Title:
Measuring Willingness to Pay for Environmental Attributes in Seafood

Author:
Villas-Boas, Sofia B, UC Berkeley
Hilger, James
Stevens, Andrew
Hallstein, Eric

Publication Date:
02-08-2016

Series:
CUDARE Working Papers

Permalink:
http://escholarship.org/uc/item/6nh0g7fg

Measuring Willingness to Pay for Environmental Attributes in Seafood*

James Hilger, Eric Hallstein, Andrew Stevens, and Sofia B. Villas-Boas†

January 31st, 2015

We investigate whether consumers are willing to pay for sustainability in seafood purchases. To do this, we estimate a demand model of seafood counter purchases using a product-level scanner dataset that includes data both before and after the implementation of an independent, non-profit, third party seafood advisory and sustainability label. The label ratings are based on seafood species, catch method, production method, and country of origin. The label presents sustainability information to consumers through the use of a traffic light label color-based system, where a color rating is assigned and labeled to each seafood stock-keeping unit. Green represents the best choices, yellow represents the "proceed with caution" choices, and red represents the worst choices. Using our panel dataset of individual seafood purchases, we estimate a random utility choice model (RUM) of consumer demand for seafood products. Each seafood product is defined as a bundle of attributes, including price, species, gear type, environmental contaminant advisory information, and sustainability rating color. Model identification comes from the partial and random implementation of the label system across the retail chain. Empirical results suggest that consumers perceive seafood as a differentiated product category with respect to sustainability labels. A second stage GLS regression of product characteristics on the RUM product fixed effects point estimates indicate that consumers prefer selective harvest methods, wild caught seafood, and U.S. caught seafood.

---

* We thank the retailer for providing us the data and the California Sea Grant College Program for financial support in the form of a 2013 Special Focus Award for Villas-Boas, Hilger, and the trainee Stevens.

† Hilger: NOAA Fisheries, Southwest Fisheries Science Center, 8901 La Jolla Shores Dr., La Jolla, CA 92037 (e-mail: james.hilger@noaa.gov); Hallstein: The Nature Conservancy, California, 201 Mission Street, 4th Floor, San Francisco, CA 94105 (e-mail: ehallstein@tnc.org); Stevens: Department of Agricultural and Resource Economics, University of California at Berkeley, 207 Giannini Hall, Berkeley, CA 94720-3310 (e-mail: stevens@berkeley.edu); Villas-Boas: Department of Agricultural and Resource Economics, University of California at Berkeley, 226 Giannini Hall, Berkeley, CA 94720-3310 (e-mail: sberto@berkeley.edu).

1. Introduction

Commercial fisheries operate in a globally competitive marketplace characterized by factors including differential levels of both customer conservation concerns and regulatory oversight across species groups and management bodies (Smith, et. al, 2010; Costello, et. al., 2008; Watson and Pauly, 2001; Delgado, et. al., 2003). Environmental sustainability labels, or eco-labels, are one tool that commercial fishers employ to increase economic viability through product differentiation in terms of sustainability. Existing empirical research has illustrated that seafood eco-labels are associated with shifts in market demand from moderate to more sustainable choices (Teisl et al 2002; Roheim et al, 2011; Hallstein and Villas-Boas, 2013; Sogn-Grundvag et al, 2013; Sogn-Grundvag et al, 2015; Blomquist, Bartolino, and Waldo, 2015), demonstrating that consumer-focused mechanisms, such as eco-labels and certification, have market impacts and are a tool to be considered in fisheries management.

Existing research on consumer-focused mechanisms, like eco-labels, rely largely on attitudinal and knowledge surveys, consumer choice experiments, and experimental auctions (see e.g., Alfnes, et. al., 2006; Johnston and Roheim, 2006; Johnston et. al., 2001; Wessels, 2002; Wessels et. al., 1999). While these studies offer valuable insight and methodological approaches, one potential weakness is that they capture consumers' stated preferences and not actual behaviors. There can be wide disparities between consumers' stated preferences and their actual purchases (Hensher and Bradley, 1993).

In the revealed preference literature, hedonic price models (Asche and Guillen, 2012; Roheim et. al., 2011; Roheim et. al., 2007; Jaffry et. al., 2004; Carroll et. al., 2001; McConnell and Strand, 2000), demand system (Teisl et. al., 2002; Chiang et. al., 2015) and case study (Roheim, 2003) approaches have been used to estimate relative values for seafood product attributes such as catch method, fishing gear choice, country of origin, product color (of salmon), and environmental sustainability. The articles most relevant to our paper are Teisl et al. (2002), Roheim et al. (2011), and Hallstein and Villas-Boas

(2013).[1] Teisl et al. (2002) use consumer purchase data to confirm that the dolphin-safe tuna label increased the market share of canned tuna with the label. Roheim et al. (2011) apply an hedonic price function approach to scanner data on the sale of frozen, processed Alaskan Pollock in the London metropolitan market to estimate a statistically significant price premium for Marine Stewardship Council certification; and Hallstein and Villas-Boas (2013) estimate a reduced-form model to measure the difference-in-difference quantity responses to the sustainability label system studied in this paper. Hallstein and Villas-Boas (2013) surprisingly find that sales of yellow-rated labeled products decreased significantly in treatment stores relative to controls, while red- and green-rated labeled products saw no change in the quantity sold. We expand on the reduced-form evidence presented in Hallerstein and Villas-Boas (2013), and utilize the same primary data to estimate a structural utility theoretic consistent revealed-preference random utility model (RUM) of consumer choice. The expansion to the RUM consumer choice framework allows the estimation of the consumer's dollar value willingness to pay for the environmental information provided by the eco-labels and other product attributes that are observable to consumers. Estimation of dollar value willingness to pay measures is not possible from the reduced form approach presented in Hallerstein and Villas-Boas (2013).

The availability of information about a product does not necessarily mean consumers will incorporate it into their decisions and alter their behavior.[2] Our study expands upon these efforts by not only directly testing whether consumers directly incorporate the available information, but also exploiting the quasi-experimental nature of our research design to estimate whether consumers' implied sensitivities towards price and preferences for product attributes change with increased product label information.

