described in Appendix A). The product-level prices and household purchase quantities are then used to construct household-specific Fisher Ideal price indexes for the twenty-one vegetables in the demand system.¹⁸ Diewert (1976) showed that the Fisher Ideal price index is superlative in that it is exact for a second-order approximation to a twice-differentiable linear homogeneous cost function. This property allows us to construct a utility-theoretic cost-of-living (COL) index for each of the twenty-one vegetables without having to estimate explicitly twenty-one product-level demand models.

To calculate a price index for the numéraire good, we first create a spatial dataset of county-level COL indexes for all goods and services by linking average COL index values for over 400 urban areas tracked by the Council for Community and Economic Research (C2ER) over the 2002–2006 period to the nearest counties. The spatial county-level price data are then combined with the Bureau of Labor Statistics' regional time-series consumer price index to construct a time-series and cross-sectional dataset of the COL index for all goods and services. Using the COL index and Fisher Ideal price indexes for the twenty-one vegetables, we then applied the method first proposed by Wohlgenant (1989, p. 172) to back out a price index for the numéraire good.¹⁹

Demand Estimation and Results

We apply the quasi-maximum likelihood estimator for correlated random effects developed by Meyerhoefer, Ranney, and Sahn (2005) to estimate the censored EASI demand. We account for clustering at the household level by constructing the cluster-robust covariance matrix estimator (Huber, 1967; White, 1982; Williams, 2000). To allow for sufficient flexibility in estimating income effects, we include both higher-order polynomial terms for income (indexed by r) with the budget share equation (1) and higher-order polynomials (indexed by s) in the Tobit heteroskedasticity link function (2). Additional terms in the Tobit link function increase the flexibility in which income can affect the probability of purchasing along the income spectrum. While it may be tempting to maximize flexibility by setting r and s to very high values, our experience indicates that setting income polynomials too high may create a nonpositive, semidefinite variance-covariance matrix for the estimated parameters—a symptom of the difficulty introduced to the estimation process by multicollinearity among the income polynomial terms.²⁰

To determine the value of an additional income parameter with the EASI demand model, we successively tested the joint significance of the b_{Rj} coefficients by minimum distance (Wooldridge, 2002, p. 444). Under the null that the R th degree of income polynomial is excludable, the test statistic is asymptotically distributed as $\chi^2(J - 1)$, where $J = 21$, the number of equations estimated (the budget share equation of the numéraire is omitted in estimation). After successively adding higher-degree polynomials in our estimation and testing for their joint significance, we found that when R is set to 4 and S to 3, the test statistic is 125.475 (p-value < 0.000). We settled on these values of R and S for the final specification of demand because further increases in the values of R and S caused difficulty in estimating the budget share equations.

¹⁷ This price is inclusive of any discounts based on volume or value-club membership that the consumer might receive.

¹⁸ See Zhen et al. (2014, equation 2) for the formula for a Homescan-based Fisher Ideal price index.

¹⁹ Zhen et al. (2014, p. 9) detailed the application of Wohlgenant's method using Homescan and C2ER data.

²⁰ We center log income, $\ln(x)$, and real income, y , around their means to reduce multicollinearity among the polynomial terms.

Parameter Estimates

With twenty-two goods in the demand systems ($J = 22$), there are 1,071 free parameters after imposing symmetry, homogeneity, and adding-up conditions on the latent demand. Of the 1,071 parameters, $(J(J - 1))/2$ are price parameters, $4(J - 1)$ are income parameters, $11(J - 1)$ are demographic parameters (including an intercept), $21(J - 1)$ are parameters on household-specific mean prices used to control for time-invariant household effects, and $4(J - 1)$ are parameters (γ) specifying the link function for variance of the residual in the Tobit model.

The parameter estimates and their cluster-robust variances are presented in the Online Supplement. While accounting for household-level clustering increases our standard errors, most of our main economic parameters are statistically significant. In our estimation, 173 of 231 (74.9%) price parameters, 120 of 231 (51.9%) demographic parameters, and 365 of 441 (82.5%) mean price parameters are significantly different than 0 at the 5% level. All of the first-order income terms for the twenty-one goods are significant, but only 46 of 63 (73.0%) higher-order income terms are. Similarly, all of the first-order coefficients in the heteroskedastic link function (γ_{1j}) were significant and 40 of the 42 γ_{sj} ($s = 2, 3$) coefficients on higher-order nominal income terms were significant.

Price and Income Elasticity Estimates

Because elasticities cannot be obtained using observed budget shares when the shares are zero, we calculated expected elasticities by replacing W with conditional means of observed budget shares and substituting marginal effects of log prices and real-income polynomials on these conditional means for a_{ij} and b_{rj} in equations (3) and (4), respectively. We calculated expected price and income elasticities at all observations.

Table 4 provides the median values of the Marshallian price elasticities.²¹ With the exception of fresh broccoli, demand is elastic for twenty of the twenty-one vegetable goods. Preserved and fresh offerings of the same vegetable are substitutes for five of the ten vegetable pairs. Table 5 presents the median income elasticities overall and by quintile along with their confidence intervals based on a Monte Carlo simulation.²² For all twenty fresh and preserved vegetables, overall median income elasticities are less than 1, consistent with the expectations of Engel's law. Preserved peppers and preserved asparagus appear to be inferior goods with negative overall income elasticities, but trivial in economic significance.

In terms of income elasticities by quintile, several trends are apparent. First, in nearly all cases, fresh and preserved vegetables are primarily luxuries in the economic sense for the lowest two income quintiles, as income elasticities exceed one in thirty-three of forty instances. Second, fresh and preserved vegetables are normal or inferior goods in most cases for the top three quintiles, as income elasticities are less than zero in thirty-seven of sixty cases and below one in all but eight cases. While this striking disparity aligns broadly with the empirical implications of Engel's Law on a household basis, it has strong implications for in-store vegetable purchases if income growth is uneven across quintiles.

