Using Scanner Data To Answer Food Policy Questions

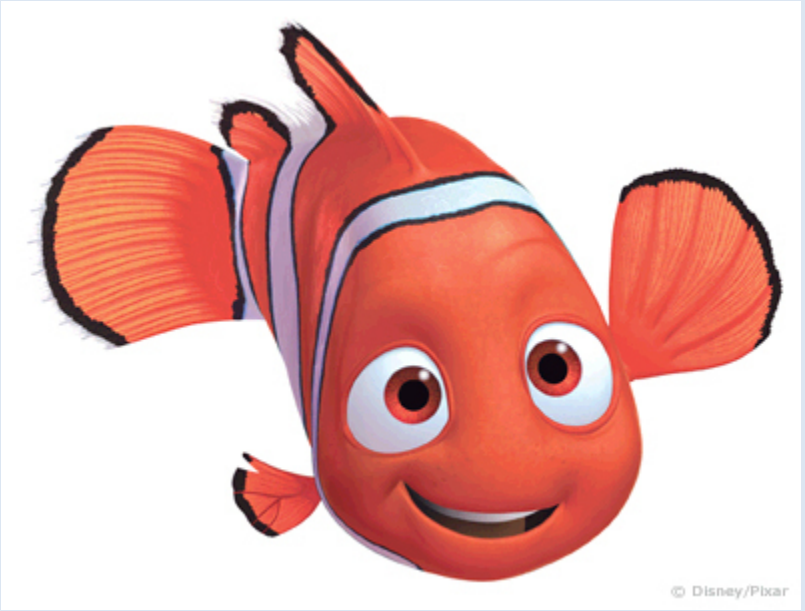
Conference

Wednesday, June 1 - Thursday, June 2, 2011

Economic Research Service
1800 M Street, NW
Waugh Auditorium
Washington, DC

Can Household Consumes Save the Wild Fish?

The FishWise Environmental Labeling Program and Retail Grocery Seafood Sales



Eric Hallstein and Sofia B. Villas-Boas
University of California, Berkeley

Using Scanner Data to Answer Food Policy Questions Conference
June 1-2, 2011

Research Question

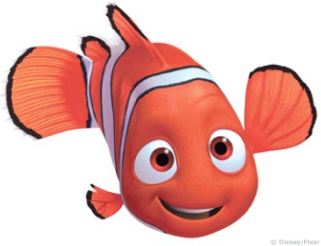
Research question: What impact does the FishWise environmental labeling program have on seafood sales in a local retail chain?

Rationale for environmental labels

- Inform consumers about unobservable or hard-to-verify product attributes

- Helps consumers select products that better match their values, resulting in greater direct utility

- Provides firms with market-based incentive to produce products with higher levels of socially-desirable attributes



FishWise project

The problem

Seafood stocks declining globally

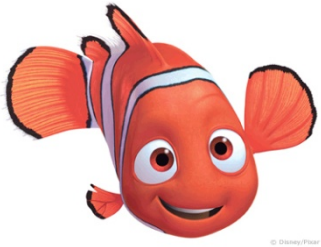
Aquaculture potential limited by environmental issues

Health concerns (and benefits)

The Fishwise solution

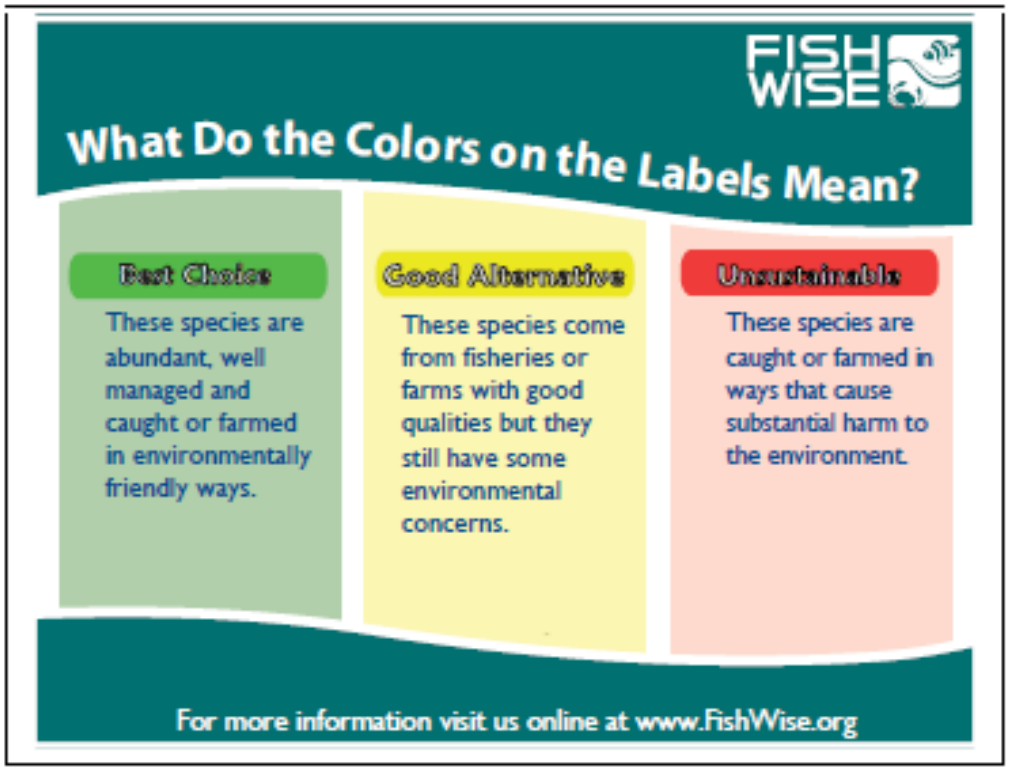
Label seafood at point of purchase

- **<u>Best choice</u> – sustainable, little impact on ecosystem**
- **<u>Proceed with caution</u> – populations healthy, other problems such as poor fishery management**
- **<u>Worst choice</u> – populations over-fished; additional problems such as habitat destruction**