---

[1] One has to look outside the seafood industry for additional empirically based studies on the efficacy of product labels. Studies on restaurants hygiene (Jin and Leslie, 2003), organic milk certification (Kiesel and Villas-Boas, 2007; Batte et al. 2007), wine quality (Hilger et al, 2011) and apparel (Nimon and Beghin, 1999) estimate significant relationships between attributes covered by these labels and price premiums or market share gains.

[2] As shown in a variety of settings, consumers do not always incorporate all available information (Ippolito and Mathios, 1995; Mathios, 2000). Teisl et al. (2002) is one of the few studies on the seafood industry using consumer purchase data to confirm that the dolphin-safe tuna label increased the market share of canned tuna. More recently, Shimschack et al. (2007) and Teisl et al. (2012), investigate the impact of seafood risk advisories for certain population groups.

Our contribution is to develop and estimate a structural model of seafood demand. That is, we analyze actual consumer retail supermarket shopping behavior for seafood products sold at the seafood counter of a coastal California supermarket chain to directly estimate revealed preference parameters and corresponding willingness to pay measures for color-ratings that proxy for sustainability characteristics. In so doing, we provide resource managers and policy makers with important information on the efficacy of eco-labels as well as a barometer reading on consumer preferences.

Empirical results suggest that consumers find the addition of a yellow label diminished from the value of seafood counter products, whereas there is no significant value associated with the application of red- or green-labels. Moreover, consumers would need to be offered an average discount of 3.56 dollars per pound to offset the impact of the addition of a yellow "proceed with caution" product label. Given the pre-labeling-period average product price of 11 dollars per pound, the percentage price discount needed to get people to continue buying the product post-implementation of the yellow-label is estimated to be roughly a third. In addition we find that, on average, consumers, prefer the use of selective harvest methods from wild fisheries, prefer wild caught seafood relative to farmed, and prefer seafood originating in the U.S.

The rest of the paper proceeds as follows. Section 2 provides more detailed background information of the study. Section 3 describes the empirical setting and the data, section 4 presents the model used, section 5 discusses the main results, and section 6 concludes.

2. Background

Using supermarket scanner data on seafood counter purchases, we estimate a demand model of consumer preferences for seafood products, where consumers' willingness to pay measures for sustainability characteristics of seafood are calculated by taking advantage of a phased roll-out of an environmental sustainability label system for seafood products. We define products as a bundle of attributes such as product type, harvest method, country of origin, and label score. The label score is communicated by means of a traffic light color rating assigned to a label on each stock-keeping unit (SKU) sold at the seafood counter based on seafood species, catch method, production method, and country or state of origin data. The labeling program is developed and by the independent group Fishwise. The label utilizes three color-based scores: Green, 'best choice,' means that the product is considered to be from a sustainable fishery. Yellow, `proceed with caution,' means that wild fish populations are healthy; however, other problems exist such as poor fisheries management.[3] Red, 'worst choice,' means that the wild fish populations are overfished and that the fishery may be characterized by other problems, such as habitat destruction. The supermarket chain randomly chose two of its stores in which to test posting the label advisory before rolling it out to all stores, creating a quasi experimental empirical setup of two treatment and eight control stores. As retail prices are common across all stores for each time period, retail prices are uncorrelated with the implementation of the label system treatment. As the data set does not contain observations for the entire year, we do not directly model seasonality; however the treatment-control approach controls for any seasonal variation effects within our data.

Prior to the in-store implementation of the sustainability labels, consumers observed species, retail price, country of origin, and catch method when making seafood choices. Using a multi-store, multi-product level panel scanner data set for individual purchases, we estimate a RUM of consumer demand for seafood following Barry's logit approach (1995) to estimate consumer preferences for the implemented seafood sustainability-rating labels, as well as the product attributes that are observable by all consumers. Using the fitted demand RUM estimates, we then estimate the marginal average willingness to pay for seafood receiving each sustainability-rating label level designation as the ratio of the marginal utility of each sustainability-rating level label designation relative to the marginal utility of price. Finally, we conduct a posterior analysis regressing product fixed effects values on time invariant characteristics, such as harvest method and country of origin.

3. Experimental setting and data

This study uses aggregate point-of-sale scanner data from a regional upscale US supermarket retailer (the Retailer) that operates stores in the San Francisco Bay area of

[3] Subsequent to our data collection, the yellow label's definition was changed from "Proceed with Caution" to "Good Alternative." This re-definition is discussed in our Results and Conclusion sections, 4 and 5 respectively.

California for 24 weeks during the period February 19th, 2006 to September 16th, 2006.[4] The Retailer piloted an established sustainability-labeling program, developed by FishWise,[5] for twelve weeks starting on June 10th, 2006, at two randomly chosen stores and then fully implemented the advisory label system at all ten stores. The gradual phase-in of the program created a quasi-experimental setup with two treatment stores and eight control stores.

The pilot-labeling program consisted of deploying a product-specific "traffic-light" sustainability label. In addition to the sustainability-rating label deployed within pilot stores during the treatment period, each seafood product offered at the seafood counter by the chain across all stores and time periods were labeled with the product species, the country of origin, the catch method, and the retail price in dollars per pound. Thus, an example of a typical label across all stores during all time periods would inform the consumer that the product is Petrale Sole, was caught in the United States (U.S.) using bottom trawl gear, and the price is $16.99 per pound. Additional labeling in pilot stores (treatment) during the treatment period would include the "traffic-light" environmental sustainability rating, such as a "yellow" "proceed with caution" rating. A consumer in a treatment store during the treatment period could choose from a large number of seafood counter products. Products often had multiple color-rated options within a seafood genus or species (e.g. salmon labels may vary by species, country of origin, stock, and gear type), and did not have seafood counter options without a color-rating label. The Retailer updated the prices weekly on Tuesday nights after the stores had closed. Seafood product label scores did not change at the seafood product level during the timeframe of our data.