Effects of Income Growth

The growth rate in demand that can be attributed to income growth across quintiles is found by multiplying the income elasticities by the growth rates of median income for each quintile and taking their average. The difference between the income elasticities for a fresh and a preserved vegetable

²¹ We account for the effect of log income in the heteroskedasticity link function on income elasticities.

²² Specifically, we draw 110 sets of demand-system parameters based on our estimated parameter and covariance estimates, calculate new elasticity estimates at every observation with each parameter set, and calculate median elasticities for each good based on each income quintile in the manner identical to our point estimates. We then calculate a mean and standard deviation based on the 110 estimates and report the range of the 95% confidence interval below the point estimate.

Table 4. Median Price Elasticities

With respect to the price of

Elasticity of the quantity of	1. Artichoke, Fr.	2. Artichoke, Pres.	3. Asparagus, Fr.	4. Asparagus, Pres.	5. Broccoli, Fr.	6. Broccoli, Pres.	7. Corn, Fr.	8. Corn, Pres.	9. Gr. Beans, Fr.	10. Gr. Beans, Pres.	11. Mushrooms, Fr.	12. Mushrooms, Pres.	13. Onions, Fr.	14. Onions, Pres.	15. Peppers, Fr.	16. Peppers, Pres.	17. Tomatoes, Fr.	18. Tomatoes, Pres.	19. Yams, Fr.	20. Yams, Pres.	21. Other Veg. Agg.	22. All Other Goods
1. Artichoke, Fr.	-1.633	-0.034	-0.268	-0.026	0.039	0.016	-0.015	-0.037	0.085	0.078	-0.125	-0.069	0.018	-0.006	-0.050	0.010	0.023	-0.111	0.045	-0.049	0.072	1.058
2. Artichoke, Pres.	-0.026	-1.497	-0.016	-0.168	-0.145	-0.062	0.038	-0.057	-0.037	-0.016	-0.068	-0.024	0.028	-0.052	-0.023	-0.082	0.000	0.059	-0.098	-0.004	0.004	1.633
3. Asparagus, Fr.	-0.231	-0.018	-2.025	-0.006	-0.087	-0.051	-0.023	-0.010	0.042	0.024	-0.020	-0.001	0.018	-0.011	0.007	-0.017	-0.108	0.023	-0.042	0.034	-0.049	1.959
4. Asparagus, Pres.	-0.020	-0.175	-0.005	-1.369	-0.047	-0.061	-0.020	-0.007	-0.005	-0.083	-0.011	0.004	-0.009	-0.088	0.037	0.049	-0.022	-0.174	0.029	-0.108	0.072	2.207
5. Broccoli, Fr.	0.053	-0.265	-0.137	-0.083	-0.908	0.067	-0.051	0.055	0.000	-0.170	0.089	0.013	0.013	-0.033	-0.148	-0.127	0.002	0.092	-0.179	-0.099	0.039	1.440
6. Broccoli, Pres.	0.017	-0.091	-0.064	-0.086	0.053	-1.311	-0.065	-0.082	0.008	-0.046	0.009	0.036	-0.020	-0.010	-0.070	-0.046	0.009	0.049	-0.110	-0.123	0.100	1.545
7. Corn, Fr.	-0.022	0.073	-0.038	-0.038	-0.055	-0.087	-1.466	0.161	-0.596	0.223	-0.012	-0.002	-0.007	0.084	-0.161	0.146	-0.286	0.155	0.346	0.196	0.137	0.827
8. Corn, Pres.	-0.045	-0.095	-0.014	-0.012	0.050	-0.092	0.135	-2.090	0.048	-0.075	0.102	-0.080	0.099	0.012	-0.009	0.024	0.158	-0.001	-0.154	-0.110	0.162	1.673
9. Green Beans, Fr.	0.121	-0.070	0.069	-0.009	0.000	0.010	-0.584	0.057	-1.200	0.149	-0.066	-0.060	0.089	0.076	0.030	-0.082	0.001	0.088	0.036	-0.035	0.185	0.596
10. Green Beans, Pres.	0.089	-0.025	0.032	-0.123	-0.143	-0.049	0.175	-0.070	0.120	-1.556	0.079	-0.060	0.078	-0.024	-0.134	-0.139	0.190	0.070	-0.209	-0.131	0.168	1.480
11. Mushrooms, Fr.	-0.150	-0.109	-0.028	-0.018	0.078	0.010	-0.010	0.100	-0.055	0.082	-1.219	0.018	0.032	-0.001	-0.038	0.025	-0.041	-0.015	-0.015	0.035	0.087	0.997
12. Mushrooms, Pres.	-0.067	-0.031	-0.001	0.005	0.009	0.032	-0.001	-0.063	-0.041	-0.051	0.015	-1.412	0.064	-0.017	0.035	0.013	0.084	-0.071	0.095	-0.075	0.157	1.051
13. Onions, Fr.	0.025	0.053	0.029	-0.017	0.013	-0.026	-0.006	0.114	0.088	0.096	0.037	0.092	-2.357	-0.059	0.127	-0.092	0.169	0.023	0.116	0.025	-0.013	1.119
14. Onions, Pres.	-0.007	-0.074	-0.014	-0.121	-0.026	-0.010	0.061	0.011	0.056	-0.022	-0.001	-0.019	-0.044	-1.133	0.025	0.009	-0.034	0.002	-0.092	-0.053	0.001	0.936
15. Peppers, Fr.	-0.059	-0.036	0.009	0.056	-0.127	-0.075	-0.129	-0.008	0.025	-0.137	-0.037	0.042	0.106	0.027	-1.224	-0.069	0.175	0.071	0.003	-0.153	0.101	1.012
16. Peppers, Pres.	0.018	-0.198	-0.035	0.115	-0.168	-0.076	0.180	0.035	-0.104	-0.218	0.038	0.024	-0.117	0.015	-0.107	-2.893	0.118	-0.068	-0.148	-0.109	0.065	3.899
17. Tomatoes, Fr.	0.017	0.000	-0.090	-0.021	0.001	0.005	-0.141	0.092	0.000	0.119	-0.025	0.061	0.086	-0.023	0.106	0.046	-1.579	0.057	0.125	-0.059	0.182	0.711
18. Tomatoes, Pres.	-0.122	0.087	0.029	-0.249	0.075	0.050	0.117	-0.001	0.068	0.067	-0.014	-0.080	0.018	0.002	0.067	-0.041	0.088	-1.426	0.002	-0.093	-0.030	1.260
19. Yams, Fr.	0.059	-0.174	-0.063	0.048	-0.173	-0.132	0.312	-0.165	0.033	-0.239	-0.017	0.129	0.109	-0.114	0.004	-0.108	0.230	0.003	-1.945	-0.238	0.078	2.138
20. Yams, Pres.	-0.064	-0.008	0.051	-0.183	-0.095	-0.147	0.177	-0.117	-0.052	-0.150	0.038	-0.101	0.024	-0.065	-0.172	-0.080	-0.107	-0.110	-0.238	-1.655	0.013	3.109
21. Other Veg. Agg.	0.026	0.001	-0.021	0.033	0.010	0.032	0.034	0.047	0.046	0.052	0.026	0.058	-0.004	0.000	0.031	0.013	0.092	-0.010	0.021	0.003	-1.351	0.589
22. All Other Goods	0.001	0.001	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.001	0.001	0.000	-0.987

