The Value of Certification: Evidence from the U.S. Organic Food Market

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

Dan Yu

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

Doctor of Philosophy

Amil Petrin, Advisor

June, 2023

Acknowledgements

I am deeply grateful and indebted to my advisor, Amil Petrin, and thesis committee members, Thomas Holmes, Joel Waldfogel, and Joseph Mullins, for their valuable advice, support, and guidance. I also want to thank the participants of the Applied Microeconomics Workshop at the University of Minnesota for their helpful suggestions. I appreciate the Graduate School at the University of Minnesota for the financial support through the Henry Sandor Fellowship in Environmental Economics. Finally, I thank Caty Bach for her help, encouragement, and life wisdom.

Dedication

To Filip, who hated the idea of having kitties and later became their best friend.

And to my two fluffy therapists, Toffi and Śnieżka.

Abstract

Consumers may value some costly production procedures that improve sustainability without affecting the observed quality of products. Without credible disclosure, consumers cannot distinguish products compliant with these procedures from others, giving rise to a lemons problem and potentially inhibiting efficient trade. This dissertation explores whether third-party certification provides an effective solution to this problem in the context of the U.S. food market. This dissertation finds that consumers are willing to pay between 4 and 116% of a product's price for organic certification. Consistent with the theory that certification provides otherwise unverifiable information, estimates of the willingness to pay across product types covary negatively with the predictability of organic certification using other product characteristics. Results from the counterfactual analysis suggest a decrease in consumer welfare from removing the organic certification program, equivalent to 0.05% of total spending on organic products.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Many observers would like firms to change their production processes in costly ways that improve resource stewardship without changing the existing product attributes. Examples include organic farming practices, better treatment of workers, sustainable fishing, and carbon footprint reduction, all of which could "save the planet" in the eyes of some observers. For instance, serious questions have been raised about the adverse effects of conventional agriculture on the environment and soil fertility, including soil compaction, erosion, and degraded productivity [Reganold et al., 1987, Gerhardt, 1997], as well as loss of biodiversity [Lotter, 2003] and increased greenhouse gas emission [Dalgaard et al., 2006]. Because using those costly but potentially beneficial inputs or processes does not affect the observable quality of products, consumers cannot distinguish products produced with these inputs or processes from those without, giving rise to a classic lemons problem [Akerlof, 1970] and potentially inhibiting efficient trade. Many consumers willing to pay extra may not be able to directly perceive these benefits, as applying harmful synthetic pesticides does not change the taste of vegetables, and clothing made by well- or poorly-treated workers wears the same. Hence, the asymmetric information lies in the difference between the *perceived value*, which accounts for only the observed attributes of a product, and the *true value* of a product to consumers, taking into consideration these beneficial inputs or processes. Instinctively, these costly but potentially beneficial processes could conceivably be mandated by regulation; while many observers advocate these changes, they are unlikely to emerge due to a lack of consensus and the difficulty of enforcement.

However, if consumers could be assured that products contained these unobservable attributes, it is possible that their willingness to pay (WTP) for them would be sufficient to allow the use of costly but beneficial production processes. Accordingly, this dissertation explores whether third-party certification provides an effective solution to this problem. For the certification to work, consumers' *true value* needs to be higher than the cost of provision, allowing for profitable entry of these attributes.

This dissertation considers the role of USDA Organic certification in solving the asymmetric information problem related to organic farming practices in the food market. As suggested by the increasing demand for certified organic products and the high price premium over their conventional counterparts, some consumers value these attributes highly enough to pay extra when they are assured that the products are produced following the standards required by the certification. Therefore, this dissertation aims to answer the following three questions. First, since the certification status of a product may be related to its other existing characteristics, it is important to learn whether the extent of how informative the certification is to consumers varies across product types. Second, given the hefty price premium of certified organic products, this dissertation investigates whether consumers are willing to pay extra and how the WTP varies by product type and household demographics. Third, considering that the existence of organic certification may affect consumers' utility from products through changes in the choice sets and equilibrium prices, regardless of the organic status of their actual purchases, this dissertation evaluates how consumer welfare changes in response to the certification program.

To answer the proposed questions, I use transaction-level data of household food purchases with detailed product characteristics and demographic information from the Nielsen Consumer Panel, supplemented with information on ingredients and nutrient values for multi-ingredient products from the USDA FoodData Central. I focus on 15 product types[1] from six food categories. First, through a reduced form analysis, I identify the informativeness of the organic certification by calculating how

[1]These 15 product types are ready-to-eat cereal, cookies, baby food, vinegar, olive oil, herbal tea, yogurt, milk, eggs, ground beef, frozen fruit, salad mix, fresh herbs, apples, and spinach.

much variation in organic certification of products can be explained by the variation in other existing characteristics, such as ingredients used. Then, following a similar two-stage estimation strategy used in Goolsbee and Petrin [2004], I estimate a discrete choice demand model with a binary variable of organic certification using both micro and macro data. In this model, I allow the price coefficient to vary across household income groups and the taste for organic certification to vary across demographic variables, including income, household size, presence of children, marital status, education, employment, and occupation of household heads. To account for price endogeneity and potential correlation between the organic attribute and other unobserved product characteristics, I use instrumental variables (IV) proposed in BLP [Berry et al., 1995]. Using estimates from the demand model, I calculate consumers' WTP for organic certification for different product types. To quantify the welfare effects of organic certification, I compare consumer welfare in different counterfactuals, with and without organic certification.

1.1 Related Literature

This dissertation contributes to two strands of literature: (a) those measuring the willingness to pay (WTP) for eco-labels and (b) studies related to product quality disclosure and certification.

First, this dissertation adds to a large body of literature on consumers' willingness to pay for various types of certification, including organic certification [Batte et al., 2007, Sriwaranun et al., 2015, Kai et al., 2013, Rodríguez et al., 2007, Van Doorn and Verhoef, 2011, Krystallis and Chryssohoidis, 2005], carbon neutral [Choi and Ritchie, 2014], fair trade [De Pelsmacker et al., 2005], and other eco-labels [Loureiro et al., 2002]. These studies show that consumers have a positive WTP for organic certification and other eco-labels. Through a contingent valuation survey in the Argentinean market, Rodríguez et al. [2007] find that consumers' WTP for organic products ranges from 12% for milk to 110% for herbs. Sriwaranun et al. [2015] find similar results in Thailand, with WTP ranging from 51% for pork to 88% for kale. Batte et al. [2007] focus on the same certification as my paper does, the USDA organic certification. Estimating a probit model with data from a consumer intercept survey, they find that consumers are willing to pay a price premium for multi-ingredient processed food, specifically cereals, and those demographic variables have little effect on WTP. There are similar results about WTP for other eco-labels, such as 5% for apples grown with environmentally sound practices [Loureiro et al., 2002], 10% for fair-trade coffee [De Pelsmacker et al., 2005], and AU$21.38 per tonne of CO2 reduced through carbon offsets per person [Choi and Ritchie, 2014]. This dissertation adds to this literature from two aspects. First, I estimate a discrete choice model using data from actual purchases made by households, whereas the existing literature, to the best of my knowledge, measures WTP only using the contingent valuation method, which according to Hausman [2012] may have some limitations since consumers' responses to a survey may not be consistent with their actions. Second, my paper focuses on more product types, which include both *simple* products, such as fresh produce, and *complex* food, such as RTE cereals, while the existing studies focus on either a single product type or one of these two categories.

My work is also related to literature that studies quality disclosure and certification, most of which is summarized in Dranove and Jin [2010]. Regarding this topic, many studies investigate the value or effects of information provided through reputation, such as scores, rating, or reviews [Jin and Leslie, 2003, Jin and Sorensen, 2006, Hastings and Weinstein, 2008, Dranove et al., 2003, Reimers and Waldfogel, 2021, Saeedi, 2019]. While reputation works as a signal of quality for experience goods, for which consumers are able to realize the ex-ante unobserved product after purchase or consumption, it may not work well for credence goods, whose qualities cannot be observed even after purchase, such as quality related to organic practices. This dissertation focuses on the organic quality of products and investigates the value of the USDA organic certification to consumers.

1.2 Background

Though the concept of "non-chemical" farming, which has shaped the definition of organic farming today, can be dated back to the 1940s, a set of national standards was not developed until the early 2000s. In response to the rapid growth of demand in the organic market and a lack of clarity on what "organic" meant, the U.S. Congress passed the Organic Foods Production Act (OFPA) of 1990, authorizing a National Organic Program (NOP) administered by the USDA. This program regulates the U.S. organic market by developing and enforcing standard organic farming practices, which were finalized and implemented in 2002. To sell the products with the USDA Organic Seal, farmers and businesses need to meet the national organic standards and get certified by a third-party certifying agent accredited by the NOP[2].

The national organic standards regulate the substances and methods used in growing and handling agricultural products. In particular, organic operations need to adhere to the USDA organic regulations when using substances (or ingredients) by following the National List of Allowed and Prohibited Substances. The National List summarizes what non-organic substances may be used in organic production and handling when there is no organic alternative. In general, the standards prohibit the use of synthetic pesticides, fertilizers, and other substances. Instead, organic farmers usually use crop rotation to break the insect life cycle or other mechanical, and biological control for pests, weeds, and disease management. For soil fertility, animal and crop waste materials are used together with other cultivation practices, such as crop rotations and cover crops. Organic farming practices, including the limited

[2]Farmers and businesses who sell less than $5000 a year are exempt from certification. They may sell their products as organic if they follow the standards but cannot display the USDA Organic Seal.

application of synthetic substances, result in healthier soil, increased resistance to pests and diseases, less pollution of groundwater, less greenhouse gas emission, and improved biodiversity, which may be contributing factors to the increase in demand. Given strict requirements on the use of synthetic substances, organic farms tend to have higher costs per acre to produce their crops because organic practices that are alternative to synthetic pesticides and herbicides demand more labor-intensive activities. Both higher marginal costs in production and entry cost of certification contribute to the high price premium of certified organic products. Based on data used in this paper, I find that organic products have price premiums ranging from 6% to 69% of the average price of their conventional counterparts.