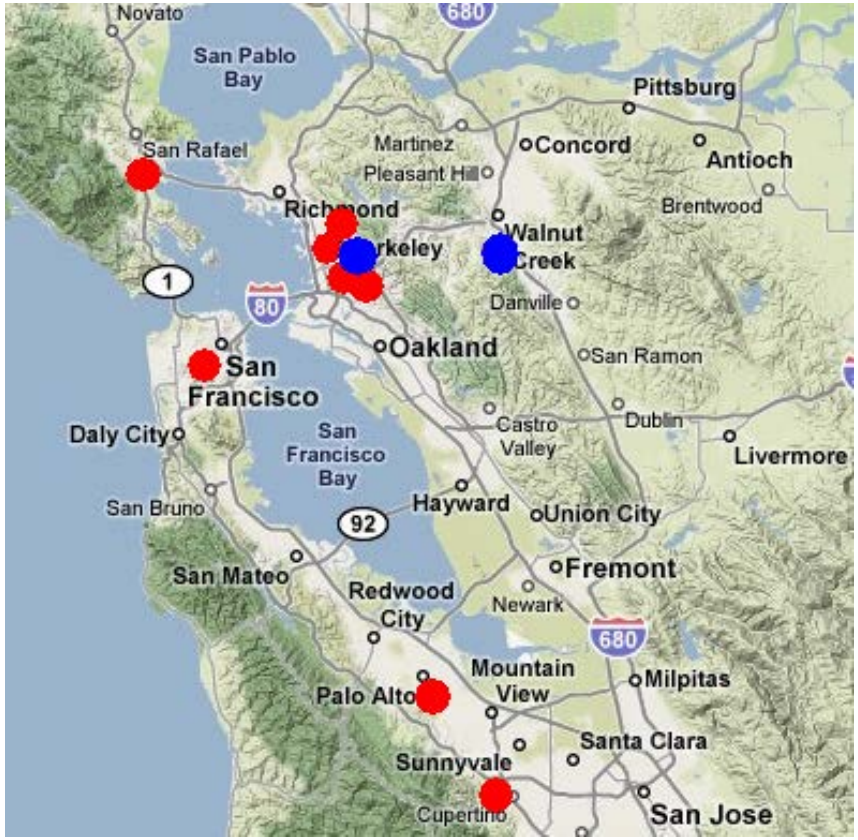
Fishwise Labels

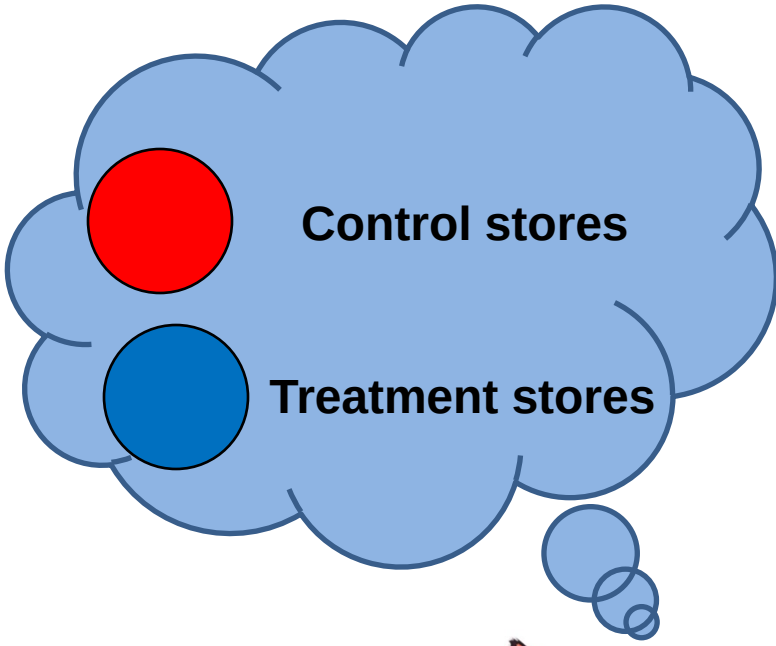
Figure 1.9: **FishWise Interpretive Poster**



This is an example of the FishWise interpretive poster that is supplied by FishWise for use by retailers. In a typical store, a poster of this information is displayed immediately adjacent to or behind the fresh seafood counter.