Notes: All price elasticities are Marshallian elasticities.

Table 5. Median Income Elasticities Estimates, by Quintile and Overall

	Quintile					Overall
	1st (bottom)	2nd	3rd	4th	5th (top)	
1. Artichoke, Fresh	1.737 [0.985,2.305]	1.106 [0.782,1.4]	0.676 [0.364,0.998]	0.730 [0.095,1.261]	1.213 [0.755,1.553]	0.997 [0.754,1.182]
2. Artichoke, Preserved	2.739 [2.26,3.176]	1.664 [1.396,1.912]	0.283 [0.031,0.531]	-0.137 [0.528,0.218]	0.211 [-0.05,0.592]	0.726 [0.578,0.888]
3. Asparagus, Fresh	2.312 [2.083,2.547]	1.415 [1.244,1.562]	0.022 [-0.198,0.206]	-0.291 [-0.8,0.108]	0.674 [0.365,1.091]	0.695 [0.564,0.806]
4. Asparagus, Preserved	2.552 [2.013,3.001]	1.415 [1.088,1.706]	-0.580 [-0.855, -0.257]	-1.378 [-1.849, -1.001]	-0.794 [-1.503, -0.187]	-0.005 [-0.156,0.178]
5. Broccoli, Fresh	1.981 [1.751,2.183]	1.221 [1.041,1.383]	-0.535 [-0.775, -0.247]	-0.478 [-1.075,0.055]	0.500 [-0.132,1.082]	0.413 [0.264,0.582]
6. Broccoli, Preserved	2.497 [2.217,2.761]	1.399 [1.199,1.591]	-0.855 [-1.054, -0.618]	-0.512 [-0.759, -0.227]	1.226 [0.422,1.972]	0.418 [0.307,0.525]
7. Corn, Fresh	1.874 [1.687,2.047]	1.027 [0.906,1.136]	-0.241 [-0.387, -0.121]	-0.334 [-0.883,0.057]	0.602 [-0.097,1.165]	0.461 [0.311,0.573]
8. Corn, Preserved	1.768 [1.623,1.937]	0.978 [0.887,1.089]	-1.179 [-1.405, -0.981]	-0.366 [-0.594, -0.122]	1.349 [0.863,1.851]	0.257 [0.203,0.311]
9. Green Beans, Fresh	2.515 [2.242,2.73]	1.594 [1.38,1.748]	-0.009 [-0.237,0.195]	-0.671 [-1.026, -0.336]	-0.118 [-0.491,0.275]	0.638 [0.484,0.762]
10. Green Beans, Preserved	1.929 [1.801,2.051]	0.997 [0.91,1.076]	-1.172 [-1.409, -0.949]	-0.810 [-1.29, -0.394]	1.304 [0.766,1.73]	0.165 [0.068,0.25]
11. Mushrooms, Fresh	0.176 [-0.944,1.158]	0.371 [0.031,0.681]	0.550 [0.243,0.869]	-1.225 [-1.884, -0.622]	0.024 [-0.678,0.618]	0.222 [-0.136,0.55]
12. Mushrooms, Preserved	2.153 [2.008,2.29]	1.308 [1.188,1.426]	-0.673 [-0.848, -0.488]	-0.750 [-1.08, -0.458]	0.516 [-0.105,1.089]	0.344 [0.211,0.461]
13. Onions, Fresh	1.156 [1.076,1.234]	0.687 [0.636,0.74]	-0.411 [-0.537, -0.271]	0.323 [0.086,0.566]	1.615 [1.369,1.873]	0.406 [0.333,0.487]
14. Onions, Preserved	2.300 [1.955,2.689]	1.585 [1.383,1.823]	0.172 [-0.058,0.37]	-1.105 [-1.502, -0.772]	-0.925 [-1.251, -0.395]	0.596 [0.459,0.749]
15. Peppers, Fresh	2.562 [2.364,2.808]	1.438 [1.303,1.593]	-0.819 [-1.07, -0.618]	-0.352 [-0.572, -0.128]	1.676 [1.171,2.333]	0.460 [0.365,0.551]
16. Peppers, Preserved	2.379 [2.151,2.611]	1.107 [0.953,1.271]	-2.149 [-2.608, -1.716]	-1.234 [-2.168, -0.458]	0.613 [-0.121,1.335]	-0.115 [-0.415,0.137]
17. Tomatoes, Fresh	1.204 [1.049,1.347]	0.716 [0.626,0.8]	-0.440 [-0.695, -0.219]	0.187 [-0.085,0.443]	1.079 [0.699,1.413]	0.348 [0.269,0.415]
18. Tomatoes, Preserved	1.236 [0.665,1.719]	0.564 [0.27,0.806]	-0.103 [-0.363,0.145]	-0.967 [-1.299, -0.657]	-0.551 [-1.024, -0.196]	0.153 [-0.018,0.276]
19. Yams, Fresh	2.436 [2.195,2.683]	1.269 [1.111,1.421]	-1.669 [-2.164, -1.222]	-1.139 [-1.654, -0.714]	1.847 [0.813,2.883]	0.141 [-0.019,0.283]
20. Yams, Preserved	2.775 [2.47,3.046]	1.644 [1.43,1.83]	-1.028 [-1.252, -0.82]	-1.977 [-2.666, -1.346]	-0.950 [-1.761, -0.241]	0.100 [-0.093,0.287]
21. Other Vegetables	0.542 [0.493,0.585]	0.422 [0.384,0.452]	0.184 [0.137,0.229]	0.080 [-0.035,0.187]	-0.018 [-0.203,0.147]	0.288 [0.249,0.325]
22. All Other Goods	0.971 [0.969,0.973]	0.976 [0.974,0.978]	0.980 [0.98,0.98]	0.985 [0.985,0.985]	0.989 [0.989,0.989]	0.979 [0.979,0.979]