Despite the high price premium of certified organic products, the organic industry has been rapidly growing since the adoption of the national organic standards, from $12 billion in 2004 to $57.5 billion in 2021. In addition, according to The USDA Economic Research Service [2021], 2,314 to 3,302 organic products from the food and beverage category were introduced each year from 2015 to 2020 in the U.S., implying that there is an increase in consumers' choice set regarding products' organicity.

1.3 Result Summary

To summarize the findings of this dissertation, I divide these product types into two groups, *simple* and *complex* products, based on the complexity of their observed product characteristics. *Simple* products are those with a single ingredient or minimally processed, such as fresh produce or meat and dairy products. *Complex* products are those processed from a list of ingredients, such as ready-to-eat (RTE) cereals and cookies.

First, I find that the informativeness of organic certification differs across product types. Take ready-to-eat (RTE) cereals and eggs as an example. For RTE cereals, 81% of the variation in organic certification can be explained by the variation of other product characteristics, such as ingredients, flavors, and brands. In contrast, this number is only 17% for eggs. The difference in the predictability of certification status implies that when purchasing eggs, consumers obtain more information from organic certification that is not already captured by other characteristics, such as size, color, and grade. Overall, organic certification is more informative for *simple* products than for *complex* products. Meanwhile, the informativeness of certification is positively correlated with the price premium associated with the certification, implying a higher WTP for certification when the organic status is less likely to be predicted using other product characteristics.

Second, as implied by the demand estimates, organic products are more price elastic than conventional ones. Take eggs for an example. In response to a 1% price increase, the share of organic eggs decreases by 13%, whereas the share of conventional ones only decreases by 6%. The estimate of semi-elasticity with respect to organic certification suggests that getting an organic egg product certified leads to a 6% increase in its market share.

Third, this dissertation finds a positive WTP for organic certification that varies by product type, ranging from 4% for yogurt to 116% for eggs of the average unit price of the conventional counterparts. Consistent with the first finding on the informativeness of organic certification, the WTP is higher for products whose organic certification status is less predictable using other characteristics because consumers tend to place a higher value on otherwise unverifiable information. For example, the WTP for organic certification on RTE cereal and eggs is 18% and 116%, respectively. On average, the WTP for organic certification is 46% for *simple* products and 30% for *complex* products.

Fourth, some household demographics have small but statistically significant effects on coefficients of both price and organic certification, and as a result, the WTP varies little by demographic variable. Households with two to four members and the presence of children obtain higher utility from certified organic products. The additional years of schooling for household heads and additional employed family members slightly lower the household utility from organic certification.

Fifth, I calculate consumer welfare change when organic certification is removed as a product characteristic under counterfactuals with different assumptions on prices and product sets. Holding prices fixed, removing organic certification as a product characteristic results in a slight decrease in consumer welfare. For example, for fresh herbs, the welfare loss is equivalent to 0.5% of total spending on organic products, and the welfare loss increases to 1.3% when the counterfactual choice set includes only conventional products. When using the equilibrium prices in the counterfactual without the certification program, the welfare loss restricted to the previously organic products is worth 0.05% of the observed expenditure.

Finally, using equilibrium prices calculated from the estimated demand model and cost parameters, I find a decrease of 35% in the implied gross profit of the previously organic products. This result supports an argument that there may be less entry of products compliant with those costly but beneficial production processes without a credible disclosure mechanism.

Chapter 2

Certification: Informativeness and Price Pattern

2.1 Data

The primary data source for this dissertation is the consumer transaction data from Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business[1]. In addition, I use the USDA FoodData Central data to acquire information on other

[1]Researcher(s)' own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

product characteristics that are not provided in the Nielsen Consumer Panel.

2.1.1 Nielsen Consumer Panel Data

The primary data used in this paper come from the Nielsen Consumer Panel from 2014 to 2018, which records transaction-level purchases with prices, quantities, product characteristics, and household demographics.

I restrict my sample to 15 types of products and 24,000 households living in 10 designated market areas (DMA)[2] in the U.S. I define a market as a city-month combination, which results in a total of 600 markets.

I choose 15 product types from six categories: dry groceries, frozen foods, dairy, deli, packaged meat, and fresh produce. The procedures for choosing are as follows. First, for all categories, I choose a product type if it has more than 40,000 purchases and at least 12% of the purchases are organic in 2018. This step gives the following product types: baby food (strained), vinegar, herbal tea, olive oil, pre-cut fresh salad mix, frozen fruit, fresh herbs, apples, and spinach[3]. Second, because no product from the dairy and packaged meat category satisfies the requirements above, I choose the three most purchased product types with over 2% organic purchases from the dairy

[2]These ten DMAs include New York NY, Philadelphia PA, Boston (Manchester) MA-NH, Washington DC(Hagerstown MD), Atlanta GA, Chicago IL, Houston TX, Dallas-Fort Worth TX, Los Angeles CA, San Francisco-Oakland-San Jose CA

[3]Baby food, vinegar, herbal tea, olive oil are from the dry grocery category. Pre-cut fresh salad mix and frozen fruit are from the deli and frozen food categories. Fresh herbs, apples, and spinach are from the fresh produce category.

category, namely refrigerated milk, refrigerated yogurt, and fresh eggs. Third, in the packaged meat category, I choose ground beef[4], which has over 4% organic purchases. Finally, to include processed food types in the sample, I add ready-to-eat cereal (RTE) and cookies to the list, both of which have many observations but a small fraction of organic purchases. Since the Nielsen Consumer Panel does not provide information on ingredients and nutrient values of processed food, I supplement the sample with these two product types with a data set from the USDA FoodData Central, described in Section 2.1.2.

In this consumer panel, each observation records a purchase of a product characterized by a unique Universal Product Code (UPC)[5]. I observe the price, quantity, and associated product characteristics and household demographics for each purchase. Products of different types are characterized by different sets of attributes, described in appendix A. In general, the set of product characteristics used in estimation includes attributes related to the package and content of the product. For package-related variables, the dataset contains size and container type, depending on product type. Other content-related variables include but are not limited to color, flavor, shape, grade, and origin, some of which are extracted from acronyms used in the product description. Here, I take fresh eggs as an example. The set

[4]In the packaged meat category, only one product type, fresh meat, has a share of organic purchases larger than 1%. Because fresh meat is so loosely defined and includes products that are very different from each other, I choose the sub-type of ground beef.

[5]A UPC is the barcode representation that uniquely identifies a company's individual product.

of attributes includes grade, color, size of the eggs, and package size. In the product description provided by the dataset, "A" and "AA' mean grade A and grade AA, respectively. For household demographic information, I include household income, family size, an indicator of marital status, average years of education received by heads of households, the presence of children, the number of employed family members, and indicators of three occupation groups in which household heads are employed.

2.1.2 USDA FoodData Central

I obtain the information on ingredient lists and nutrient values for the RTE cereal and cookies from the USDA FoodData Central. This food data set can match 90% and 80% of the purchases of RTE cereal and cookies, respectively, recorded in the Nielsen Consumer Panel. I create indicator variables for ingredients that appear in at least 3% of observed purchases. For nutrition information, each variable gives the value of the corresponding nutrient per unit size. The set of nutrients includes protein, fat, carbohydrate, and the breakdowns of these three macronutrients. In addition, the data set also included calories per unit size and content values of minerals and vitamins, such as calcium, iron, zinc, vitamin A, vitamin E, etc.

I supplement the food data containing ingredients and nutrient values only for RTE cereal and cookies because the Nielsen data set provides enough information on the attributes of other product types. There are three cases in which the ingredient list and nutrient values do not add more information on product attributes. First, all products of the same type have the same single ingredient, such as apples, eggs, and ground beef. Second, the products of the same type have single ingredients that may vary by product but can be distinguished using the information provided in the Nielsen data[6]. Third, the products consist of multiple ingredients, but the ingredient lists are very similar conditional on other product attributes provided in the Nielsen data[7].

2.2 Descriptive Evidence

2.2.1 Price Premium of Organic Products

Organic products are more expensive than their conventional counterparts, conditional on other product characteristics. The price premium of organic products varies by product type. To study the price premium related to organic certification and its relation with other product characteristics, I quantify it in two steps. First, I regress the log of unit price only on the organic certification variable to measure the unconditional price premium of organic products. Then, I include other product characteristics and fixed effects of the brand and year to measure the price premium

[6]For example, fresh herbs include basil, rosemary, thyme, etc.

[7]For example, all banana baby food products include banana, lemon juice concentrate/citric acid, and vitamin C (ascorbic acid).

conditional on these variables, specified by the following equation:

$$logUnitPrice_j = \lambda_0^p + \lambda_1^p OrgCert_j + \lambda_X^p X_j + \eta_j^p, \tag{2.1}$$

where the left-hand side variable is the logarithm of unit price, and the right-hand side variables include an indicator of organic certification, and a vector of other product characteristics, X_j.