Analysis exploits a natural experiment

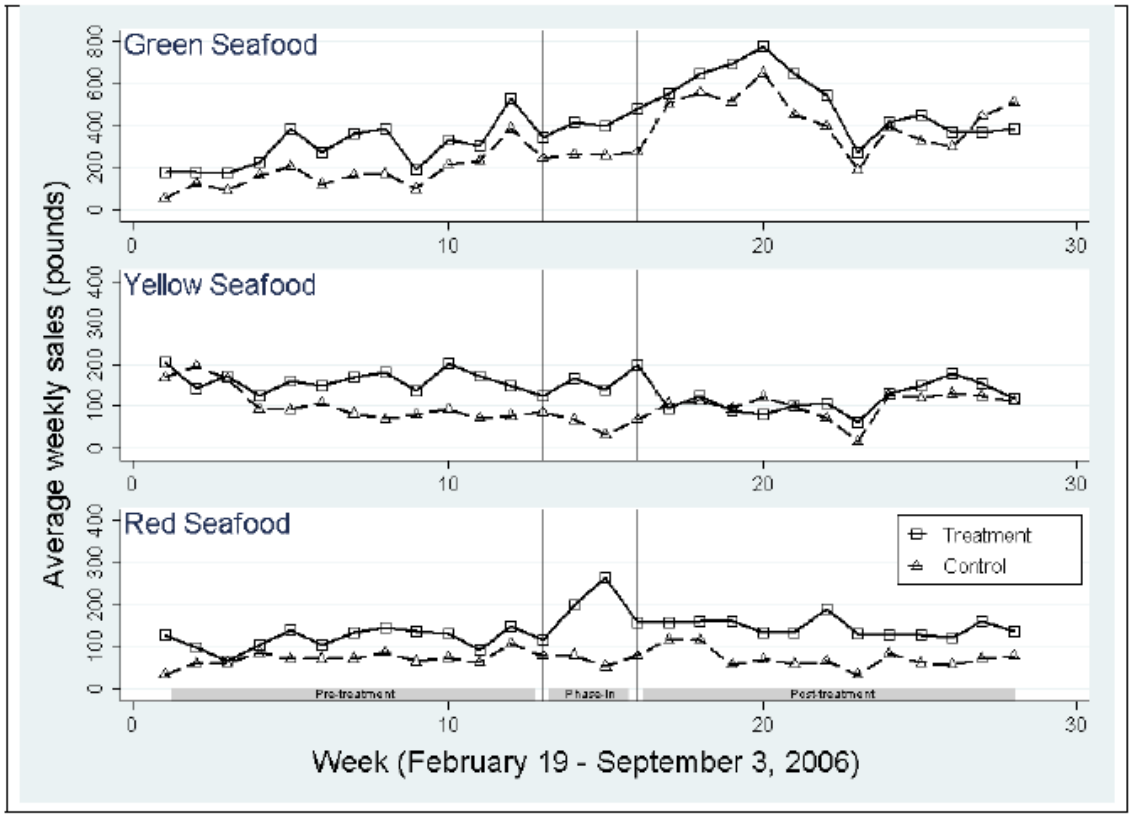


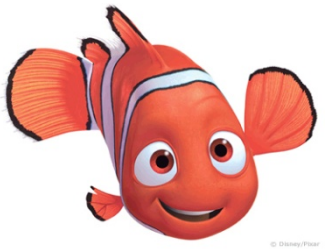


When comparing T and C stores, in pre and also in post period:
Prices, promotional activity, and product choice set, did not respond to labeling
As we find no statistically significant difference in the price, promotions, number of seafood product choices for an average week n the pre-treatment and treatment periods
All those marketing variables were common to T and C stores always, in the pre period as well as in the post treatment period.

Comparison T and C Stores – pre period

- Weekly sales' trends/levels by color





- T versus C stores' characteristics not signif diff

Data

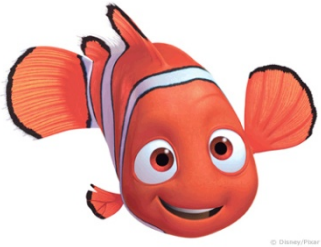
- 5 years (million + records) of weekly scanner data by SKU from retail chain
  - Revenue
  - Pounds sold
  - Retail price
  - Sale price