In this analysis, we assume that the sustainability label system weakly increased the Retailer's customers' knowledge (information) about the environmental sustainability and healthiness of the available seafood products. Other possible sources of information about seafood sustainability include popular media, other sustainability labeling programs, and scientific literature. For example, starting in 1999, the Monterey Bay Aquarium's

---

[4] This data was previously analyzed in a reduced-form analysis by Hallstein and Villas-Boas (2013).

[5] Fishwise is a sustainable seafood consultancy that promotes the health and recovery of ocean ecosystems through environmentally responsible practices. Fishwise scores are assigned based on the Monterey Bay Aquarium methodology while taking into account additional third-party standards from organizations including the Marine Stewardship Council, the Aquaculture Stewardship Council, the Environmental Defense Fund, and government agencies.

Seafood Watch program began distributing pocket guides and doing outreach to inform consumers about seafood sustainability. We do not have information about which consumers shopping at the Retailer also utilized other sources of information. We discuss the potential impact that the baseline of treatment and control store consumer sustainability knowledge from other sources could have on our estimates in our interpretation of results and our robustness checks.

The analysis for this study uses several unique data sets. From the Retailer, we obtained scanner panel data for the weeks before and after the labeling roll out. This data contained the following fields: date (*T*=24*)*, store (*S*=10), unique product number (*J*=172), general seafood type (e.g. salmon), product name (e.g. King Salmon fillet), unit sales, dollar sales, full retail price per unit, discounted price per unit, and country of origin (N=7841).[6] Second, in order to control for differences in consumer socio-demographic characteristics between stores, zip code level socio-demographic data from the United States Census Bureau was compiled for all stores by store zip code. The above datasets were merged for the empirical analysis.

The Retailer did not adjust pricing and promotional activity in response to the color-rating labels during the weeks immediately before and after implementation of the labeling program. That is, marketing-mix practices such as pricing and promotions are initially exogenous to the label color-rating. In the longer term, the Retailer may have altered its marketing mix practices and prices as a response to demand for the labels and color-ratings. In this analysis, we assume variation in retail price is exogenous to changes in demand at the Retailer.

For the empirical analysis the data are aggregated at the weekly store level. If a product is not sold at any store in the data set during a particular week, that product is absent from the data and therefore removed from the choice set. The data contain 3,899 product store-week level observations for the pre-labeling period and 3,942 for the post-labeling period.

Table 1 report that salmon, halibut, and sole in that order, are the top three types of seafood by revenue.

6 The data does not contain customer information.

Table 1. Summary Statistics by Fish Species Before Labeling

	Average Weekly Pounds Sold	Average Weekly Sales in $	Average $/Pound	Number of Observations	Percent Weekly Share
Octopus, Squid	4.03	22.72	5.76	98	2.51%
Sole	20.13	255.12	12.54	240	6.16%
Halibut	38.27	546.49	15.17	253	6.49%
Catfish	7.44	59.59	8.19	259	6.64%
Seafood Salad	5.42	43.46	9.46	262	6.72%
Cod	20.56	186.21	9.42	336	8.62%
Salmon	48.29	637.38	13.18	552	14.16%
Shrimp	7.92	106.04	14.33	1103	28.29%

"Before Labeling" refers to the pre-treatment period, which covers twelve weeks starting on February 19, 2006 toMay 17, 2006.
All figures are weekly averages except for the number of observations, which is a total for the entire period.
Pricing is in terms of Dollars per pound.

Summary statistics of the scanner quantity and price data stratified by the pre-labeling and post-labeling periods are reported in Table 2. We find that there is no statistically significant difference between the pre- and post-labeling periods for the natural logarithm of the quantity of units sold or average price per pound.

Table 2. Summary Statistics of the Scanner Data

		Variable	
Time Period	Variable	Mean	Std
Pre Labeling Period	Log (Quantity)	1.99	1.38
	Price	12.27	5.43
	Number of Observations	3899	
Post Labeling Period	Log (Quantity)	2.06	1.37
	Price	12.56	5.21
	Number of Observations	3942	

In Table 3 we see first that the proportion of sales by weight, for each sustainability color-rating, varies by treatment and control stores over the pre-treatment and post-treatment periods. Comparisons between the treatment and control stores in the pre-treatment show that sales by weight in the treatment stores are characterized by the treatment store selling more "green" products, and fewer "yellow" and "red" products relative to the control stores. During the post-treatment period the treatment stores are characterized as selling more “red” products, fewer “yellow” products, and roughly equivalent shares of “green” products relative to the control stores.

There is no statistically significant difference in the number of seafood product choices, the color-label score received, or other product characteristics such as mercury warning or proportion of products that are wild caught or farmed for an average week and store in the pre-treatment and treatment periods. At the average treatment store during the pre-treatment period, customers could choose between 33 different seafood products during an average week versus 32 seafood products for a control store. If these products had been labeled using the FishWise Advisory, 13 of the products would have been green, 9 of them would have been yellow and 11 of them would have been red in a treatment store versus 11 green, 10 yellow, and 11 red in a control store. There is some evidence that the number of red products decreased in both treatment and control stores during the treatment period (11 to 9 products in the treatment stores and 11 to 10 products in the control stores). Otherwise, the number of product choices was similar between pre-treatment and treatment periods.