Table 2.1 summarizes the price premium of products with organic certification conditional on other product characteristics. Column (3) lists the share of organic purchases of each product, which ranges from 2% for cookies to 44% for fresh herbs. Column (1) shows the unconditional price premium of organic products for each type, implying that organic products are usually more expensive than conventional ones. Column (2) shows the conditional price premium, for which I include other product characteristics, such as package size, flavor, and ingredients, as well as fixed effects for brand and year. Among these 15 product types, vinegar has the highest conditional price premium, 69%, and yogurt has the lowest, 6%. The difference between unconditional and conditional price premiums indicates that other covariates included in the regression are correlated with the organic certification status. For example, some “specialty” brands sell only certified organic products with prices higher than organic products from firms who sell both organic and conventional products, so the inclusion of dummy variables for these brands reduces the coefficient of organic

Table 2.1: Price Premium of Organic Products

	(1) Unconditional[a]	(2) Conditional[b]	(3) Organic Share[c]
Vinegar	1.201 (0.009)	0.688 (0.007)	0.09
Eggs	0.714 (0.002)	0.580 (0.002)	0.16
Milk	0.758 (0.002)	0.545 (0.001)	0.19
Ground Beef	0.529 (0.005)	0.479 (0.004)	0.07
Salad Mix	0.646 (0.002)	0.392 (0.002)	0.22
Frozen Fruit	0.384 (0.004)	0.360 (0.003)	0.15
Cookies	0.427 (0.006)	0.321 (0.005)	0.02
Spinach	0.520 (0.003)	0.305 (0.003)	0.47
Apples	0.273 (0.003)	0.234 (0.003)	0.26
Baby Food	0.364 (0.003)	0.231 (0.005)	0.42
Fresh Herbs	0.085 (0.007)	0.206 (0.005)	0.44
Olive Oil	0.201 (0.006)	0.198 (0.005)	0.10
Herbal Tea	0.531 (0.005)	0.184 (0.005)	0.20
RTE Cereal	0.253 (0.003)	0.172 (0.005)	0.06
Yogurt	−0.259 (0.002)	0.057 (0.002)	0.16

[a]Regress log unit price on organic certification.

[b]Covariates include other product characteristics and fixed effects of brand and year.

[c]The share of purchases that are organic.

certification by taking away the price premium associated with the brands. Certain product characteristics are also recognized as premium. For example, fancier packages, such as glass jars, which are not related to any organic attributes, are more costly, but organic producers are more likely to use them to signal higher quality as a marketing tool.

2.2.2 Informativeness of the Organic Certification

Organic certifications for different product types do not mean the same things to consumers. One way to learn the difference in organic certification is through how informative it is to consumers. In other words, consumers can form predictions about how "organic" a product is based on its other characteristics, and this prediction has different levels of accuracy for different products. For example, the organic certification adds much more information to fresh eggs than RTE cereals because organic and conventional eggs look the same, but organic RTE cereals may have very different ingredient lists from the conventional ones[8]. To measure the informativeness of the organic certification for each product type, I regress the binary variable of organic certification on all other product characteristics, their squares and interactions, brands, and interactions among characteristics and brands, as well as fixed effects of year and geography. The specification is as follows:

$$\begin{aligned} OrgCert = \ & \beta_{o,0} + \beta_{o,1} Chara + \mu_{o,brand} + \mu_{o,year} + \mu_{o,city} \\ & + \beta_{o,2}(Chara \times Chara) + \beta_{o,3}(Chara \times Brand) + \epsilon_o \end{aligned} \tag{2.2}$$

[8] Assuming consumers are knowledgeable about organic standards, they would know a box of RTE cereal is not organic if they find BHT (butylated hydroxytoluene), a preservative prohibited in certified organic product, in the ingredient list.

, which allows for a very flexible specification to capture potential nonlinear relations between organic certification and other variables. The R-squared from this regression, reported in Figure 2.1, shows how much variation in the organic variable can be explained by the variation in other product characteristics, which measures how predictable the organic certification is across product types.

Figure 2.1: Predictability for Organic Certification

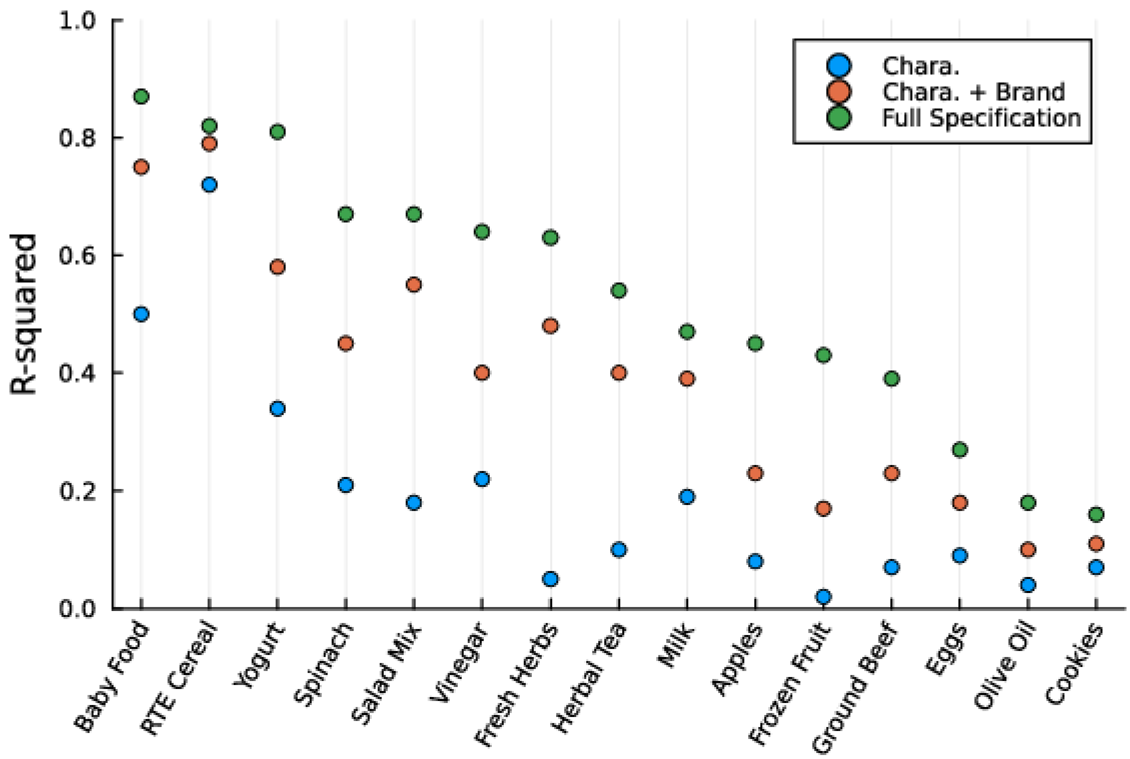


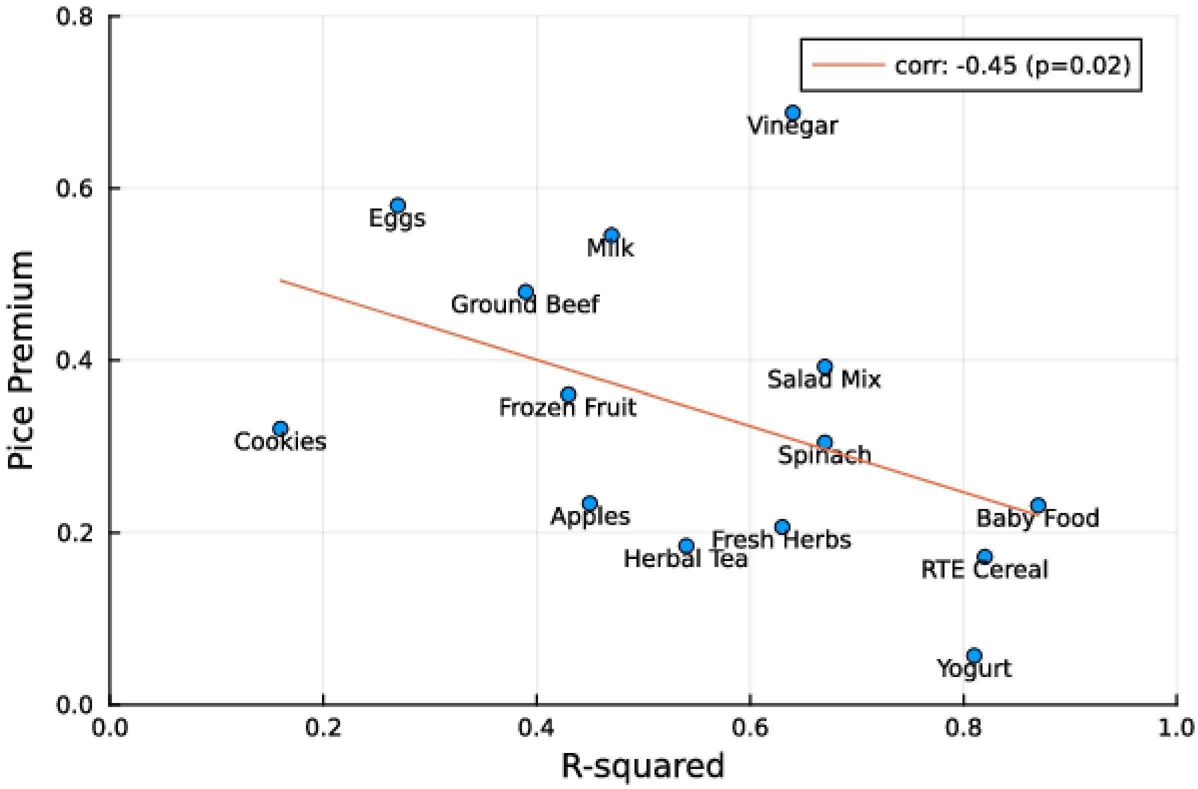
Figure 2.1 compares the R-squared values from the regression described above with three sets of variables included. The first set of variables, corresponding to the blue