- Historical advertisements schedule from retailer

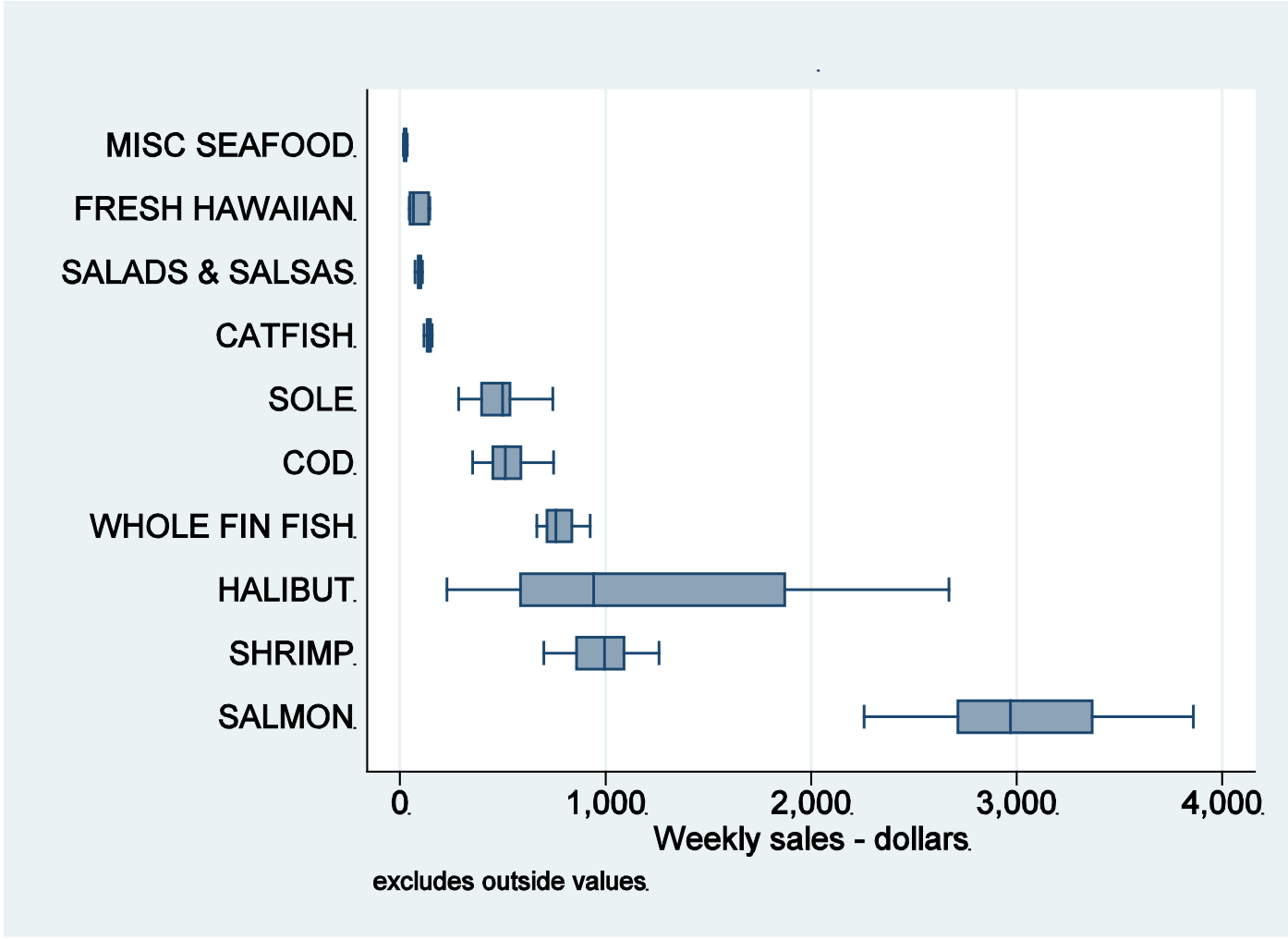
- FishWise color label guide
  - Color code
  - Catch method (e.g. bottom trawl)
  - Production method (e.g. wild)
  - Country of origin

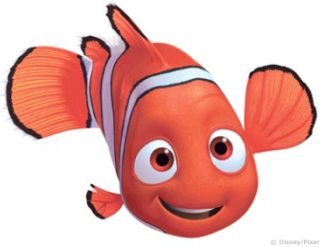
- FishWise mercury and PCB-safe list

- Data used before and after phase-in of Fishwise Label Program in Treated and in control stores



Average weekly sales by seafood type





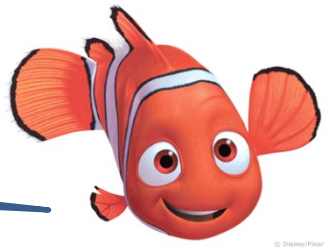
An observation are ounces sold of a certain seafood product at a certain store during a certain week.

Mean Changes

Table 2.1: **Store Characteristics**

	Treatment stores			Control stores		
	pre-treat	**treat**	**delta**	**pre-treat**	**treat**	**delta**
Weekly store sales ($)	404,707	397,769	-1.7%	338,822	326,430	-3.8%
	(125,318)	(112,912)		(120,515)	(109,707)	
Weekly meat sales ($)	30,653	29,525	-3.8%	25,954	25,408	-2.1%
	(7,977)	(6,664)		(11,542)	(10,889)	
Weekly seafood sales ($)	16,259	14,901	-9.1%	11,898	11,264	-5.6%
	(2,051)	(1,572)		(4,715)	(4,607)	
Weekly seafood sales (lbs)	732	718	-1.9%	541	551	1.9%
	(210)	(177)		(229)	(213)	
Weekly store traffic (shoppers)	14,449	14,649	1.4%	11,126	11,384	2.3%
	(3,159)	(3,349)		(2,586)	(2,532)	
Median annual HH income ($)	45,015	45,015	0.0%	78,842	78,842	0.0%
	(6,540)	(6,540)		(26,434)	(26,434)	
Number of stores	2	2		8	8	

From pure differences in means we see that seafood dollar sales drop in treatment stores and in controls too, but looks like they drop more in treatment stores



Difference-in-difference analysis

Difference-in-difference analysis exploits fact that the retailer piloted FishWise in three stores

$$\hat{\gamma}_{ist} = (\overline{Q}_{after,treat} - \overline{Q}_{after,control}) - (\overline{Q}_{before,treat} - \overline{Q}_{before,control})$$

Econometric specification

$$Q_{ist} = \alpha + \delta t_{it} + \beta T_{is} + \gamma (T_{is} * t_{it}) + \varepsilon_{it}$$

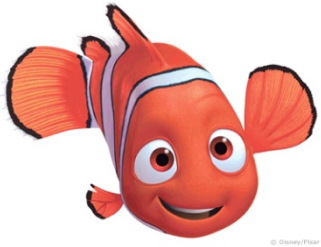
Q_{ist} : pounds of seafood i sold in store s during month t

t_{it} : dummy variable equal to 1 during pilot, 0 otherwise

T_{ist} : vector of dummy variables $(R_{ist}, Y_{ist}, G_{ist})$
where each element is equal to 1 if labeled red, yellow or green, 0 otherwise

And, although not in notation, the equation includes product –store fixed effects.