Table 3. Summary Statistics for Treated and Control Stores

	Treatment Stores			Control Stores		
	pre-treat	treat	Change	pre-treat	treat	Change
Sales in units	21.8	22.1	1.7%	16.8	16.7	-0.4%
	(35.7)	(30.7)		(30.4)	(25.6)	
Price	12.1	12.4	1.9%	12.3	12.6	2.4%
	(5.0)	(5.2)		(5.5)	(5.2)	
Sales by weight (in %)						
Green	68.9	60.8	8.2%	60.1	62	1.9%
	(5.4)	(10.3)		(10.8)	(10.4)	
Yellow	13.3	13.8	0.5%	14.8	15.3	0.4%
	(2.8)	(2.8)		(3.7)	(4.0)	
Red	17.8	25.5	7.7%	25.1	22.8	-2.3%
	(5.3)	(11.9)		(10.6)	(9.3)	
Number of Choices						
Green	13.2	13.8	4.4%	11.6	13.3	14.6%
	(2.1)	(2.3)		(2.5)	(2.9)	
Yellow	9.3	10.3	10.8%	9.5	9.7	2.4%
	(1.6)	(1.8)		(2.4)	(2.6)	
Red	11.2	8.5	-24.5%	11.1	9.9	-10.7%
	(1.9)	(2.1)		(2.5)	(2.9)	
Product Characteristics (in %)						
Low Mercury	56.7	55.2	-1.6%	53.5	50.3	-3.2%
	(4.5)	(3.4)		(6.4)	(5.6)	
On Sale	30.4	28.5	-1.9%	31.6	27.2	-4.5%
	(6.0)	(3.5)		(7.9)	(5.9)	
Wild	52.2	55.2	2.9%	53.9	57.5	3.7%
	(5.1)	(4.5)		(6.4)	(5.6)	
Caught in USA	38.6	41.2	4.6%	33.9	40.9	7.0%
	(2.6)	(3.8)		(6.8)	(6.5)	
Number of Observations	807	779		3092	3163	

"Pre- treat" refers to the pre-treatment period, which covers twelve weeks starting on February 19, 2006 and ending on May 17, 2006.
"Treat" refers to the treatment period, which covers twelve weeks starting on June 10, 2006 and ending on September 3, 2006
All figures are weekly averages except for the number of observations, which is a total for the entire period.
The number of choices refers to the number of product options available to a consumer at the point of purchase.
Label characteristics refers to the percentage of products with each of these characteristics.
Pricing is weighted by sales in pounds.
Number of observations are the number of product-store-week purchases in the data.

There is no statistically significant difference in product characteristics for an average store in the pre-treatment and treatment periods during an average week. At the average treatment store during the pre-treatment period, 56.7% of the products would have been listed on the low-mercury list if the FishWise Advisory had been in place, 30.4% of products were on sale, 52.2% of products were wild (versus farmed) and 36.6% of products were caught in the United States during an average week. These numbers are very similar to those for the control stores, and these figures do not vary much between pre-treatment and treatment periods**.**

3. Model

In our analysis on the impact of information on consumer choice, we define product-specific information provision via labels as an additional or differentiated product attribute. Recognizing that consumer products can be defined as a bundle of perceived product attributes provide the framework to compute consumers' willingness to pay for additional labeling information in a straightforward way. The utilized discrete choice model (e.g. Berry, Levinsohn and Pakes, 1995; McFadden and Train, 2000; Nevo, 2000; Nevo, 2003; Swait et al., 2004) also offers flexibility in incorporating consumer heterogeneity with regard to seafood characteristics such as environmental sustainability, environmental contaminates, and origin.

Consumer choice is modeled with a random utility model framework logit. This modeling approach combined with the unique quasi-experimental setting and resulting data variation for seafood purchases allows us to estimate consumers' valuation for environmental seafood sustainability (McFadden, 1974; Train, 2003).

Starting from a random utility framework where both the product attributes as well as a random term are assumed to enter linearly, the utility from consuming a certain seafood product *j* at time *t* in store *s* can be described as

$$U_{jts} = X_{jts}\beta + L_{jts}\gamma + \xi_{jts} + \varepsilon_{jts} \ , \qquad (1)$$

where the matrix X_{jt} contains the attributes of the seafood product, $L_{jt}=1$ after the labeling implementation for the labeled products (and equal to zero otherwise), the vector β represents the marginal utility placed on each of the X attributes, γ is the marginal utility with respect to the label, and ξ_{jts} are unobserved (to the researcher) determinants of utility but observed by consumers, and ε_{jts} denotes remaining unobserved determinants of utility, where we assume that ε_{jts} is iid type I extreme value distributed.

Assuming that a *consumer i* purchase one unit of product *j* among all the possible products available at a certain time *t* at store *s* that maximizes their indirect utility, then the market share of product *j* during week *t* at store *s* is given by the probability that good *j* is chosen, that is,

$$\int_{(\varepsilon_{ijts}):\, U_{ijts} \geq U_{iht},\, h=0,\ldots N} dF(\varepsilon) \ . \qquad (2)$$

We estimate demand parameters, following Berry (1994) and Berry et al. (1995), by equating the estimated product market shares given by (2) to the observed shares, and solving for the mean utility across all consumers, defined as

$$\delta_{jts} = X_{jts}\beta + L_{jts}\gamma + \xi_{jts} \ . \qquad (3)$$

Where ξ_{jts} denotes unobserved product characteristics, such as changes in shelf display on a subset of retailers.

The logit model is estimated using Berry's (1994) approach to linearize the choice model equation to estimate. Given the predicted market shares or probabilities equal to

$$Prob_j = s_j = \frac{e^{X_j\beta+\alpha p_j+\xi_j}}{\sum_{k=0}^{h} e^{X_k\beta+\alpha p_k+\xi_k}}, \qquad (4)$$

and given that the mean utility of not buying any alternative, that we define as the outside option good j=0, is normalized to 0, then

$$Prob_0 = s_0 = \frac{e^0}{e^0 + \sum_{k=1}^{h} e^{X_k\beta+\alpha p_k+\xi_k}} = \frac{1}{1 + \sum_{k=1}^{h} e^{X_k\beta+\alpha p_k+\xi_k}} \quad (5)\,.$$

Taking the natural logarithm of the probability in (4) and subtracting the log of the probability of not buying (5) yields

$$lnProb_j - lnProb_0 = ln\left(\frac{e^{X_j\beta+\alpha p_j+\xi_j}}{\sum_{k=0}^{h} e^{X_k\beta+\alpha p_k+\xi_k}}\right) - ln\left(\frac{1}{1 + \sum_{k=1}^{h} e^{X_k\beta+\alpha p_k+\xi_k}}\right) \Leftrightarrow$$

$$lnProb_j - lnProb_0 = X_j\beta + \alpha p_j + \xi_j \qquad (6)$$

It follows that estimation of the logit model is obtained by regressing the dependent variable that is the log of each product's observed market share minus the log of the market share of not purchasing on the variables entering the mean utility, such as label scores dummies, product attributes, and retail prices.