dots, contains only other product characteristics as well as ingredients and nutrient values for RTE cereal and cookies. The regression shown by orange dots includes, in addition to the first set of variables, binary variables for brands. Accordingly, the regression shown by green dots includes, in addition to the second set, fixed effects of year and city as well as interaction and squared terms. There are two main findings from figure 2.1. First, for each product, there is a significant increase in R-squared when the regression includes the brand fixed effects, which implies that the organic certification has a large between-brand variation. This result is consistent with the fact that many brands featured as certified organic choose to certify most, if not all, of their products. One exception is RTE cereal because I include variables of ingredients and nutrients in the first regression, which explain most of the variation in the organic variable and is correlated with the brand variable. Thus, adding brand variables has little effect on the R-squared value. Second, the organic certification has different levels of informativeness across product types, for there is a large variation in the values of R-squared. For RTE cereal, 82% of variation in the organic variable is explained by observed variables. For eggs, only 27% of variation in the organic variable is explained by variation in other characteristics. This difference is consistent with the intuition that consumers are able to obtain much more information from other existing attributes of RTE cereals than eggs, leading to a higher predictive power on whether a product is organic or not for RTE cereals. Overall, consumers can learn much more from the organic certification on *simple* products than from that on *complex* products. Cookies are an exception, indicated by the last group of dots in figure 2.1 for the following reasons. First, the share of cookies that are organic is much lower compared to other product types. Hence, it is difficult to form accurate predictions based on a limited number of observations of certified organic products. Second, the product type of cookies contains a large set of products, which are more heterogeneous in characteristic space and cannot substitute for each other as products from other types can. Take the comparison of cookies and RTE cereals, for example. Consumers are more likely to substitute Froot Loops with Cocoa Puffs, but they may treat Oreos and Sugar Wafers as completely different products. Thus, the information on the existing characteristics of one product may have little implication on the organic status of the other, resulting in a lower value of R-squared.

Figure 2.2 shows the correlation between the price premium of certified organic products and the predictability of organic status using other product characteristics. Overall, there is a negative correlation of -0.45 with a p-value of 0.02[9]. It is consistent with the intuition that the price premium of organic is higher for those product types for which it's more difficult to predict whether it's organic based on other product characteristics.

[9]In figure 2.2, olive oil is not plotted because it is treated as an outlier of the 15 product types. The same figure of correlation between price premium and predictability for organic certification with all product types can be found in appendix A

Figure 2.2: Correlation: Price Premium and Predictability for Organic Certification by Product Type



Chapter 3

Consumers' Willingness to Pay for Certification

3.1 Model: Demand

I use a discrete choice specification to model household demand. In the model, markets are defined as city-period combinations. In each "market" $m \in \mathcal{M}$, there is a set of households, indexed by $i \in \mathcal{I}_m$, and a set of products, indexed by $j \in \mathcal{J}_m$, and an outside good, which accounts for options to purchase from other product types or to not buy any product.

The utility for household i from product j in market m consists of two parts, a market-level one and household-level one, given by:

$$U_{ijm} = \delta_{jm} + \zeta_{ijm}. \tag{3.1}$$

The market-mean utility from j, δ_{jm}, is the utility from product j common to all consumers in market m, which can be written as

$$\delta_{jm} = \alpha_0 p_{jm} + \gamma_0 \phi_j + X_j \beta + \mu_{brand} + \mu_{retailer} + \mu_{year} + \xi_{jm} \tag{3.2}$$

where p_{jm} is the price of product j in market m, ϕ_j is a binary variable that indicates whether a product is certified organic, X_j is a vector of other observable product characteristics, μ_{brand}, $\mu_{retailer}$, and μ_{year} are fixed effects of the brand, retailer and year, respectively, and ξ_{jm} is the market-level unobserved utility. The second term, ζ_{ijm}, is the utility of household i from j that is not captured in δ_{jm}, which is given by:

$$\zeta_{ijm} = \sum_{g=2}^{G} d_{ig} \alpha_g p_{jm} + \sum_{\ell \in \mathcal{L}} Z_{i\ell} \gamma_\ell \phi_j + \epsilon_{ijm}. \tag{3.3}$$

$g \in \{1, ..., G\}$ indicates household income groups, where G is the total number of income groups and $g = 1$ is the baseline group. The dummy variable d_{ig} is equal to one if household i is in income group g and zero otherwise. $\mathcal{L}$ is a set of household demographics, including household income groups, household size, marital status, the existence of a child, education, and occupation. $Z_{i\ell}$ is household i's value of demographics ℓ. ϵ_{ijm} is a household-level unobserved term and follows a Type I extreme value distribution.

In this model, I allow the household price coefficient to vary by income group and the coefficient on organic certification to vary by a subset of or all demographics. The price coefficient of household from income group g is $\alpha_0 + \alpha_g$, where α_0 is both the market-level price coefficient and the price coefficient of households within income group $g = 1$ because $g = 1$ is the baseline group and α_1 is normalized to zero. In other words, α_g with $g \neq 1$ is the difference in price coefficient of income group g and income group 1. Similarly, I allow the utility from organic certification to vary across household demographics, given by $\gamma_0 + \sum_\ell Z_{i\ell}\gamma_\ell$. γ_0 is the utility obtained from the organic certification by household with the baseline demographics, such as income group 1, family size 1, not married, etc. γ_ℓ is the utility from organic certification that is associated with demographic ℓ. For example, if $Z_{i\ell}$ is the average number of years of schooling among household heads, then γ_ℓ is the utility gained from organic certification for each additional year of schooling.

Given the distribution assumption on the household-level unobserved error term, ϵ_{ijm}, the probability of household i purchasing product j in market m is given by

$$s_{ijm}(p_m, X_m, \phi_m) = \frac{exp(\delta_{jm} + \sum_{g=2}^{G} d_{ig}\alpha_g p_{jm} + \sum_{\ell \in \mathcal{L}} Z_{i\ell}\gamma_\ell\phi_j)}{1 + \sum_{h \in J_m} exp(\delta_{hm} + \sum_{g=2}^{G} d_{ig}\alpha_g p_{hm} + \sum_{\ell \in \mathcal{L}} Z_{i\ell}\gamma_\ell\phi_h)}. \quad (3.4)$$

Then, I obtain the market share of product j in market m by integrating s_{ijm} over the distribution of households, which is given by

$$s_{jm}(p_m, X_m, \phi_m) = \int_i s_{ijm}(p_m, X_m, \phi_m) di. \quad (3.5)$$

The distribution of households can be retrieved from the micro-data.

3.2 Estimation

Following a similar two-step estimation strategy used in Goolsbee and Petrin [2004], I estimate market and household parameters separately using both transaction data and aggregate market share data.

3.2.1 First Stage: Household Parameters θ_h

In the first stage, I estimate the household parameters, denoted by θ_h, consisting of coefficients on price $\{\alpha_g\}_{g=2}^G$ and organic certification $\{\gamma_\ell\}_\ell$. This stage uses market share data to pin down the market-mean utility δ for any given θ_h and then uses transaction data to maximize the likelihood of household purchasing decisions by choosing θ_h. Berry et al. [1995] and Berry [1994] show that the market-mean utility δ can uniquely match, for a given θ_h, the market shares from data to those implied by the model. I find the fixed point $\delta(\theta_h)$ of the following contraction mapping:

$$T(s_m^D, \theta_h)(\delta_m) = \delta_m + log(s_m^D) - log(s_m(\theta_h, \delta_m)) \tag{3.6}$$

where s_m^D is the market shares observed in data and $s_m(\theta_h, \delta_m)$ is given by equation (3.5).

For given household parameters θ_h and a vector of market-mean utilities δ, the probability of household i choosing product j in market m is given by:

$$s_{ijm}(\theta_h, \delta) = \frac{exp(V_{ijm})}{1 + \sum_{h \in J_m} exp(V_{ihm})} \tag{3.7}$$

where J_m is the set of products available in market m and V_{ijm} is the indirect utility implied by the given parameters:

$$V_{ijm} = \delta_{jm} + \sum_{g=2}^{G} d_{ig}\alpha_g p_{jm} + \sum_{\ell} Z_{i\ell}\gamma_\ell \phi_j. \tag{3.8}$$

Next, I choose θ_h to maximize the likelihood function given by:

$$L = \prod_{m=1}^{M} \prod_{i \in \mathcal{I}_m} \prod_{j \in \mathcal{J}_\Updownarrow} s_{ijm}(\theta_h, \delta(\theta_h))^{Y_{ijm}} \tag{3.9}$$

where $s_{ijm}(\cdot)$ is given by equation (3.7), and Y_{ijm} is equal to 1 if household i purchases product j in market m and 0 otherwise.

3.2.2 Second Stage: Market Parameters θ_m

In the second stage, I uncover the market parameters $\theta_m = \{\alpha_0, \gamma_0\}$ using the implied market-mean utility $\hat{\delta} = \delta(\hat{\theta}_2)$ where $\hat{\theta}_2$ is the estimates from stage 1.

As shown in equation (3.2), δ is linear in market parameters, which can be identified by linear regression on price, organic certification, and other product characteristics. I include fixed effects of the brand, retailer, and market in the regression.