where γ is interpreted as the treatment effect

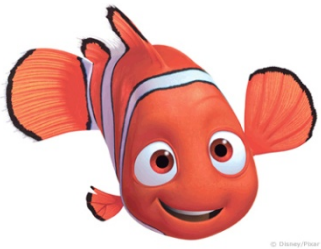


ATE- Diff-in-difference analysis

Dependent variable : Ln (pounds) sold of seafood in store s during period t

	(1)	(2)	(3)	(4)	(5)
Treatment period dummy	-0.0201	0.1438	-0.0488	-0.0741	-0.1386
	(0.0515)	(0.1503)	(0.1678)	(0.1754)	(0.1898)
Treatment effect	-0.1298	-0.1187	-0.1321	-0.1451	-0.1530*
	(0.0818)	(0.0818)	(0.0820)	(0.0825)	(0.0759)
Ln(price)			-0.0228	-0.0197	-0.0114
			(0.8056)	(0.8031)	(0.8013)
Discount			0.3132	0.3265	0.3438
			(1.0426)	(1.0453)	(1.0419)
Promotion indicator			0.6297***	0.6285***	0.6247***
			(0.0716)	(0.0728)	(0.0726)
Constant	4.8230***	5.1287***	5.0106**	4.7395*	5.2898**
	(0.0217)	(0.0731)	(2.0076)	(2.1042)	(2.0524)
week by seafood type		yes	yes	yes	yes
fish share of meat control				yes	yes
mercury list share and color share					yes
Observations	7841	7841	7841	7841	7841
r2	0.0013	0.113	0.1574	0.1578	0.1584
ll	-9515.9013	-9050.8715	-8849.6386	-8847.3896	-8844.9559

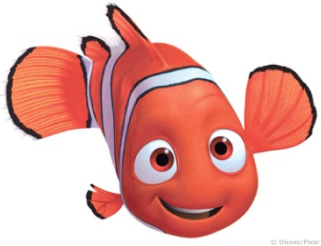
All regressions have store product fixed effects. Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01



Econometric specification (effects by color)

$$
\begin{aligned}
Ln(Q_{ist}) \quad = \quad & \alpha + \beta_1 t_{it} + \beta_2 T_{is} + \upsilon X_{ist} + \\
& \sigma\left(Store_{is} * Promo_{ist}\right) + \gamma\left(Color_{its} * T_{is}\right) + \\
& \rho\left(Color_{ist} * t_{it}\right) + \delta\left(Color_{ist} * T_{is} * t_{it}\right) + \epsilon_{its}
\end{aligned}
$$

Q_{ist} : ounces sold of seafood i in store s in period t
t_{it} : time dummy
T_{it} : treatment dummy
X_{ist} : matrix containing ln(price), discount and fixed effects for SKU, week, catch method, production method and country of origin
$Color_{ist}$: color of SKU i in store s in period t



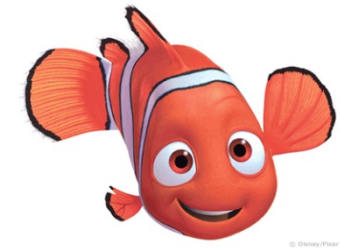
Average treatment effect by label color

Dependent variable : Ln (pounds) sold of seafood in store s during period t

	(1)	(2)	(3)	(4)	(5)
Green: treatment effect	-0.0705	-0.0572	-0.067	-0.0809	-0.0883
	(0.1034)	(0.1057)	(0.1286)	(0.1308)	(0.1229)
Yellow: treatment effect	-0.3233***	-0.3113***	-0.3286***	-0.3416***	-0.3490***
	(0.0606)	(0.0457)	(0.0407)	(0.0340)	(0.0330)
Red: treatment effect	-0.0705	-0.0653	-0.0774	-0.0904	-0.0993
	(0.1593)	(0.1521)	(0.1042)	(0.1015)	(0.0975)
Ln(price)			0.0534	0.0585	0.0642
			(0.8284)	(0.8253)	(0.8233)
Constant	4.8273***	5.1360***	4.8281**	4.5592*	5.0993**
	(0.0209)	(0.0716)	(2.0630)	(2.1596)	(2.1034)
Observations	7841	7841	7841	7841	7841
r2	0.0074	0.1162	0.1599	0.1604	0.1609
ll	-9491.6416	-9036.567	-8837.951	-8835.7041	-8833.344

Standard errors in parentheses

* p<0.10, ** p<0.05, *** p<0.01



Recap

•Overall decrease in total sales but the three-color label system has mixed results

• People did note label content, as color mattered

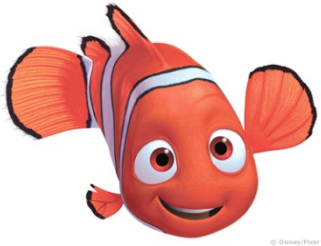
• Labels do decrease sales of **yellow**-labeled products in treatment stores relative to control stores

• Labels do not make a statistically significant difference in **green** and or **red**-labeled product sales in treatment stores relative to control stores

• That consumers do not decrease for red labels presents obstacle to change. Profit driven retailers are likely going to continue stocking red labeled seafood as long as consumers buy it…

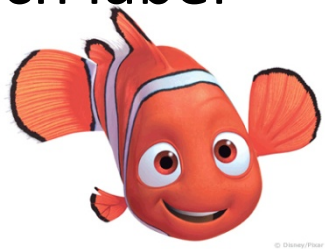
Next steps

- Use exogenous change in labels to estimate demand and supply model
  - Estimate elasticities before and after labels
- Simulate price and welfare changes due to policy simulations