Notes: Figures in brackets are 95% confidence intervals.

Table 6. Calculated Consumption/Expenditure Shifts from Income Increases, 1991–2013

Commodity	Percentage Change in Consumption Based on Median Income Growth by Quintile					Average across Quintiles	Median Income Growth	Average Income Growth
	1st (bottom)	2nd	3rd	4th	5th (top)			
1. Artichoke, Fresh	-0.84%	0.10%	2.54%	6.96%	22.06%	6.16%	3.67%	19.02%
2. Artichoke, Preserved	-1.32%	0.15%	1.06%	-1.31%	3.84%	0.49%	3.20%	16.58%
3. Asparagus, Fresh	-1.11%	0.13%	0.08%	-2.77%	12.25%	1.71%	2.78%	14.39%
4. Asparagus, Preserved	-1.23%	0.13%	-2.18%	-13.12%	-14.44%	-6.17%	0.82%	4.23%
5. Broccoli, Fresh	-0.95%	0.11%	-2.01%	-4.55%	9.09%	0.34%	1.81%	9.36%
6. Broccoli, Preserved	-1.20%	0.13%	-3.21%	-4.88%	22.28%	2.62%	2.52%	13.07%
7. Corn, Fresh	-0.90%	0.09%	-0.91%	-3.18%	10.94%	1.21%	1.97%	10.19%
8. Corn, Preserved	-0.85%	0.09%	-4.43%	-3.49%	24.53%	3.17%	1.71%	8.88%
9. Green Beans, Fresh	-1.21%	0.14%	-0.03%	-6.39%	-2.14%	-1.93%	2.23%	11.53%
10. Green Beans, Preserved	-0.93%	0.09%	-4.40%	-7.71%	23.72%	2.15%	1.51%	7.83%
11. Mushrooms, Fresh	-0.08%	0.03%	2.06%	-11.67%	0.43%	-1.84%	-0.07%	-0.36%
12. Mushrooms, Preserved	-1.04%	0.12%	-2.53%	-7.15%	9.37%	-0.24%	1.72%	8.89%
13. Onions, Fresh	-0.56%	0.06%	-1.54%	3.07%	29.37%	6.08%	2.27%	11.74%
14. Onions, Preserved	-1.11%	0.14%	0.65%	-10.52%	-16.81%	-5.53%	1.36%	7.06%
15. Peppers, Fresh	-1.23%	0.13%	-3.08%	-3.36%	30.47%	4.59%	3.03%	15.69%
16. Peppers, Preserved	-1.14%	0.10%	-8.07%	-11.75%	11.15%	-1.94%	0.48%	2.50%
17. Tomatoes, Fresh	-0.58%	0.06%	-1.65%	1.78%	19.61%	3.85%	1.84%	9.56%
18. Tomatoes, Preserved	-0.59%	0.05%	-0.39%	-9.22%	-10.01%	-4.03%	0.12%	0.62%
19. Yams, Fresh	-1.17%	0.11%	-6.27%	-10.85%	33.58%	3.08%	1.84%	9.56%
20. Yams, Preserved	-1.33%	0.15%	-3.86%	-18.83%	-17.28%	-8.23%	0.31%	1.61%
21. Other Vegetables	-0.26%	0.04%	0.69%	0.76%	-0.33%	0.18%	0.81%	4.21%
Average of 10 Fresh Vegetables						3.43%	2.04%	10.55%
Average of 10 Preserved Vegetables						0.48%	1.57%	8.14%

Notes: All values are calculated by multiplying median income elasticities with the corresponding income growth rates. Average of 10 Preserved Vegetables is weighted by vegetable budget shares.

gives an approximate estimate of the difference between consumption growth rates that can be attributed to a 1% change in income.²³ The cumulative effect will be larger if income changes appreciably. This method, which bases the expenditure growth estimates on income growth at the quintiles, contrasts with the simpler method of estimating expenditure growth by taking the overall median income elasticity and multiplying it by either the overall median or mean income growth.

Table 6 shows the difference in calculated expenditure changes using actual income changes between 1991 and 2013 based on these three methods. In 2013, median and average incomes were \$51,939 and \$84,687 per household, respectively; in 1991, those figures were \$50,249 and \$72,129, indicating growth rates of 3.36% and 17.41%, respectively. In the quintile method, the magnitude of expenditure change is much smaller for the bottom two quintiles owing to the minor income growth for these households. Moreover, because vegetables are normal goods for consumers in the lowest quintile, negative income growth is estimated to have reduced demand slightly. The driving effects, however, occurred with the upper three income quintiles. Using different median income growth rates (as they differ across income quintile), we estimate that fresh demand grew across all ten vegetable pairs by 3.43% while preserved demand grew by 0.48%, a difference of 2.95 percentage points. As a counterexample, when we naïvely assume that incomes grew uniformly across quintiles, this difference is far smaller. When income growth is assumed to be uniform across all income groups at the growth rate for the average household, we estimate fresh demand grew 10.55% while preserved demand grew 8.14%, a difference of 2.41 percentage points. When income growth is assumed to be uniform at the growth rate for the median household, we estimate that fresh demand grew 2.04% while preserved demand grew 1.57%, a difference of 0.47 percentage points. In short, assuming uniform growth across income groups yields a far smaller effect of income growth on the relative shares of fresh and preserved vegetables than when we allow them to differ within quintiles.