A product's price p_{jm} and its organic certification status ϕ_j may be correlated with other characteristics observed by consumers but not by researchers, ξ_{jm}. Following the choice of IVs used in BLP (Berry et al. [1995]), I construct the following IVs for prices: the sum of the k the characteristics of other products produced by the same firm f, $\sum_{h\in\mathcal{J}_f, h\neq j} X_{hk}$, and of other products produced by other firms, $\sum_{h\notin\mathcal{J}_f} X_{hk}$.

3.2.3 Elasticities

I calculate the price elasticity and semi-elasticity with respect to organic certification for each product.

The own-price elasticity of product j is given by:

$$\sigma_j = p_j \frac{\int (\alpha_0 + \sum_{g=2}^{G} \alpha_g d_{ig}) s_{ij}(1 - s_{ij}) di}{\int s_{ij} di} \tag{3.10}$$

The semi-elasticity of demand with respect to organic certification is given by:

$$\eta_j = \frac{\partial log(s_j)}{\partial \phi_j} = \frac{\int (\gamma_0 + \sum_\ell Z_{i\ell}\gamma_\ell) s_{ij}(1 - s_{ij})di}{\int s_{ij}di} \tag{3.11}$$

3.3 Results

3.3.1 Demand Estimates

In the first stage, I use the group of households with income higher than 100k as the baseline group and estimate the household-level income parameters on groups with income lower than 50k and between 50k to 100k. I estimate the model with two specifications. In the first case, I allow household taste on organic certification to vary only by income group, and, in the second case, to vary by both income group and other demographics, which include family size, education, marital status, presence of a child, employment, and occupation. Table 3.1 shows a comparison of estimates from these specifications for eggs. Column (1) shows results from the limited model, in which I interact the organic certification with only income groups, and column (2) shows the results from the full model, in which I add average years of schooling among household heads, the number of employed household heads, and binary variables of household size, marital status, presence of children under 18, and three groups of occupations. From column (2), coefficients on organic certification interacted with

Table 3.1: Demand Estimates for Fresh Eggs

	(1) Limited	(2) All
Price Parameter		
Income $< 50k$	0.019	0.019
	(0.018)	(0.018)
Income $50 - 100k$	0.154^{***}	0.154^{***}
	(0.017)	(0.017)
Organic Certification		
Income $< 50k$	0.0	0.017
	(0.011)	(0.012)
Income $50 - 100k$	-0.021^{*}	-0.016
	(0.01)	(0.01)
HH Size 2		0.045^{***}
		(0.012)
HH Size 3-4		0.046^{***}
		(0.013)
HH Size ≥ 5		-0.01
		(0.016)
Education[a]		-0.007^{***}
		(0.002)
Married		0.03^{*}
		(0.012)
Single[b]		0.007
		(0.012)
Presence of Children[c]		0.067^{***}
		(0.014)
Employment[d]		-0.004^{***}
		(0.001)
Occupation Group 1[e]		0.01
		(0.009)
Occupation Group 2		-0.014
		(0.01)
Occupation Group 3		-0.04^{**}
		(0.013)
2nd-stage Price	-44.358^{***}	-44.362^{***}
	(2.963)	(2.963)
2nd-stage Organic	9.2^{***}	9.214^{***}
	(1.23)	(1.23)

[a]Average years of schooling among household heads.
[b]Excluding divorced/separated and widowed. [c]Under age of 18.
[d]The number of employed household heads.
[e]Defined by data. See Appendix **??** for details of occupation groups.

households with two to four members and the presence of children, 0.05 and 0.07, are statistically significant but small compared to the base coefficient $\gamma_0 = 9.2$. Similarly, an additional year of schooling and an additional employed household head decreases the utility by a small amount. Comparing these two columns, I find that adding variables of other household demographics has little influence on the estimates from the second stage. The estimation for other product types exhibits similar results that including more household demographics in the first stage does not change the second stage estimates much, and I only show the results without other demographic variables for other products in the following sections.

Table 3.2 summarizes the demand estimates for the chosen 15 product types. Column (1) shows the price coefficients for the base group of households, α_0, which is the high income group. Columns (2) and (3) list the household-level difference in price coefficients, denoted by α_g. The marginal dis-utility from price for households outside the base group is $\alpha_0 + \alpha_g$. Columns (1)–(3) show that the base price coefficients, α_0, are negative and statistically significant, and the difference in price coefficients for other groups, α_g, are very small in absolute value compared to α_0, indicating that household income has little effect on the marginal dis-utility of price. For cookies and ground beef, the medium-income group has a smaller coefficient than the high-income group, and the difference is statistically significant. Similarly, columns (4)–(6) summarize the estimates of coefficients on organic certification. The marginal

Table 3.2: First and Second Stage Estimates

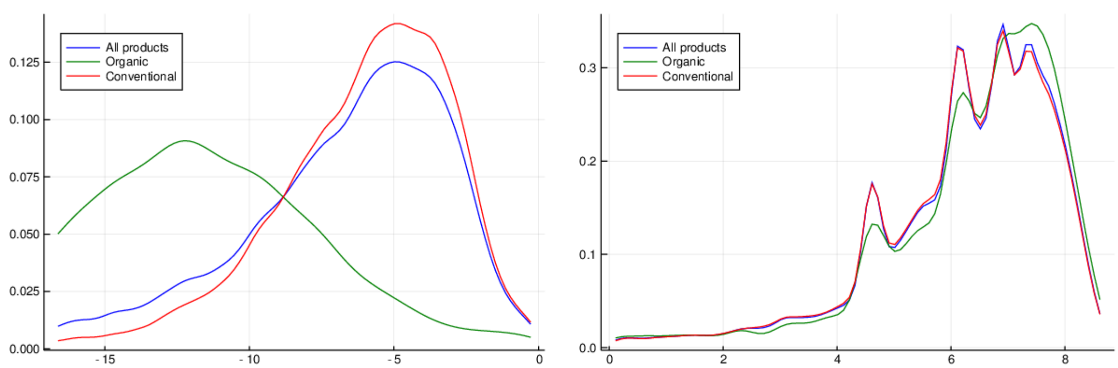
Product	Price coefficients			Organic coefficients		
	base group	difference for other groups		base group	difference for other groups	
	$> 100k$	$< 50k$	$50 - 100k$	$> 100k$	$< 50k$	$50 - 100k$
	(1)	(2)	(3)	(4)	(5)	(6)
Milk	-89.681^{***}	1.008^{***}	0.710^{***}	1.478^{***}	0.009	0.030^{***}
	(2.487)	(0.072)	(0.069)	(0.061)	(0.007)	(0.007)
Egg	-44.358^{***}	0.019	0.154^{***}	9.200^{***}	0.000	-0.021^{*}
	(2.963)	(0.018)	(0.017)	(1.230)	(0.011)	(0.010)
Frozen Fruit	-22.125^{***}	0.095^{*}	0.050	2.094^{***}	-0.005	-0.013
	(1.307)	(0.048)	(0.044)	(0.136)	(0.028)	(0.025)
Salad Mix	-19.274^{***}	0.211^{***}	0.160^{***}	1.880^{***}	-0.021^{*}	-0.035^{***}
	(0.958)	(0.014)	(0.013)	(0.119)	(0.010)	(0.009)
Spinach	-18.329^{***}	0.030	-0.012	1.814^{***}	0.017	0.001
	(1.619)	(0.034)	(0.029)	(0.184)	(0.022)	(0.018)
Ground Beef	-10.718^{***}	-0.056	-0.108^{***}	1.113^{***}	-0.041	0.017
	(1.017)	(0.031)	(0.030)	(0.130)	(0.035)	(0.032)
Herbal Tea	-9.152^{***}	-0.017	-0.036	0.083^{*}	0.016	0.010
	(1.100)	(0.046)	(0.042)	(0.039)	(0.026)	(0.025)
Yogurt	-7.728^{***}	0.028	0.046^{*}	0.051^{**}	0.004	0.005
	(0.285)	(0.021)	(0.018)	(0.017)	(0.008)	(0.008)
Baby Food	-5.024^{***}	-0.010	0.033	0.190^{***}	-0.021	-0.001
	(0.275)	(0.041)	(0.035)	(0.029)	(0.020)	(0.017)
RTE Cereal	-4.523^{***}	-0.003	-0.011^{*}	0.230^{*}	-0.048^{***}	0.002
	(0.492)	(0.003)	(0.005)	(0.105)	(0.013)	(0.012)
Apple	-4.300^{***}	0.004	0.008^{*}	0.740^{***}	-0.040^{*}	0.002
	(0.585)	(0.004)	(0.004)	(0.172)	(0.017)	(0.015)
Olive Oil	-3.945^{***}	0.147^{***}	0.058	0.124^{**}	-0.032	-0.034
	(0.481)	(0.040)	(0.039)	(0.041)	(0.031)	(0.029)
Vinegar	-3.163^{***}	0.147^{***}	0.058	0.185^{***}	-0.032	-0.034
	(0.499)	(0.040)	(0.039)	(0.049)	(0.031)	(0.029)
Cookie	-0.473^{***}	-0.014	-0.043^{***}	0.100^{***}	0.034	0.011
	(0.042)	(0.008)	(0.008)	(0.020)	(0.021)	(0.022)
Fresh Herb	-0.145^{***}	-0.005	-0.004	0.498^{***}	-0.000	-0.009
	(0.029)	(0.003)	(0.002)	(0.111)	(0.027)	(0.021)

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

utility from organic certification for demographic group ℓ household is given by $\gamma_\ell + \gamma_0$. From column (4), for each product type, the base group has a positive and statistically significant estimate of the organic coefficient, γ_0. Similar to household-level price coefficients, columns (5) and (6) show that for most product types the difference in household-level organic coefficients for different income group are small compared to γ_0. For olive oil and vinegar, there is a relatively large decrease in the organic coefficient for households from lower-income groups, but the differences are not statistically significant. Allowing the coefficient on organic certification to vary only by income group, I find that household income has either ambiguous or insignificant effect on how much utility consumers obtain from organic certification. Note that the values of α_0 or γ_0 should not be compared among product types because I estimate the coefficients for each type separately. Only the ratio of coefficients on organic certification and price, γ_0/α_0, can be interpreted as the WTP for organic certification and compared among product types, which is described in detail in section 3.3.3.