Structural Analysis - demand

- Discrete choice model, where consumers max U among product choice set
- In the choice set some observable product attributes change in the middle of the sample by the addition of the color labels.
- Label change is orthogonal to any other demand marketing variable
- Use scanner data to estimate demand.
- Obtain demand elasticities before and after label change
- Obtain estimates of willingness to pay (WTP) for each label color



Demand

Discrete choice model for differentiated products

Indirect latent utility from consumer *i* choosing product *j* (brand-store) at time *t*

$U_{ijt} = D_t + d_j + x_{jt}\,\beta_i - \alpha_i\, p_{jt} + \xi_{jt} + \varepsilon_{ijt}$, d_j product dummy variables ,

D_t seasonal dummies

x_{jt} observed product characteristics

ε_{ijt} distribution of consumer preferences about unobserved product characteristics (will be integrated out)

What is in ξ_{jt} ? Changes in unobserved consumer preferences, other unobserved market specific conditions (e.g. unobserved promotions, previous sales, changes in shelf display)

Specifying consumer heterogeneity in that $\alpha_i = \alpha + \Phi_v \nu_i$ where ν_i are unobserved consumer characteristics

Note: if $\alpha = \alpha_i$ (and ε extreme value) $\Rightarrow$ Logit.

Demand

Consumer purchases one unit of the good that gives the highest utility conditional on characteristics, prices and outside good.

Aggregate market share of product *j*

$s_{jt} = \int(\{(D_i, \nu_i, \varepsilon_i) | U_{ijt} \geq U_{iht} \;\; h = 0, \ldots N\})\, dF(\varepsilon)\, dF(\nu)\, dF(D)$

Estimate demand parameters that produce predicted aggregate market shares close to observed ones – non linear estimation

- linearise by log difference of share of choice j relative to share of choice of no buying (good zero denoted by s0) assuming ε_{ijt} is distributed *iid* extreme value

$\text{Log}(s_{jt}) - \log(s0t) = D_t + d_j + x_{jt}\,\beta_i - \alpha_i\, p_{jt} + \xi_{jt}$ *(a)*

Problem of estimation of (a) - prices are correlated with ξ_{jt}

Solution: Use instruments for prices – we use wholesale prices

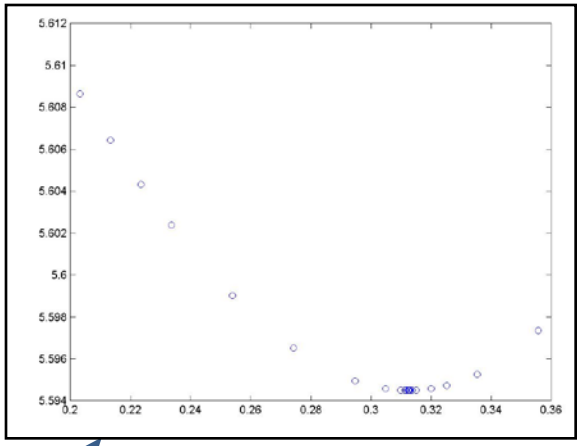
Note: explain how we get no buying every t observation s0t

Structural Analysis - demand

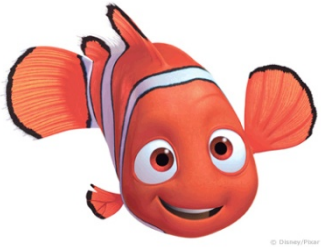
Table 4: **Demand Estimates**

Dependent variable: ln (pounds) sold of seafood i in store s during period t.

	Logit IV		RC GMM	
Variable	Estimate	Std Error	Estimate	Std Error
Price	-1.717	(0.023)	-1.830	(0.399)
Std Dev Price			0.313	(0.047)
Green	0.445	(0.058)	0.476	(0.117)
Red	-0.298	(0.057)	-0.315	(0.150)
Green label	-0.066	(0.050)	-0.076	(0.117)
Yellow label	-0.432	(0.042)	-0.428	(0.133)
Red label	0.081	(0.049)	0.070	(0.126)
Store 1	-0.632	(0.058)	-0.624	(0.208)
Store 2	-1.233	(0.064)	-1.230	(0.192)
Store 3	-0.446	(0.049)	-0.441	(0.237)
Store 4	-1.414	(0.051)	-1.412	(0.213)
Store 5	-0.602	(0.057)	-0.593	(0.208)
Store 6	-0.424	(0.063)	-0.417	(0.202)
Store 7	0.128	(0.055)	0.140	(0.220)
Store 8	-0.728	(0.051)	-0.736	(0.234)
Store 9	-0.201	(0.060)	-0.204	(0.322)
First Stage R2	0.996		0.996	
First Stage F Stat p value	0.000		0.000	
Number Observations	2634		2634	



gmm



Structural Analysis - WTP

1. Estimate consumers WTP for different label colors

	Logit IV		RC GMM	
Variable	Estimate	Std Error	Estimate	Std Error
Price	-1.717	(0.023)	-1.830	(0.399)
Std Dev Price			0.313	(0.047)
Green	0.445	(0.058)	0.476	(0.117)
Red	-0.298	(0.057)	-0.315	(0.150)
TreatGreen	-0.066	(0.050)	-0.076	(0.117)
TreatYellow	-0.432	(0.042)	-0.428	(0.133)
TreatRed	0.081	(0.049)	0.070	(0.126)

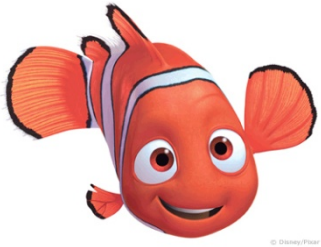
WTP = mg U label /mg U price

Only significant marginal U is yellow

WTP yellow is – 0.42/1.83 = - 23 cents

People do not like Yellow Labels !