Secondary to the primary analysis, we conduct a posterior analysis on the estimated product-store fixed effects, $\widehat{\xi_{Js}}$ from the logit analysis (6). As demonstrated by Pagan (1984), the reliability of drawing such inferences with a simple 2-step OLS based model may be questioned as the OLS estimator of standard errors may be inconsistent in the general "generated regressor" setting. Following Chamberlain (1982), Hoffman (1987), and Nevo (2000) we employ a 2-step approach utilizing generalized least squares (GLS) in the second step to regress the generated product-store fixed effects on several time invariant product characteristics. GLS can be used to perform linear regression when there is a certain degree of correlation between the explanatory variables (independent variables) of the regression. In these cases, ordinary least squares and weighted least squares can be statistically inefficient, or even give misleading inferences. We model fitted product-store fixed effects, $\widehat{\xi_{Js}}$, as:

$$\widehat{\xi_{Js}} = X_j\beta + R_s\gamma + \varepsilon_{js} \qquad (7)$$

where X_j is a $J \times K$ ($K<J$) matrix of several time-invariant product characteristics that are consistent across pre- and post-labeling periods: mercury content, use of selective catch method[7], wild-caught, and U.S. caught, R is a $S\times 1$ vector of store dummy variables, and ε_{js} denotes the remaining unobserved product-store qualities. Results for Barry's logit (6) and posterior GLS (7) analyses are presented in the following section.

4. Results

Parameter estimates for the logit discrete choice demand model of the probability of purchasing a particular seafood product at a store in a week as a function of seafood products attributes and store and week fixed effects, as given by equation (6), are reported in Table 4.

Product attributes consist of (i) time changing attributes such as price and the implementation of posting the labels, and (ii) time invariant product and store determinants of demand captured by a product attributes and store fixed effects. Table 4 reports the parameter estimates for 6 specifications of the model. The dependent variable for all columns is the log of the market share of each product minus the log of the market share of not buying. Seafood receiving a yellow rating is the base category or reference group. In column (1) we present the base specification results, in column (2) we add product by store fixed effects, while in column (3) we add week fixed effects in addition to (2). Columns (4), (5), and (6) add retail price as an additional control to the specification in columns (1), (2), and (3) respectively. Model specifications presented in columns (1) and (3) are estimated with a random effects model, and columns (2), (3), (5), and (6) are estimated with fixed effects. All columns report clustered standard errors with the cluster being defined by product-store.

[7] Using published information from the Monterey Bay Aquarium Seafood Watch, an aggregate group of fishing methods identified as having lower by catch are assembled: midwater trawl, handline, pole, troll, setline, bottom longline, traps, and salmon gear (Seafoodwatch, 2014).

Table 4. Logit Discrete choice Demand Model Estimates

	(1) RE	(2) FE	(3) FE	(4) RE	(5) FE	(6) FE
Constant	-9.842***	-9.618***	-9.337***	-9.120***	-8.522***	-8.222***
	(-0.059)	(-0.021)	(-0.046)	(-0.19)	(-0.469)	(-0.437)
Red	-0.801***			-0.640***		
	(-0.052)			(-0.056)		
Green	0.168*			0.138		
	(-0.095)			(-0.102)		
Red * Treatment Time	-0.017	-0.018	-0.176	-0.013	-0.01	-0.148*
	(-0.061)	(-0.07)	(-0.103)	(-0.053)	(-0.053)	(-0.079)
Yellow*Treatment Time	-0.199***	-0.204***	-0.368***	-0.192***	-0.199***	-0.343***
	(-0.052)	(-0.053)	(-0.075)	(-0.047)	(-0.044)	(-0.077)
Green*Treatment Time	0.101	0.102	-0.06	0.099	0.091	-0.052
	(-0.074)	(-0.076)	(-0.11)	(-0.074)	(-0.076)	(-0.074)
Red* Treatment Time * Treated Store	-0.047	-0.071	-0.071	-0.051	-0.073	-0.073
	(-0.084)	(-0.103)	(-0.104)	(-0.087)	(-0.109)	(-0.108)
Yellow*Treatment Time*Treated Store	-0.277***	-0.324***	-0.327***	-0.268***	-0.309***	-0.314***
	(-0.048)	(-0.061)	(-0.056)	(-0.047)	(-0.055)	(-0.051)
Green*Treatment Time*Treated Store	-0.017	-0.071	-0.076	-0.002	-0.046	-0.051
	(-0.131)	(-0.159)	(-0.157)	(-0.121)	(-0.137)	(-0.135)
Price				-0.060***	-0.088**	-0.088**
				(-0.014)	(-0.037)	(-0.036)
Seafood Product by Store Fixed Effects	NO	YES	YES	NO	YES	YES
Week Fixed Effects	NO	NO	YES	NO	NO	YES
R-squared	0.043	0.007	0.046	0.019	0.050	0.087
Number Observations	7841	7841	7841	7841	7841	7841

Dependent Variable in all specifications is log(sharej)-log(share not buying). Product-Store Clustered SE in parenthesis. Source: Author's Calculations form the scanner data. Interactions of Lower terms omitted to save space. *** p<0.01, ** p<0.05 , * p<0.1 level

Starting with columns (1) and (4), the green and red point estimates, respectively, suggest that consumers prefer green products relative to yellow products and they dislike red products relative to yellow products. When we include product by store fixed effects in columns (2), (3), (4), and (5), we no longer have the color type average marginal utility estimates as in columns (1) and (4). We note that the negative and significant sign on the price parameter in column (4), (5), and (6).