3.3.2 Elasticity

Figure 3.1: Distribution of Elasticity (Fresh eggs)



(a) Own-price elasticity (b) Organic semi-elasticity

Figure 3.1 shows the distribution of (a) own-price elasticities and (b) organic semi-elasticities of fresh eggs. Comparing the distribution of own-price elasticities of organic and conventional eggs, plotted by the green and red curves, we can see that organic eggs are more price elastic. A one percent increase in price decreases the market share of conventional eggs by 6%, whereas the decrease in market share of organic eggs is 13%. This result is consistent with the fact that, compared with conventional products, organic ones are closer to "luxury goods," whose demand is more sensitive to price change. Since the overall market share of organic eggs is 6%, the pooled distribution of price elasticities is very close to the distribution of price elasticities of conventional eggs. Similar results also apply to other product types, which are shown in appendix B In figure 3.1 (b), organic and conventional eggs have a similar distribution of semi-elasticities with respect to organic certification. On average, getting a conventional product certified as USDA organic increases its market share by 6.2%, and removing the certification of an organic product reduces its market share by 6.4%.

3.3.3 Willingness to Pay (WTP)

I calculate the WTP of consumers from income group g for the organic certification by dividing the marginal utility from the certification, $\gamma_0 + \gamma_g$, by their price coefficient, $\alpha_0 + \alpha_g$.

Table 3.3 summarizes the WTP for organic certification both as the percentage of average prices and in dollar terms for each product with a common unit size. The first row shows that consumers are willing to pay $2.5 more for a dozen organic eggs,

Table 3.3: WTP Dollar Value vs. Price Percentage

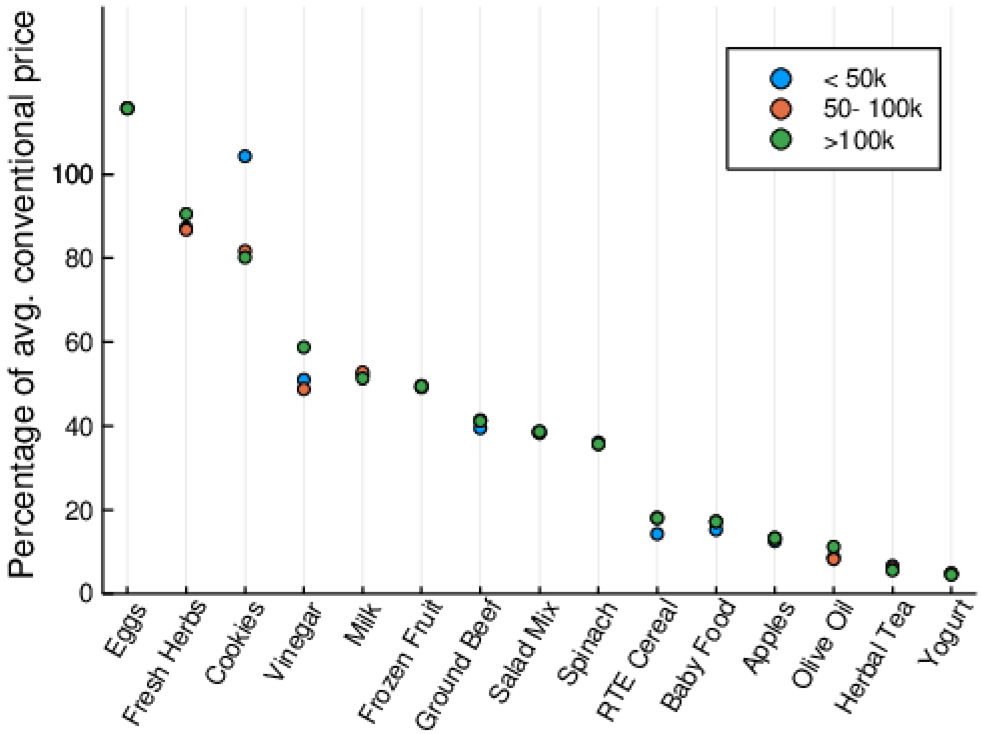
Product	Unit	WTP (dollar)	WTP (%)[a]
Eggs	1 dozen	2.49	116
Fresh Herbs	1 oz	3.42	91
Cookies	16 oz	3.39	80
Vinegar	16 oz	0.94	59
Milk	1 gallon	2.11	51
Frozen Fruit	48 oz	4.54	49
Ground Beef	1 lb	1.66	41
Salad Mix	16 oz	1.56	39
Spinach	1 lb	1.58	36
RTE Cereal	16 oz	0.81	18
Baby Food	4 oz	0.15	17
Apples	1 lb	0.17	13
Olive Oil	16 oz	0.5	11
Herbal Tea	20 bags	0.18	6
Yogurt	16 oz	0.11	5

[a] Percentage of the avg. unit price of conventional products.

which is equivalent to 116% of the average price of a dozen conventional eggs. The last row shows that consumers are only willing to pay 5% more for organic yogurt compared to conventional one, which can be translated to 11 cents for a 16 OZ tub. All values reported in table 3.3 are based on estimated parameters for the baseline income group, for the WTP varies little by household income group.

Figure 3.2 depicts the variation of WTP among income groups. For most product types, the WTP is almost identical for all income groups. For cookies, the WTP by the low-income group is higher than the other two groups, which is counterintuitive and could potentially be explained by the extremely small share of organic

Figure 3.2: WTP by Household Income Group

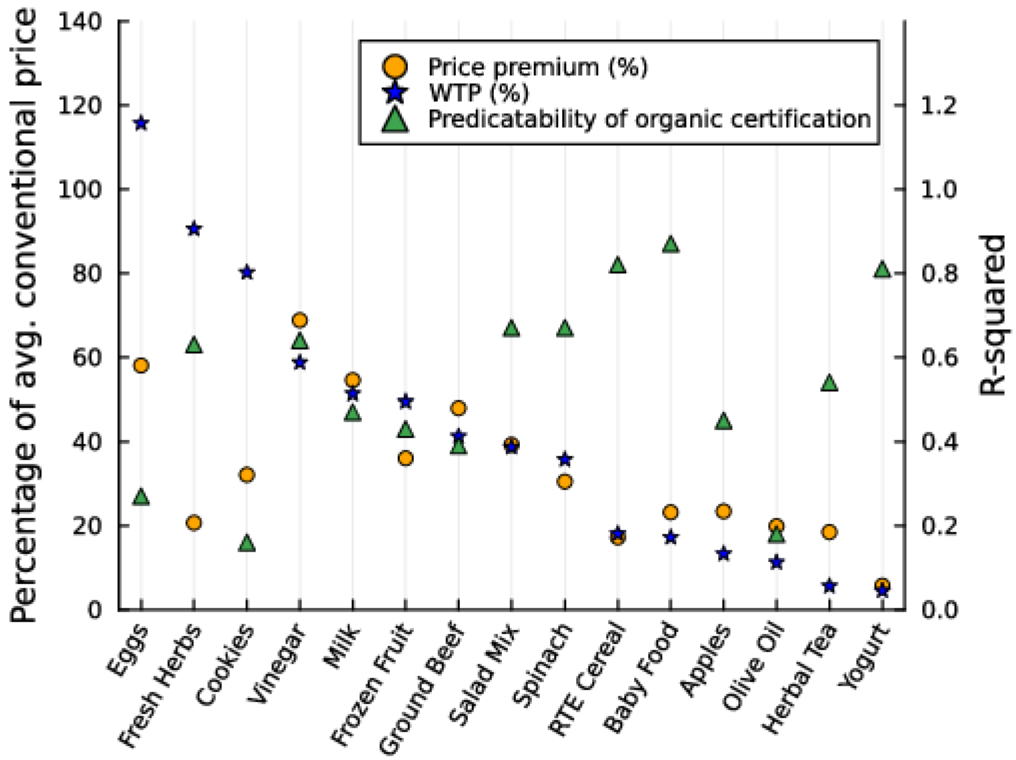


purchases[1]. For fresh herbs, vinegar, RTE cereal, and olive oil, the WTP of the high-income group is only slighter larger than that of the other two groups. Given the little variation in WTP among different income groups, I use only the estimates of the baseline group, the high-income group, in reporting results.