People would need a 23 cents discount per ounce to consume yellow labeled products and are not willing to pay or be paid anything to buy red or green.

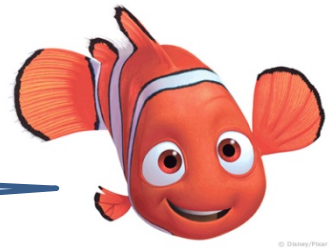


Structural Analysis - Demand

2. Estimate demand and elasticities before and after labeling using two years of data available.

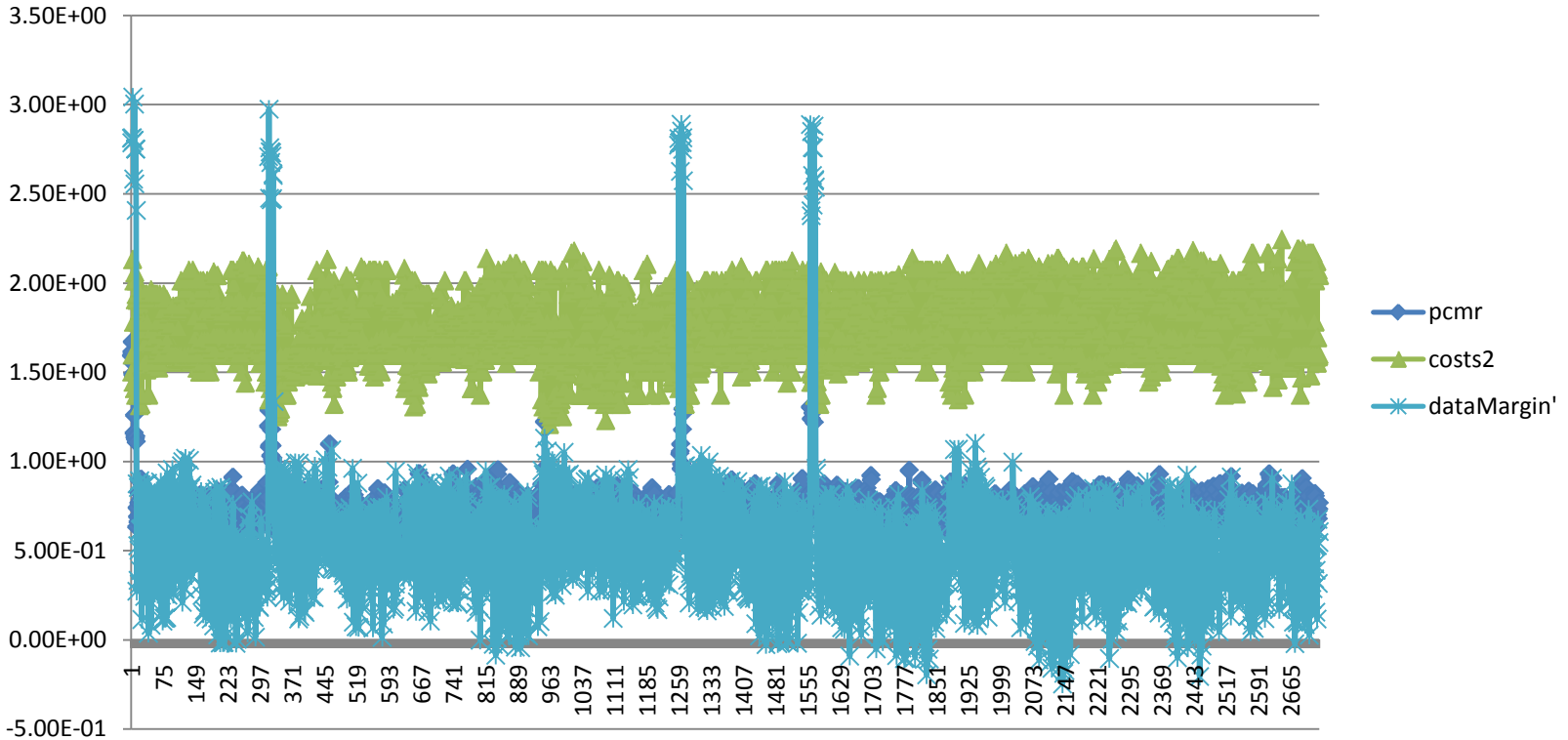
No Label				
		Green	Yellow	Red
	Green	-2.2697	0.0824	0.0881
	Yellow	0.0217	-1.9854	0.014
	Red	0.0361	0.0216	-2.057
Label				
		Green	Yellow	Red
	Green	-1.8083	0.0172	0.0119
	Yellow	0.0296	-1.8899	0.0002
	Red	0.0049	0.0001	-1.5809

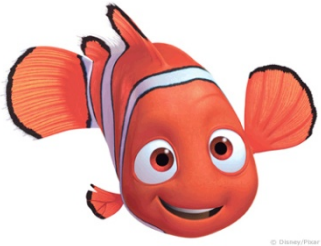
Products become less elastic.
Cross price elasticities are smaller



Structural Analysis – in progress

1. Given a supply model of local multiproduct monopolistic retailer, estimate price cost margins by label color (pcmr)
2. Contrast estimated margins with p-wholesale price data
3. Recover costs = p-estimated margins (green line)
4. Break costs up by color (to do)