The intuition for this finding is in table 7, which shows the differences between the income elasticities between fresh goods and their preserved analogs and the difference in calculated expenditure change between 1991 and 2013 using the quintile method. For the bottom two income quintiles, the fresh income elasticity exceeded its preserved analogue in seven of twenty instances (35%). For the top income quintiles, the fresh income elasticity exceeded its preserved analogue in twenty-two of thirty instances (70%). Moreover, across the ten commodity pairs, the weighted average of the differences and the (related) differences of the weighted averages are negative for the bottom two quintiles and positive for the top three quintiles. This suggests a marginal income increase reduces the share of fresh vegetables for the bottom two income quintiles but increases it for the top three. However, little change in market-wide demand can be attributed to income change for consumers in the bottom income quintiles because those groups experienced very little income growth over the period in question. As a matter of interpretation for why low-income households (in the bottom two income quintiles) have higher income elasticities for preserved versus fresh vegetables (in a way not reflected in high-income households), we speculate that the convenience characteristics of preserved vegetables (greater flexibility in the time of use and shopping trip planning, less likelihood of waste as food goes bad) are more useful for low-income households than for high-income household (less prep time for readily consumable or lightly processed foods.)

Table 7 also shows the commodity-by-commodity effect of income growth for our ten vegetable pairs. In eight of these pairs, income growth causes a shift to fresh vegetables. The largest calculated changes occurred for asparagus, artichokes, onions, and yams; in these cases, income growth caused demand for fresh vegetables to grow at least 5% faster than demand for their preserved counterparts. Conversely, with broccoli and mushrooms, income growth caused demand to grow faster for preserved goods than fresh goods, although this rate was less than 2%.

Based on the estimated total share changes due to income growth in table 6, we estimate a predicted change in the fresh shares of vegetables and compare it to the actual change in that share as drawn from the *ERS Vegetable and Pulses Yearbook* (2016). For our estimate of the all-vegetables

²³ This method assumes that the income elasticities for the respective goods, estimated using 2002–2006 Nielsen data, remain constant throughout the period of our simulation (1991–2013.)

Table 7. Differences in Income Elasticities and in Expenditure Growth rates between Fresh and Preserved Vegetables

Commodity	Income Elasticity Difference between Fresh and Preserved Vegetables Quintiles					Overall Mean	Difference between Fresh and Preserved Expenditure Growth Rates
	1st	2nd	3rd	4th	5th		
Artichoke	-1.002	-0.558	0.392	0.867	1.002	0.140	5.7%
Asparagus	-0.240	0.000	0.602	1.087	1.468	0.583	7.9%
Broccoli	-0.516	-0.178	0.321	0.035	-0.726	-0.213	-2.3%
Corn	0.106	0.048	0.938	0.032	-0.747	0.075	-2.0%
Green Beans	0.586	0.597	1.163	0.139	-1.422	0.213	-4.1%
Mushrooms	-1.977	-0.937	1.223	-0.474	-0.492	-0.531	-1.6%
Onions	-1.143	-0.898	-0.583	1.428	2.540	0.269	11.6%
Peppers	0.183	0.331	1.330	0.881	1.062	0.757	6.5%
Tomatoes	-0.033	0.152	-0.338	1.155	1.630	0.513	7.9%
Yams	-0.339	-0.375	-0.640	0.838	2.797	0.456	11.3%
Wtd. Avg of Diff's	-0.369	-0.138	0.298	0.649	0.768	0.160	
Diff. of Wtd. Avg.	-0.575	-0.266	0.509	0.668	0.320	0.144	

Notes: Growth rates are cumulative for 1991–2013. The weighted average of the differences computes the difference between the fresh and preserved income elasticity for each commodity and averages of those differences (weighted by the total for each commodity pair). The difference of the weighted averages first computes the averages of the fresh income elasticities (weighted by their shares of fresh consumption) and preserved income elasticities (weighted by their shares of preserved consumption) and then takes the difference of the two.

Table 8. Expenditure Share Changes Attributable to Income Changes

Year	Observed Fresh Share		Predicted 2013 Shares	Predicted Difference between 1991 and 2013 Shares	Actual Difference
	1991	2013			
Asparagus, Fr	61.9%	86.7%	63.7%	1.9%	24.8%
Broccoli, Fresh	57.5%	73.6%	56.9%	-0.6%	16.2%
Corn, Fresh	22.4%	41.0%	22.1%	-0.3%	18.5%
Green Bean, Fresh	16.0%	24.6%	15.5%	-0.5%	8.6%
Onions, Fresh	40.9%	45.1%	43.7%	2.8%	4.2%
Tomatoes, Fresh	16.6%	23.5%	17.7%	1.1%	6.9%
All Vegetables Fresh	47.0%	56.0%	47.7%	0.7%	9.0%

Notes: Predicted differences are based on EASI demand elasticities and historical income growth. Actual differences are based on the USDA's *Vegetable and Pulses Yearbook* (U.S. Department of Agriculture, Economic Research Service, 2016).

aggregate, we used our share weighted average of the predicted income effects for the ten vegetable pairs. Table 8 shows the actual fresh shares from 1991 and 2013, the actual change in fresh shares, and our predicted change in 2013 budget shares resulting from income growth.²⁴ Unfortunately, actual share changes for specific commodities are only available for six of our ten pairs of fresh and preserved goods over the full time period.