According to table 3.2, households from different income groups have almost identical coefficients on price and organic certification, which are the only determinants of WTP for the organic certification, so I only report the WTP of the households from

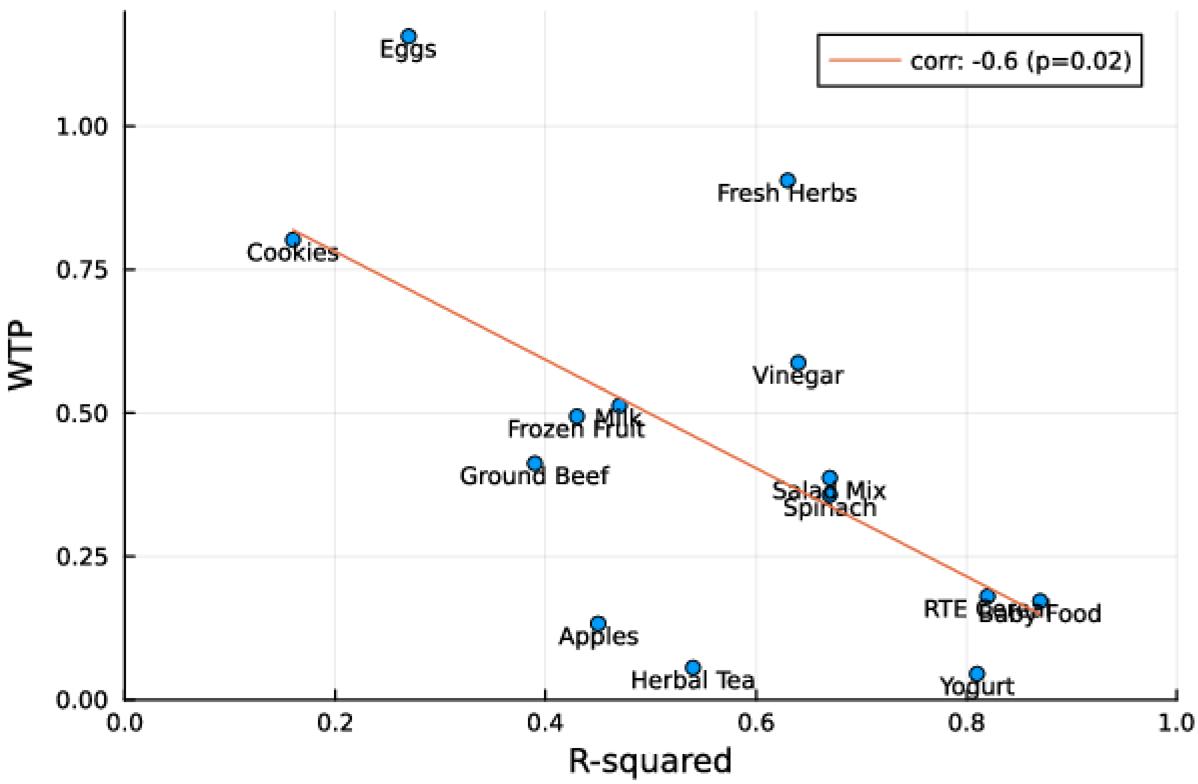
[1]Table 2.1 shows that only 2% of cookie purchases are organic.

Figure 3.3: WTP, Price Premium, and Predictability for Organic Certification by Product Type



the control group. Figure 3.3 compares the price premium, WTP, and predictability of organic certification by product type. Price premium and WTP are expressed in percentages of the average unit price of conventional products, and the predictability of organic certification is measured in terms of R-squared from regressing organic certification variable on all other characteristics. There are three main findings. First, consumers' WTP for organic certification is positive for all chosen product types, which supports the growing sales and market share of organic food in the past two

Figure 3.4: Correlation: WTP and Predictability for Organic Certification by Product Type



decades. Second, for most product types, except fresh herbs and cookies, the WTP follows a similar order of types as the price premium. There is a big gap between the price premium and WTP for eggs, fresh herbs, and cookies. Third, the WTP varies by product type, ranging from 4% for yogurt to 116% for eggs. Consistent with the theory that certification provides otherwise unverifiable information, estimates of the willingness to pay across product types covary negatively with the predictability of organic certification using other product characteristics. Figure 3.3

reports a negative correlation of -0.6 with a p-value of -0.02[2]. According to these two figures, the predictability of eggs' certification status using other characteristics is much less than that of RTE cereals, implying that the organic certification on eggs brings much more new information to consumers and therefore has a larger WTP. As shown in figure 3.3, the WTP of organic eggs is 116%, which far exceeds that of RTE cereal, 18%. Also, consumers may value different aspects of organic products. For example, if consumers perceive organic as pesticide-free, they may be willing to pay more for organic produce that is directly consumed than those that can be easily cleaned, which could explain why the WTP for salad mix, 39%, is higher than that of apples, 13%. If consumers want to avoid chemicals, such as antibiotics and growth hormones, they would value the certification more for meat and dairy products. Among the chosen four meat and dairy products, three of them have relatively high WTP, namely eggs, milk, and ground beef, with WTP 116%, 51%, and 41%.

[2] In figure 3.4, olive oil is not plotted because it is treated as an outlier of the 15 product types. The same figure of correlation between price premium and predictability for organic certification with all product types can be found in appendix B

Chapter 4

Counterfactual Analysis: Welfare Effects of Certification

In this chapter, I evaluate the welfare effects of the USDA Organic certification using estimates from the demand model and estimated cost parameters. In addition, based on equilibrium prices calculated using firms' first-order conditions, as well as the implied marginal costs (MC) and market shares, I am able to compare product profit and draw inferences on potential product entry decisions. My estimates support a story in which the removal of the organic certification program decreases consumer welfare through two channels: elimination of information provided by the certification and a potential reduction in the product set due to some organic products failing to enter the market.

4.1 Model: Supply

Each firm f produces a set of products $\mathcal{J}_f$. For each product, $j \in \mathcal{J}_f$, firm f chooses its price p_{jm}, organic certification decision ϕ_j, and other characteristics W_j, which include the set of characteristics observed by the consumer, X_j.

The MC of producing j depends on j's characteristics, W_j, and the firm's certification decision ϕ_j. If j is produced following the rules required by the certification, its MC is higher compared to conventional production. This higher MC of organic products may come from a higher requirement of labor intensity and skills and more expensive alternatives of material used. I use $c_j(W_j)$ to denote the MC of producing j when it is produced in a conventional way. For the same product with the same product characteristics, the MC is scaled up by e^{κ} if it is produced following the organic standards, denoted by $e^{\kappa}c_j(W_j)$. Each certified organic product has a fixed cost, c_F, which includes the certification fee and cost of machinery, equipment, and marketing that do not vary by product quantities. Since firms cannot convince consumers that they follow the organic requirements without certification, they do not have the incentive to produce a product in an organic way with a higher cost without getting it certified.

The profit maximization problem of firm f is given by:

$$\max_{\{p_{jm},\phi_j\}_{j\in\mathcal{J}_f}} \sum_{j\in\mathcal{J}_f} \Big\{ \big[p_{jm} - e^{\kappa\phi_j} c_j(W_j)\big] s_{jm}(p_j, \phi_j, X_j) - \phi_j c_F \Big\} \tag{4.1}$$

where $s_{jm}(p_m, \phi, X)$ is the market share of product j in market m. Assuming firms can only choose prices in the short run, the vector form of the first-order conditions of all firms with respect to prices is given by

$$s(p, \phi, X) - \Delta(p, \phi, X)\big[p - e^{\kappa\phi} \odot c\big] = 0 \tag{4.2}$$

where

$$\Delta_{jr} = \begin{cases} -\frac{\partial s_{jr}}{\partial p_j} & \text{if } j \text{ and } r \text{ are produced by the same firm} \\ 0 & \text{otherwise} \end{cases} \tag{4.3}$$

The equilibrium prices are given by

$$p^* = e^{\kappa\phi}c + \Delta^{-1}(p^*, \phi, X)s(p^*, \phi, X) \tag{4.4}$$

4.2 Estimation

Given by equation 4.4, with the observed equilibrium prices and market shares and Δ implied by estimated demand parameters, I can uncover the MC, $mc_j = e^{\kappa\phi_j}c_j$, for each product. Assume that the logarithm of MC if produced conventionally, $log(c_j)$, is linear in product characteristics, W_j. Then, to uncover κ and other cost parameters, I estimate the following equation

$$ln(mc_j) = \kappa\phi_j + W_j\tau + \mu^c_{brand} + \mu^c_{city} + \mu^c_{year} + \omega_j. \tag{4.5}$$

The MC are calculated using 4.4 with prices and market shares from the data. The identification of the key cost parameter, κ, relies on the assumption that the error term, ω_j, is uncorrelated with a product's organic status when controlling for its brand, the city where it is sold, and the year.

With estimated MC, I am able to solve the new equilibrium price vector p for different counterfactuals using firms' pricing strategy given by equations 4.4.

4.3 Results

The parameter of main interest is the ratio of MC for organic and conventional production (organic-conventional MC ratio), denoted by e^κ. Estimating equation 4.5, I uncover parameter κ for each product type, show in table 4.1.

Column (1) of table 4.1 shows that all types of product have positive κ without any restriction in estimation and, therefore, $e^\kappa > 1$, implying that the MC for organic production is higher than that of conventional production, consistent with the expectation. The estimated κ is statistically significant for all product types. Implied by e^κ in Column (2), five out of the chosen 15 product types have a MC of organic

Table 4.1: Cost parameter for organic production (κ) by product type

Product	(1) κ	(2) e^{κ}
Baby Food	1.195 (0.024)	3.305
Olive Oil	1.125 (0.033)	3.079
Milk	0.797 (0.005)	2.220
Herbal Tea	0.743 (0.023)	2.103
Ground Beef	0.740 (0.017)	2.096
RTE Cereal	0.657 (0.038)	1.929
Eggs	0.652 (0.006)	1.918
Salad Mix	0.480 (0.007)	1.616
Frozen Fruit	0.466 (0.007)	1.594
Yogurt	0.392 (0.016)	1.480
Vinegar	0.319 (0.025)	1.376
Spinach	0.287 (0.010)	1.332
Apples	0.272 (0.007)	1.312
Fresh Herbs	0.122 (0.020)	1.130
Cookies	0.002 (0.002)	1.002

production more than two times as much as the conventional production, closely followed by RTE cereal and eggs with organic MC almost two times as much. The large estimate of e^{κ} supports a story in which, without organic certification, firms have a large incentive not to follow organic standards because it is much more expensive to do so, and consumers also rationally anticipate products are no longer organic. For cookies and fresh herbs, this ratio is very close to 1, implying that organic cookies or herbs are only slightly more costly to produce compared to conventional ones.