During the experiment period (interpreted as a simple pre-treatment and post-treatment period change), red and green product sales did not significantly change as can be seen in the rows labeled “Red*Treatment Time” and “Green*Treatment time”, respectively, across all specifications. However, sales of yellow rated products dropped significantly as can be seen for all specifications in the row “Yellow*Treatment Time”.

Focusing at the difference-in-difference estimates associated with the triple interactions of the rating color, Treatment Time and Treated Store, we can explore the changes in quantity sold between the control and treatment periods for products in treated stores relative to the contemporaneous changes in the same products in the control stores by label color. Parameter estimates for all specifications result in negative and significant estimates for the yellow label triple-interaction parameters and suggests that the quantity of yellow labeled products sold do decrease due to the labeling implementation. There is no significant impact on red and green products. When we consider additional controls for product – store fixed effects (Column 2), weekly level determinants of quantity (Columns 3), and controls for retail prices (Columns 5 and 6), the results are robust: the only statistically significant triple interaction difference-in-difference parameter estimate is that for yellow-labeled products. Results indicate that yellow-labeled projects experienced a drop in sales by roughly 31% (Columns 5 and 6). We select the specification of Column 6 as the preferred specification due to the inclusion of important covariates and the goodness of fit.

Willingness to pay estimates can be calculated using the estimated coefficients from the RUM framework. For example, the WTP for yellow labels is computed by taking the ratio of the marginal utility of a yellow treatment label of -0.314 by the absolute value of the marginal utility of price that is equal to -0.088. This ratio suggests that consumers find products receiving a yellow label as not preferable and would need to be offered an average discount of 3.56 dollars per pound to purchase offset the impact of labeling a product with a yellow rating. Given the pre-labeling-period average product price of 11 dollars per pound, the estimated price discount needed to get people to continue buying the yellow-labeled alternatives is roughly a third.

While at first, the empirical finding that the implementation of the "yellow" labels has a robust and significant negative impact may appear puzzling. However a programmatic development subsequent to our data collection both supports our results and offers an explanation: the yellow label's definition was changed from "Proceed with Caution" to "Good Alternative." We posit that the label program managers understood that consumers were shifting away from yellow-labeled options, but were not substituting to green-labeled options (as suggested by the yellow and green treatment-time*treatment-store triple interactions). Perhaps, consumers had held higher priors regarding the sustainability of the yellow-labeled options. From a programmatic perspective, while the substitution away from yellow would be a desired outcome if consumers substituted green-labeled options for yellow ones, it may be considered a undesirable outcome if consumer substituted away from seafood products.[8] Therefore therefor softened the yellow label's messaging.

Results of the posterior GLS analysis on the estimated product-store fixed effects from the logit analysis are reported in Table 5. The specification (Equation 7) projects the estimated product-store fixed effects, $\widehat{\xi_{JS}}$, from the specification in Column 6 of Table 4 on several time invariant product characteristics that are consistent across pre- and post-labeling periods: mercury content, use of selective catch method[9], wild-caught, and U.S. caught, and store fixed effects.

Table 5: GLS Fixed Effects Regression

	(1)	(2)	(3)	(4)	(5)
Constant	-1.029***	-1.050***	-1.362***	-1.740***	-0.1198
	(0.078)	(0.151)	(0.258)	(0.525)	(0.259)
Low Mercury	-0.062	-0.052	0.058	-0.120	-0.100
	(0.101)	(0.101)	(0.140)	(0.115)	(0.113)
Selective Gear	0.855***	0.854***	0.533***		0.900***
	(0.121)	(0.121)	(0.180)		(0.127)
Wild Caught	0.324***	0.318***	0.701***		0.451***
	(0.114)	(0.114)	(0.140)		(0.144)
USA	0.178*	0.171*	0.260**	0.187*	
	(0.099)	(0.099)	(0.100)	(0.097)	
Store FE		Yes	Yes	Yes	Yes

[8] Our data is limited to seafood sales data. We therefor cannot explore substitution away from seafood purchases.

[9] Using published information from the Monterey Bay Aquarium Seafood Watch, an aggregate group of fishing methods identified as having lower by catch are assembled: midwater trawl, handline, pole, troll, setline, bottom longline, traps, and salmon gear (Seafoodwatch, 2014).

Species FE			Yes		
Gear FE				Yes	
Country FE					Yes
Observations	864	864	864	864	864
AIC	2940.27	2953.10	2892.29	2887.09	2929.57
BIC	2964.08	3019.76	3001.80	2987.08	3062.90
DF	5	14	23	21	28

Standard errors in parentheses
* (p<0.10), ** (p<0.05), *** (p<0.01)

In the first column of Table 5 we control for time invariant product characteristics. In column (2), we additionally control for store fixed effects. In column (3), we control for species group using species fixed effects as well. In column (4), we substitute harvest method fixed effects for species fixed effects. In column (5), we substitute country of origin fixed effects for species fixed effects. In column (3), the species groups we control for are coastal pelagic species, highly migratory species, snapper, bass, freshwater species, groundfish, halibut, salmon, shellfish and other crustaceans, a multi-species sea-food salad category, and a residual "other" category for all other seafood counter products. In column (4), the harvest methods we control for are purse seine, gillnet, traps, handline/pole/troll, longline, trawl, farmed, and a category for mussels and wild salmon. In column (5), we control for 10 countries and an "exempt from disclosure" category. Due to collinearity with the product specific fixed effects, the parameter estimates for some of the primary time invariant product characteristics vary by model. Specifically, in column (3) the parameter estimate for "wild caught" more than doubles from columns (1) and (2) due to the inclusion of species fixed effects.

Empirical results from the posterior GLS estimation suggest that, on average, consumers prefer the use of selective harvest methods, wild caught seafood over farmed, and seafood originating in the U.S. over internationally. The coefficient on low-mercury products is not significantly different from zero for all specifications. These results are robust across all specifications apart from column 4 for which we expect that the gear type fixed effect is collinear with the both the selective harvest method and wild caught binary variables.