Based on vegetables aggregate, we estimate that, when income growth is specified as differing across quintiles, the fresh share of vegetable expenditure grows from 47.0% to 47.7% based on the 3.43% and 0.48% average growth rate of the fresh and preserved vegetable consumption (table 6). Therefore, when its unevenness is accounted for, income growth accounts for 0.7 percentage points out of the 9.0 percentage point total increase in the fresh consumption share between 1991 and 2013. When income is assumed to grow uniformly at the median growth rate across all income quintiles, the fresh share increases from 47.0% to 47.1% based on the 2.04% and 1.57% average growth rates

²⁴ Importantly, ERS yearbook data are based on farm-gate-level sales, while our demand estimates are based on household-level purchases in stores. In making this comparison, we assume that the proportion of fresh to preserved vegetables in food away from home consumption (i.e., restaurants, cafeterias) follows the same trends as vegetables purchased for at-home consumption.

of fresh and preserved vegetable consumption. Hence, when its unevenness is not accounted for, income growth amounts to only 0.1 percentage points of the 9.0 percentage point increase in the fresh consumption share. In our disaggregated analysis of six fresh and preserved vegetables pairs, income growth causes a shift in the fresh budget share to grow for three goods (asparagus, 1.9; onions, 2.8; and tomatoes 1.1 percentage points) but income growth caused the fresh share to decrease for three goods (broccoli, -0.6, corn, -0.3, and green beans, -0.5 percentage points). For onions, income growth caused the largest share change relative to the total share change (2.6 points of the total 4.2 percentage point shift). Nonetheless, in both the aggregate and individual cases, income explains, at most, a small portion of the total shift. The convenience, quality, and price tradeoffs between fresh and preserved goods differ substantially across vegetables. Having controlled for price effects in the demand system, our analysis indicates that the demand for convenience and quality attributes embodied in fresh and preserved vegetables can differ substantially across income groups, possibly due to the manner in which different income groups view convenience.

Conclusion

Our results suggest that the effect of income on commodity demand can be obscured if sufficient flexibility is not incorporated in policy analysis. This problem is exacerbated if simulation exercises ignore the striking unevenness of recent U.S. income growth. As incomes grow, consumption shares shift to goods with larger income elasticities. As a general pattern, we find that households in the top three income quintiles have higher income elasticities for fresh vegetables compared to preserved vegetables, but this pattern does not hold for households in the lower quintiles. Because income growth was far larger for the top three income quintiles between 1991 and 2013, we find that income growth explains 0.7 percentage point of the 9.0 percentage point total increase in share of vegetables consumed fresh. When the unevenness of income growth across income quintiles is ignored, median income growth explains only 0.1 percentage point of the 9.0 percentage point change in that share. We find that incorporating diversity in income growth across consumer groups is necessary whenever applied economic analysis is likely to have large effects on predictions based on the model.

Given the relatively small effect of income growth compared to the total change in the fresh consumption share, other factors have also likely contributed to the shift. Unlike income, however, factors affecting price and consumer preferences are more difficult to isolate causally in a generalized demand framework. For instance, increased import access for specific fresh vegetables that face phytosanitary restrictions may be more likely to be sought after a country obtains a more generalized tariff reductions with a trade agreement (Peterson et al., 2013), and such trade access may be affected by idiosyncratic factors specific to the commodity (Ferrier, 2014). The effect of improved trade access may be difficult to distinguish from the effect of shipping and storage innovations (primarily affecting fresh products) that can simultaneously reduce prices and increase out-of-season availability.

Income growth allows consumers to purchase more convenience items. Historically, preserved foods were considered more convenient in that they allowed consumers to reduce the number of shopping trips and afforded them more flexibility in timing the use of foods to avoid waste. More recently, changes in store-level processing, supply chain technology, and packaging may have improved the convenience characteristics of fresh goods, particularly for higher-income consumers. Given the limited role of income in explaining the shift to fresh consumption and the ongoing concerns regarding its consequences for household nutrition and food loss, explaining why households are buying larger shares of fresh goods requires looking to other factors as well.

[Received October 2016; final revision received September 2017.]