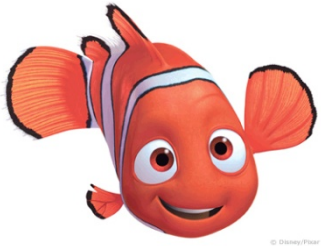
Structural Analysis – in progress

Simulate policy changes:

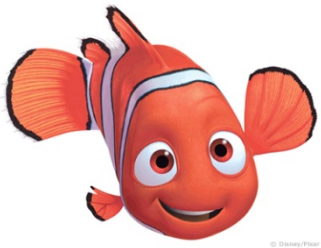
- What if regulation banning red in the markets?– what would happen to prices if consumers have no red fish in the choice set? To consumer welfare? To profits of retailer?
- What if a red label fish became a yellow by red fisheries becoming yellow cost fisheries? Would that be profitable to the fishery given new prices post simulation?

Take Away

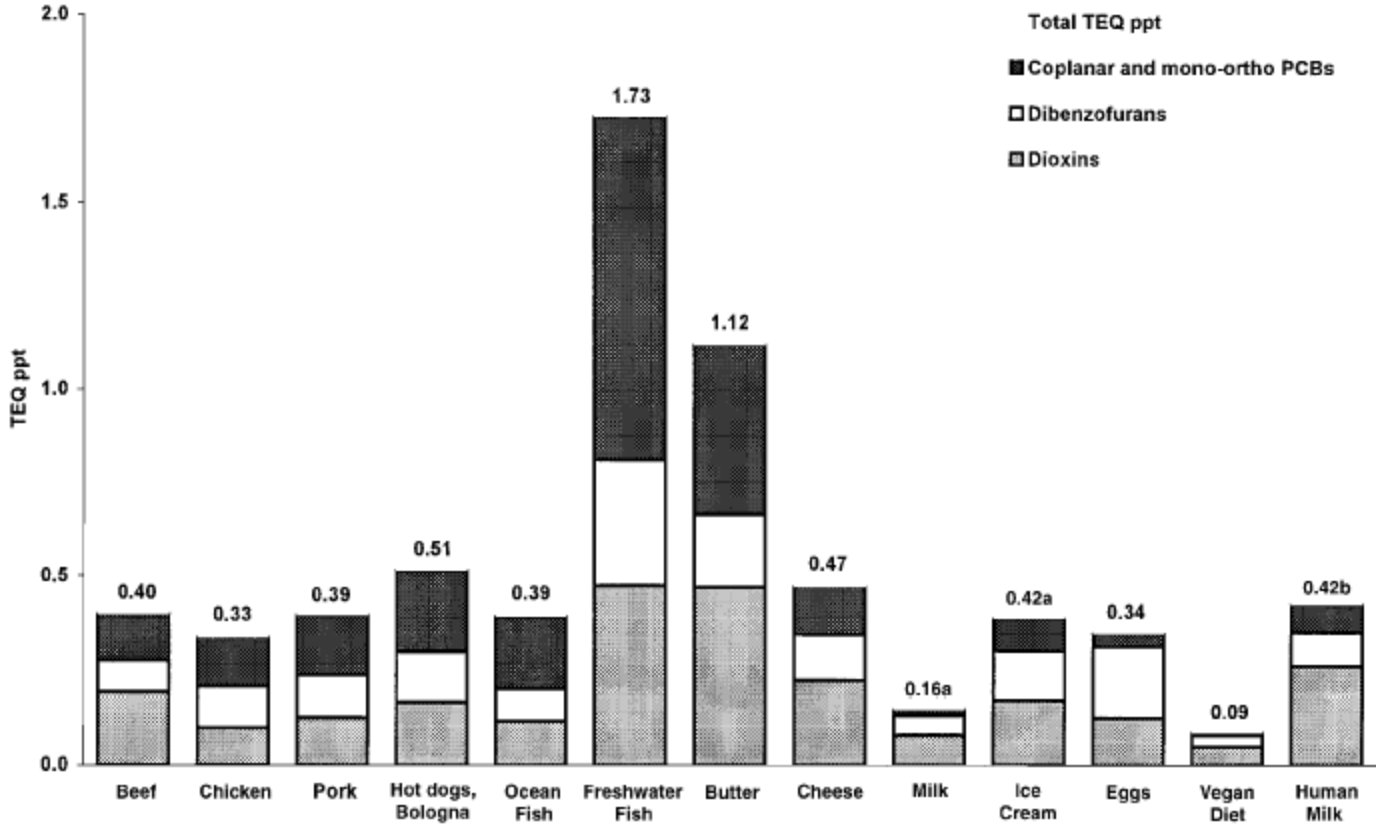
- Overall decrease in total sales but the three-color label system has mixed results
- People did note label content, as color mattered
- Labels do decrease sales of **yellow**-labeled products in treatment stores relative to control stores
- Labels do not make a statistically significant difference in **green** and or **red**-labeled product sales in treatment stores relative to control stores
- That consumers do not decrease for red labels presents obstacle to change. Profit driven retailers are likely going to continue stocking red labeled seafood as long as consumers buy it…
- Maybe explore complementary solutions such as transitioning consumers to alternatives to red products



They did find Nemo… Thank you



Just in case slides



Levels of Dioxin in U. S. Food Supply, 1995.

Chart from May 2001 study by Arnold Schecter et. al., Journal of Toxicology and Environmental Health, Part A, 63:1–18]