5. Conclusions

This paper provides the first detailed empirical RUM analysis concerning the impact of point-of-sale labeling changes on purchase patterns between seafood products using retail data. We empirically investigate consumer responses to one such seafood "traffic light" eco-label system implemented at a coastal California supermarket chain.

Through marketplace signals, eco-labels are purported to economically reward successful stewardship. In theory, eco-labels provide consumers easy-to-use relative-rating information, allowing the differentiation of products on the basis of a complex set of attributes that are beyond the general knowledge of consumers. From the commercial fishery operation's perspective, the utility of eco-labels is their ability to allow the differentiation of products along environmental and sustainable attributes that may allow the passing of costs associated with best-practices onto consumers who value sustainability.

Our findings inform fishery participants, managers, retailers, and consumers of the expected price premium and demand shift impacts of seafood labels, and have implications for the current debate regarding sustainability labels as a tool in coastal and marine resource conservation and management. Through the testing of the presence and scale of an effect, the research provides information on the estimated benefits of the eco-label tool, and highlights the concomitant risk of placing too much confidence in the efficacy of eco-labels to alter consumer behavior and demand for seafood.

Based on modeling actual seafood purchase records, we find consumers reveal preferences that are generally consistent with placing a positive value on sustainability. Evidence from the posterior GLS analysis suggests positive and significant consumer willingness-to-pay for seafood characterized as being caught by selective gear, wild-caught, and of US origin.

Evidence on consumer willingness-to-pay for sustainability, as measured by the implementation of the eco-label, is more nuanced. In this data, when consumers are provided additional information on the sustainability rating via a "traffic-light" based product rating system, the consumers do not switch to the better alternatives within those rankings. Neither do we see consumers significantly switching away from the worst ranked products. However, we do find that when provided with the eco-label consumers would need to receive a discount of roughly 1/3 to purchase products receiving a Yellow, `proceed with caution,' score.

One possible explanation of the empirical finding of a negative and significant treatment parameter estimate for yellow-rated seafood is that the aggregate pre-label perception of consumers is that the seafood is relatively more sustainable than it is as measured by the third-party rating agency; and in the light of the provision of sustainability information via the sustainability-ratings labels consumers selected to reduce purchases of yellow-rated seafood that they had previously believed to be of a high level of sustainability. Although the interpretation of the result should be tested through further study, the possible interpretation illustrates a possible behavioral consumer response in which consumers value attributes associated with sustainability. And while consumers may have already acquired information on seafood sustainability regarding “red” and “green” products through other information channels, the eco-label program provides new information to consumers with which they, in aggregate, alter their actual purchases. The differentiated nature of the consumers that shop at the high-end retailer should be considered in the context of hypothesis.

Consistent with the above explanation, we note a programmatic development subsequent to our data collection where the yellow label’s definition was changed from “Proceed with Caution” to “Good Alternative.” From a programmatic perspective, while the substitution away from yellow would be a desired outcome if consumers substituted green-labeled options for yellow ones, it may be considered a undesirable outcome if consumer substituted away from seafood products. In light of this, program administrators may desired to fine-tune their messaging through the softening of the yellow label’s definition.

These findings provide mixed information to the fishing industry and dependent communities regarding the willingness of individual consumers to pay a price premium for sustainably caught seafood and the impact of the implantation of the eco-label program. Within our sample of data, we find evidence that in general consumers value sustainability, yet we reject that the labels have a significant positive label efficacy. Thus our findings may, taken as given, not encourage businesses to adopt additional programs that disseminate product sustainability attributes. Results suggest that the sustainability labels under study had the short-run affect of steering consumers away from seafood (yellow ratings) that did not meet the standards of a "best choice" (green), and consumers did not substitute to the "best choice" alternative due to a mix of price and tastes. It is possible that consumer behavior may change over the long-run as new product markets mature with consumer demand.

Researchers and practitioners have considerable work ahead of them in order to a) better understand the impact of confounding factors on eco-labeling, environmental contaminates, and origin labeling efficacy, b) understand consumer perceptions and understanding of product labeling programs, and c) design more effective, uniform and standardized product labeling programs. In all three areas, it is crucial that conservation and management professionals understand the expected strengths and weaknesses of product labeling programs. Poorly designed labeling programs may have the potential to shift consumer purchase patterns with negative consequences, reducing consumer and producer welfare and reducing overall fishery sustainability.

References

Alfnes, F., A.G. Guttormensen, G. Steine, and K. Kolstad. "Consumer's Willingness to Pay for the Color of Salmon: A Choice Experiment with Real Economic Incentives." *American Journal of Agricultural Economics*, 88 (2006): 1050–1061.

Asche, F. and J. Guillen. "The Importance of Fishing Method, Gear and Origin: The Spanish Hake Market." *Marine Policy*, 36 (2012): 365–369.

Batte, M. T., N. H. Hooker, T. C. Haab, and J. Beaverson. "Putting their money where their mouths are: Consumer willingness to pay for multi-ingredient, processed organic food products." *Food Policy*, 32 (2007): 145–149.

Berry, S. "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 25-2 (1994): 242-262.

Berry, S., J. Levinsohn and A. Pakes. "Automobile Prices in Market Equilibrium." *Econometrica,* 63 (1995): 841-890.

Blomquist, J., V. Bartolino, and S. Waldo. "Price Premiums for Providing Eco-Labelled Seafood: Evidence from MSC-Certified Cod in Sweden." *Journal of Agricultural Economics*, February (2015), n/a - n/a. doi:10.1111/1477-9552.12106.

Chamberlain, G. ,(1982): "Multi Variate Regression Models for Panel Data," *Journal of Econometrics*, 18 (1982), 5-46.

Carroll, M., J. Anderson, and J. Martinez-Garmendia. "Pricing U.S. North Atlantic Bluefin Tuna and Implications for Management." *Agribusiness*, 17 (2001):243–254.