References

- Awuah, G. B., H. S. Ramaswamy, and A. Economides. "Thermal Processing and Quality: Principles and Overview." *Chemical Engineering and Processing: Process Intensification* 46(2007):584–602. doi: 10.1016/j.cep.2006.08.004.
- Banks, J., R. Blundell, and A. Lewbel. "Quadratic Engel Curves and Consumer Demand." *Review of Economics and Statistics* 79(1997):527–539. doi: 10.1162/003465397557015.
- Bertail, P., and F. Caillavet. "Fruit and Vegetable Consumption Patterns: A Segmentation Approach." *American Journal of Agricultural Economics* 90(2008):827–842.
- Buzby, J. C., H. F. Wells, and J. Hyman. "The Estimated Amount, Value, and Calories of Postharvest Food Losses at the Retail and Consumer Levels in the United States." Economic Information Bulletin EIB-121, U.S. Department of Agriculture, Economic Research Service, Washington, DC, 2014. Available online at <https://www.ers.usda.gov/publications/pub-details/?pubid=43836>.
- Cawley, J., and C. Meyerhoefer. "The Medical Care Costs of Obesity: An Instrumental Variables Approach." *Journal of Health Economics* 31(2012):219–230.
- Christensen, L. R., D. W. Jorgenson, and L. J. Lau. "Transcendental Logarithmic Production Frontiers." *Review of Economics and Statistics* 55(1973):28–45.
- Creed, P. "Quality and Safety of Frozen Ready Meal." In D.-W. Sun, ed., *Handbook of Frozen Food Processing and Packaging*, Boca Raton, LA: CRC Press, 2005.
- Darmon, N., and A. Drewnowski. "Does Social Class Predict Diet Quality?" *American Journal of Clinical Nutrition* 87(2008):1107–1117.
- Deaton, A. "Policy Implications of the Gradient of Health and Wealth." *Health Affairs* 21(2002):13–30. doi: 10.1377/hlthaff.21.2.13.
- Deaton, A., and J. Muellbauer. *Economics and Consumer Behavior*. Cambridge: Cambridge University Press, 1980.
- Diewert, W. E. "Exact and Superlative Index Numbers." *Journal of Econometrics* 4(1976):115–145. doi: 10.1016/0304-4076(76)90009-9.
- Dong, D., and B.-H. Lin. "Fruit and Vegetable Consumption by Low-Income Americans: Would a Price Reduction Make a Difference?" Economic Research Report ERR-70, U. S. Department of Agriculture, Economic Research Service, Washington, DC, 2009. Available online at <https://www.ers.usda.gov/publications/pub-details/?pubid=46131>.
- Einav, L., E. Leibtag, and A. Nevo. "Recording Discrepancies in Nielsen Homescan Data: Are They Present and Do They Matter?" *Quantitative Marketing and Economics* 8(2010):207–239.
- Ferrier, P. "The Effects of Phytosanitary Regulations on U.S. Imports of Fresh Fruits and Vegetables." Economic Research Report ERR-168, U.S. Department of Agriculture, Economic Research Service, Washington, DC, 2014. Available online at <https://www.ers.usda.gov/publications/pub-details/?pubid=45223>.
- Ferrier, P., and C. Zhen. "The Producer Welfare Effects of Trade Liberalization When Goods Are Perishable and Habit-Forming: The Case of Asparagus." *Agricultural Economics* 45(2014):129–141. doi: 10.1111/agec.12020.
- Finkelstein, E. A., J. G. Trogdon, J. W. Cohen, and W. Dietz. "Annual Medical Spending Attributable to Obesity: Payer- and Service-Specific Estimates." *Health Affairs* 28(2009):w822–w831.
- Gorman, W. "Some Engel Curves." In A. Deaton, ed., *Essays in the Theory and Measurement of Consumer Behaviour: In Honour of Sir Richard Stone*, Cambridge, UK: Cambridge University Press, 1981, 7–30.
- Gustavsen, G. W., and K. Rickertsen. "A Censored Quantile Regression Analysis of Vegetable Demand: The Effects of Changes in Prices and Total Expenditure." *Canadian Journal of Agricultural Economics/Revue Canadienne d'Agroéconomie* 54(2006):631–645.
- Heckman, J., and E. Vytlacil. "Instrumental Variables Methods for the Correlated Random Coefficient Model: Estimating the Average Rate of Return to Schooling When the Return is Correlated with Schooling." *Journal of Human Resources* 33(1998):974–987.

- Huber, P. J. "The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions." In L. M. Le Cam, ed., *Proceedings*, vol. 2/1, Contributions to Probability Theory. Berkeley: University of California Press, 1967, 221–223.
- Just, D. R., J. Lund, and J. Price. "The Role of Variety in Increasing the Consumption of Fruits and Vegetables Among Children." *Agricultural and Resource Economics Review* 41(2012):72–81. doi: 10.1017/S1068280500004196.
- Kasteridis, P., and S. T. Yen. "U.S. Demand for Organic and Conventional Vegetables: A Bayesian Censored System Approach." *Australian Journal of Agricultural and Resource Economics* 56(2012):405–425. doi: 10.1111/j.1467-8489.2012.00589.x.
- Kurlansky, M. *Birdseye: The Adventures of a Curious Man*. New York: Penguin-Random House, 2009.
- LaFrance, J. T., and W. M. Hanemann. "The Dual Structure of Incomplete Demand Systems." *American Journal of Agricultural Economics* 71(1989):262–274. doi: 10.2307/1241583.
- Lee, L.-F., and M. M. Pitt. "Microeconomic Demand System with Binding Nonnegativity Constraints: The Dual Approach." *Econometrica* 54(1986):1237–1242. doi: 10.2307/1912330.
- Lewbel, A. "A Demand System Rank Theorem." *Econometrica* 57(1989):701–705. doi: 10.2307/1911061.
- Lewbel, A., and K. Pendakur. "Tricks with Hicks: The EASI Demand System." *American Economic Review* 99(2009):827–863. doi: 10.1257/aer.99.3.827.
- Lucier, G., S. Pollack, M. Ali, and A. Perez. "Fruit and Vegetable Backgrounder." Vegetables and Pulses Outlook VGS-31301, U.S. Department of Agriculture, Economic Research Service, Washington, DC, 2006. Available online at <https://www.ers.usda.gov/publications/pub-details/?pubid=39508>.
- Meyerhoefer, C. D., C. K. Ranney, and D. E. Sahn. "Consistent Estimation of Censored Demand Systems Using Panel Data." *American Journal of Agricultural Economics* 87(2005):660–672. doi: 10.1111/j.1467-8276.2005.00754.x.
- Moore, L. V., and F. E. Thompson. "Adults Meeting Fruit and Vegetable Intake Recommendations." *Morbidity and Mortality Weekly Report* 64(2015):709–713. Available online at <https://www.cdc.gov/mmwr/preview/mmwrhtml/mm6426a1.htm>.
- Niu, L., and M. K. Wohlgenant. "A Censored Demand System Analysis of Fruits and Vegetables by Different Income Groups Using Micro Data." 2012. Paper presented at the annual meeting of the Agricultural & Applied Economics Association, August 12–14, Seattle, Washington.
- Nzaku, K., J. E. Houston, and E. G. Foseh. "Analysis of U.S. Demand for Fresh Fruit and Vegetable Imports." *Journal of Agribusiness* 28(2010):163–181.
- Perali, F., and J.-P. Chavas. "Estimation of Censored Demand Equations from Large Cross-Section Data." *American Journal of Agricultural Economics* 82(2000):1022–1037. doi: 10.1111/0002-9092.00100.
- Peterson, E., J. Grant, D. Roberts, and V. Karov. "Evaluating the Trade Restrictiveness of Phytosanitary Measures on U.S. Fresh Fruit and Vegetable Imports." *American Journal of Agricultural Economics* 95(2013):842–858. doi: 10.1093/ajae/aat015.
- Piggott, N. E. "The Nested PIGLOG Model: An Application to U.S. Food Demand." *American Journal of Agricultural Economics* 85(2003):1–15. doi: 10.1111/1467-8276.00099.
- Pollack, S. L. "Consumer Demand for Fruit and Vegetables: The U.S. Example." In A. Regmi, ed., *Changing Structure of Global Food Consumption and Trade*, No. WRS-01-1 in Agriculture and Trade Report. Washington, DC: U.S. Department of Agriculture, Economic Research Service, 2001, 49–54.
- Pollan, M. *In Defense of Food: An Eater's Manifesto*. New York: Penguin Press, 2008.
- Regmi, A., and M. Gehlhar. "Consumer Preferences and Concerns Shape Global Food Trade." *Food Review: The Magazine of Food Economics* 24(2001):2–8.