Chiang, F., J. Sun, M. Owens, and D. Squires. "Will American Consumers Pay More for Eco-Friendly Canned Tuna?" Working Paper. January, 2015.

Costello, C., S.D. Gaines, and J. Lynham. "Can Catch Shares Prevent Fisheries Collapse?" *Science*, 321 (2008): 1678–1681.

Delgado, C. L., N. Wada, M. W. Rosegrant, S. Meijer, and M. Ahmed. "Outlook for Fish to 2020: Meeting Global Demand." Technical report, International Food Policy Research Institute, October 2003.

Hallstein, E. and S. B. Villas-Boas. "Can Household Consumers Save the Wild Fish? Lessons from a sustainable seafood advisory." *Journal of Environmental Economics and Management,* 66-1 (2013): 52-71.

Hammarlund, C. The Big, the Bad, and the Average: Hedonic Prices and Inverse Demand for Baltic Cod. *Marine Resource Economics*, volume 30 (2015), number 2.

Hensher, D. A., and M. Bradley, "Using Stated Response Data to Enrich Revealed Preference Discrete choice Models," *Marketing Letters*, 4-2 (1993):139-152.

Hilger, J. G. Rafert, and S. B. Villas-Boas. "Expert Opinion and the Demand for Experience Goods: An Experimental Approach in the Retail Wine Market." *Review of Economics and Statistics*, 93-4 (2011): 1289-1296.

Hoffman, D. "Two-Step Generalized Least Squares Estimators in Multi-Equation Generated Regressor Models." *The Review of Economics and Statistics,* 69-2 (1987): 336-346.

Ippolito, P. and A. D. Mathios. "Information and Advertising: The Case of Fat Consumption in the United States." *The American Economic Review*, 85-2 (1995): 91–95.

Jaffry, S., H. Pickering, Y. Ghulam, D. Whitmarsh, and P. Wattage. "Consumer Choices for Quality and Sustainability Labeled Seafood Products in the U.K." *Food Policy* 29 (2004): 215–28.

Jin, G. and P. Leslie. "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards." *The Quarterly Journal of Economics*, (2003): 409–451.

Johnston, R. J., and C.A. Roheim. "A Battle of Taste and Environmental Convictions for Eco-labeled Seafood: A Contingent Ranking Experiment." *Journal of Agricultural and Resource Economics*, 31 (2006): 283–300.

Johnston, R. J., C.R. Wessells, H. Donath, and F. Asche. "Measuring Consumer Preferences for Eco-labeled Seafood: An International Comparisons." *Journal of Agricultural and Resource Economics*, 26 (2001): 20–39.

Kiesel, K. and S. B. Villas-Boas. "Got Organic Milk? Consumer Valuations of Milk Labels after the Implementation of the USDA Organic Seal." *Journal of Agricultural and Food Industrial Organization*, 5-4 (2007).

Mathios, A. D. “The impact of mandatory disclosure laws on product choices: An analysis of the salad dressing market.” *Journal of Law and Economics,* 42-2 (2000): 651-677.

McConnell, K. and I. Strand. “Hedonic Prices for Fish: Tuna Prices in Hawaii.” *American Journal of Agricultural Economics*, 82 (2000):133–44.

McFadden, D. “The Measurement of Urban Travel Demand.” *Journal of Public Economics* 3(1974): 303-28.

McFadden, D. and K.E Train. “Mixed MNL Models for Discrete Response.” *Journal of Applied Econometrics* 15(2000): 447-470.

Nevo, A. “A Practitioner’s Guide to Estimation of Random-Coefficients Logit Models of Demand*.” Journal of Economics & Management Strategy* 9 (2000): 513-548.

Nevo, A. “New Products, Quality Changes and Welfare Measures Computed from Estimated Demand Systems.” *Review of Economics and Statistics* 85(2003): 266-275.

Nimon, W. and J. Beghin. “Are Eco-Labels Valuable? Evidence from the Apparel Industry.” *American Journal of Agricultural Economics*, 81-4(1999): 801–11

Pagan, A. “Econometric Issues in the Analysis of Regressions with Generated Regressors.” *International Economic Review*, 25-1(1984): 221-247.

Roheim, C. A. “Early Indications of Market Impacts from the Marine Stewardship Council’s Eco-labeling of Seafood.” *Marine Resource Economics*, 18 (2003): 95–104.

Roheim, C. A., F. Asche, and J.L. Santos. “The Elusive Price Premium for Eco-labeled Products: Evidence from Seafood in the UK Market.” *Journal of Agricultural Economics*, 62-3(2011): 655–68.

Roheim, C. A., L. Gardiner, and F. Asche. "Value of Brands and Other Attributes: Hedonic Analysis of Retail Frozen Fish in the UK." *Marine Resource Economics*, 22-3(2007): 239–53.

Seafoodwatch. "Fishing and Farming Methods." Monterey Bay Aquarium. May 6th, 2014 http://www.seafoodwatch.org/cr/cr_seafoodwatch/sfw_gear.aspx

Shimshack, J. , M. Ward and T. Beatty. "Mercury Advisories: Information, Education, and Fish Consumption." *Journal of Environmental Economics and Management* 53-2(2007): 158-79.

Smith, M. D., C.A. Roheim, L.B. Crowder, B.S. Halpern, M. Turnipseed, J.L. Ander- son, F. Asche, L. Bourillón, A.G. Guttormsen, A. Khan, L.A. Liguori, A. McNevin, M.O. O'Connor, D. Squires, P. Tyedmers, C. Brownstein, K. Carden, D.H. Klinger, R. Sagarin, and K.A. Selkoe. "Sustainability and Global Seafood." *Science*, 327(2010): 784–86.

Sogn-Grundvåg, G., T. Larsen, and J. Young. "The Value of Line-Caught and Other Attributes: An Exploration of Price Premiums for Chilled Fish in UK Supermarkets." Marine Policy 38 (2013): 41-44. doi:10.1016/j.marpol.2012.05.017.

Train, Kenneth. Discrete Choice Methods with Simulation. Cambridge University Press, 2003.