- Rickard, B. J., J. Liaukonyte, H. M. Kaiser, and T. J. Richards. "Consumer Response to Commodity-Specific and Broad-Based Promotion Programs for Fruits and Vegetables." *American Journal of Agricultural Economics* 93(2011):1312–1327. doi: 10.1093/ajae/aar042.
- Sagar, V. R., and P. S. Kumar. "Recent Advances in Drying and Dehydration of Fruits and Vegetables: A Review." *Journal of Food Science and Technology* 47(2010):15–26. doi: 10.1007/s13197-010-0010-8.
- Stack, S. *AMA Supports Newest Dietary Guidelines to Improve Public Health* | American Medical Association. Chicago, IL: American Medical Association, 2016. Available online at <https://www.ama-assn.org/content/ama-supports-newest-dietary-guidelines-americans-improve-public-health>.
- Stewart, H., J. Hyman, J. C. Buzby, E. Frazão, and A. Carlson. "How Much Do Fruits and Vegetables Cost?" Economic Information Bulletin EIB-71, U.S. Department of Agriculture, Economic Research Service, Washington, DC, 2011. Available online at <https://www.ers.usda.gov/publications/pub-details/?pubid=44520>.
- Stewart, M. B. "On Least Squares Estimation when the Dependent Variable is Grouped." *Review of Economic Studies* 50(1983):737–753. doi: 10.2307/2297773.
- Strazdins, L., A. L. Griffin, D. H. Broom, C. Banwell, R. Korda, J. Dixon, F. Paolucci, and J. Glover. "Time Scarcity: Another Health Inequality?" *Environment and Planning A* 43(2011):545–559. doi: 10.1068/a4360.
- Stuart, T. *Waste: Uncovering the Global Food Scandal*. New York: W. W. Norton & Company, 2009.
- Union of Concerned Scientists. *The \$11 Trillion Reward: How Simple Dietary Changes Can Save Lives and Money, and How We Get There*. Cambridge, MA: Union of Concerned Scientists, 2013. Available online at http://www.ucsusa.org/sites/default/files/legacy/assets/documents/food_and_agriculture/11-trillion-reward.pdf.
- U.S. Bureau of Labor Statistics. *Income Data Tables: Income Inequality (Table H-1. Income Limits for Each Fifth - All Races)*. Washington, DC: U.S. Census Bureau, 2017.
- U.S. Department of Agriculture, Economic Research Service. *Vegetables and Pulses Yearbook*. Washington, DC: U.S. Department of Agriculture, 2016. Available online at <http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1858>.
- U.S. Department of Health and Human Services, and U.S. Department of Agriculture. *Dietary Guidelines for Americans 2015–2010*. Washington, DC: Government Printing Office, 2015, 8th ed. Available online at <https://health.gov/dietaryguidelines/2015/guidelines/>.
- White, H. "Maximum Likelihood Estimation of Misspecified Models." *Econometrica* 50(1982):1–25. doi: 10.2307/1912526.
- Williams, R. L. "A Note on Robust Variance Estimation for Cluster-Correlated Data." *Biometrics* 56(2000):645–646. doi: 10.1111/j.0006-341X.2000.00645.x.
- Wohlgenant, M. K. "Effects of the Changing Composition of Beef Consumption on the Elasticities for Beef and Poultry." In R. C. Buse, ed., *The Economics of Meat Demand*, Madison, WI: University of Wisconsin, 1989, 170–186.
- Wooldridge, J. M. *Econometric Analysis of Cross-Section and Panel Data*. Cambridge, MA: MIT Press, 2002.
- World Health Organization. *Increasing Fruit and Vegetable Consumption to Reduce the Risk of Noncommunicable Diseases*. Rome: World Health Organization, 2014. Available online at http://www.who.int/elena/titles/fruit_vegetables_ncds/en/.
- Zhen, C., E. A. Finkelstein, J. M. Nonnemaker, S. A. Karns, and J. E. Todd. "Predicting the Effects of Sugar-Sweetened Beverage Taxes on Food and Beverage Demand in a Large Demand System." *American Journal of Agricultural Economics* 96(2014):1–25. doi: 10.1093/ajae/aat049.
- Zimmerman, C. C. "Ernst Engel's Law of Expenditures for Food." *Quarterly Journal of Economics* 47(1932):78–101. doi: 10.2307/1885186.

Appendix A: Imputing Missing Prices for Categories Not Purchased by Homescan Households

We estimate missing prices with the following regression:

$$(A1) \quad p_{hjt} = \beta_0 + \beta_t \times time_t + \beta_m \times ID_m + \beta_j \times SC_j + \beta_{jt} \times time_t \times SC_j + \beta_{mt} \times ID_m \times time_t,$$

where p_{hjt} is the price for household h , product subcategory j in time period t ; $time$ is a dummy variable for the time period; SC is a dummy for the product subcategory; and ID is a dummy for each Homescan market (indexed by m). The estimated prices were created automatically in SAS using the PROC GLMSELECT function with $time$, ID , and SC as class variables.

In several instances, no household in a market purchased a specific subcategory of product in a given month. In these instances, SAS will not generate an estimate, because the relevant β_{jt} or β_{mt} terms are unidentified. For these missing prices, we sequentially dropped the market ID interaction (β_{mt}) and product subcategory interaction terms (β_{mt} and β_{jt}), re-estimated the model, and used the predicted prices. This imputation procedure assigns the same product-level price to all households that did not purchase the product in a particular market and time period.