As reported in estimated annual production costs and returns by USDA, the cost of producing organic milk is 1.84 times as much as the cost of producing conventional milk. The same estimate for milk from my model is 2.22, shown in table 4.1, which is comparable to the USDA estimate and speaks in favor of the reliability of my empirical strategy of cost estimation. The discrepancy in the organic-conventional MC ratio can be explained by that the USDA reports the farm-level costs of milk production, whereas this dissertation focuses on the costs of final products, which additionally include handling, marketing, etc.

Figure 4.1 summarizes the difference in MC of organic and conventional products for each product type and compares it with the consumers' WTP in dollar terms. ΔMC represents the difference between organic MC, $e^{\kappa}C$, and conventional MC, C. For most product types, the WTP for organic products can compensate the MC difference in production. For baby food and apples, the WTP for organic products can cover most of the increase in the MC of organic production. For herbal tea, salad mix, ground beef, and milk, the increase in MC is much higher than the WTP for organic products.

Figure 4.1: ΔMC vs WTP

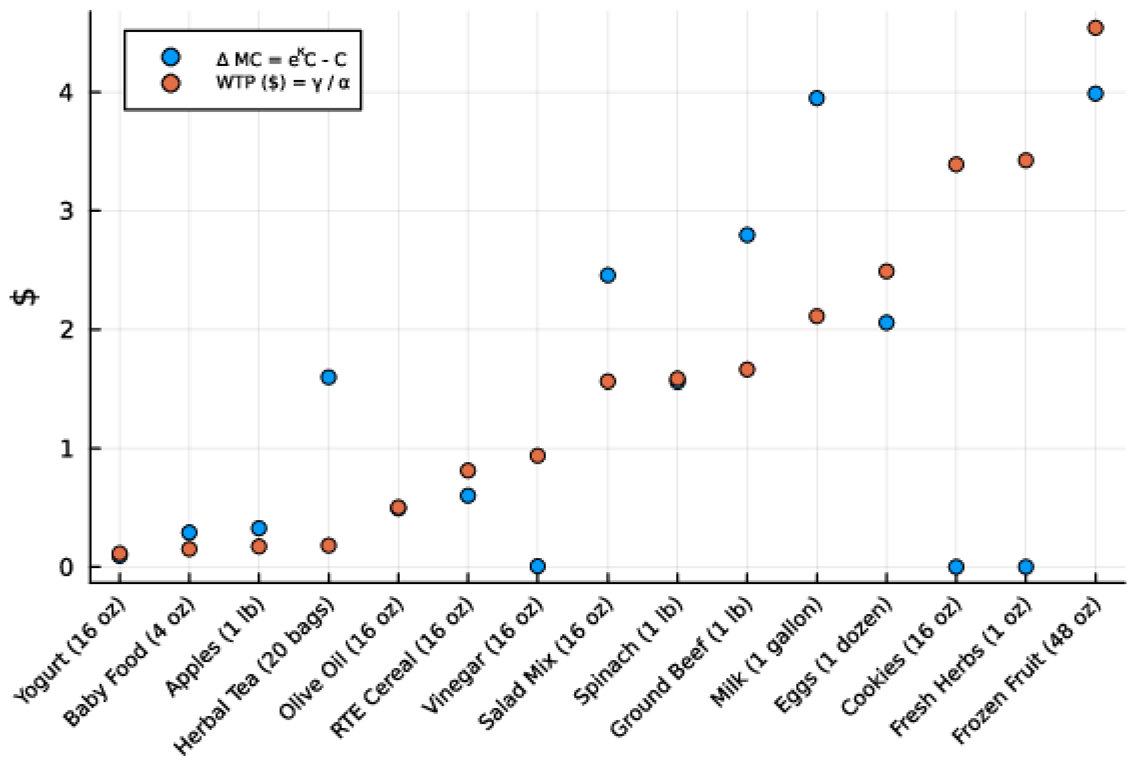


Table 4.2: Policy Counterfactuals

	Counterfactual	Description
(1)	*hide certification*	Same prices and products without certification. No prior knowledge about organic.
(2)	*no organic product*	(1) + all organic products removed.
(3)	*new equilibrium with cost reduction*	Evaluated at equilibrium prices when firms produce all products conventionally.

4.4 Counterfactuals

Table 4.2 summarizes the assumptions associated with prices, product characteristics, and choice sets under each counterfactual. I denote the current policy by *equilibrium with certification* and consider the following counterfactuals in which organic certification is removed as a product characteristic. For the first two counterfactuals, I use the observed equilibrium prices with certification and compare the consumer welfare in cases with different product sets. In the *hide-certification* counterfactual, both prices and the set of products are the same as in *equilibrium with certification*. Assuming consumers do not believe that the organic quality is present without certification, the consumer welfare change, in this case, is driven by the lack of information that is previously provided by the organic certification. Then, in the *no organic product* counterfactual, I eliminate entry of all organic products and evaluate this extreme case in which the cost of organic production is so expensive that no such product can have profitable entry. This counterfactual does not take into consideration the supply-side decision on pricing and certification and hence only provides a measure of how much consumers' welfare decreases in response to a reduced product set. In counterfactual of *new equilibrium with cost reduction*, when the USDA certification program shuts down, firms choose to switch those previously organic products to conventional with lower production costs given that consumers can no longer verify the organic status and therefore rationally anticipate no organic product is provided. In this setting, consumer welfare is evaluated at the new equilibrium prices.

The following approach describes how consumer welfare change is measured. First, I calculate the compensating variation (CV) by evaluating the difference in consumers' expected utility with and without certification in monetary terms. Recall the utility of consumer i from product j in market m with certification from equation 3.1, which is denoted by $U^{(0)}_{ijm}$ in this section. Under *hide-certification* denoted by (1), when facing the same set of products with the same prices but without certification, the consumer i's utility is

$$U^{(1)}_{ijm} = U^{(0)}_{ijm} - (\gamma_0 + \sum_{\ell} Z_{i\ell}\gamma_\ell)\phi_j \tag{4.6}$$

Under *no organic product* counterfactual, consumers have the same utility specification but different choice sets. In *new equilibrium with cost reduction*, consumer utility is evaluated at new equilibrium prices and the same product set. Given table 3.2, the difference in household level coefficients is essentially zero relative to the coefficients for the base group, so I use estimates for the base group, α_0, and γ_0, in calculating the CV. The CV in counterfactual (x) is given by

$$CV^{(x)} = \frac{\mathbb{E}[U^{(x)}|\mathcal{J}^{(x)}] - \mathbb{E}[U^{(0)}]}{\alpha_0}, \tag{4.7}$$

where $\mathcal{J}^{(x)}$ is the product set assumed under counterfactual (x), and $\mathbb{E}[U^{(x)}]$ is the expected utility from each purchase[1].

Then, consumer welfare change from removing the certification program is measured by the CV divided by the average spending on a product. Take eggs, for example. The CV from each purchase of a dozen eggs is around minus 2.5 cents, and the average price of a dozen organic eggs is 4.22 dollars, resulting in a 0.6% welfare decrease in *hide-certification*. Since the CV and prices are market-specific, the average welfare change is weighted by the number of purchases in each market.

Table 4.3 summarizes the welfare change for each product type under the first two counterfactuals, *hide-certification* and *no organic products*. Column (1) shows that for most product types, there is a slight welfare decrease when the organic certification is removed as a product characteristic. Under *hide-certification* counterfactual, salad mix has the most significant welfare decrease, measured as 0.36% of spending on organic products. In this counterfactual, I use the current equilibrium prices and product sets, so the welfare change is solely from removing the information provided by the organic certification. Moreover, as shown in column (2), as I use the choice sets without organic products, the expected utility becomes smaller compared to the current case, resulting in a larger welfare decrease from product types with higher

[1] The expected utility also takes into consideration of the outside goods. Since I do not observe when consumers decide not to purchase any product from a certain type and the expected utility is the same regardless of the final choice, I consider the value of $\mathbb{E}[U^{(x)}]$ as the expected utility per purchase.

Table 4.3: Welfare change (%) under counterfactuals (1) & (2)

Product	(1) *hide certification*	(2) *no organic product*
Yogurt	-0.17	-2.76
Fresh Herbs	-0.50	-1.27
Cookies	-0.05	-0.54
Baby Food	-0.07	-0.43
Salad Mix	-0.36	-0.42
Milk	-0.32	-0.42
Vinegar	-0.05	-0.31
Spinach	-0.15	-0.18
RTE Cereal	-0.03	-0.16
Apples	-0.06	-0.12
Eggs	-0.08	-0.08
Frozen Fruit	-0.06	-0.07
Ground Beef	-0.05	-0.07
Olive Oil	-0.01	-0.07
Herbal Tea	-0.00	-0.06

organic shares under *equilibrium with certification*. Fresh herbs, for example, have the highest organic share of 44% among the chosen 15 types. In the *hide-certification* counterfactual where prices and product sets are fixed, the welfare decrease is 0.5% of the total spending on organic fresh herbs. When I assume that all organic products would not enter the market without the certification program, the welfare loss is equivalent to 1.3% of the total organic spending. Overall, the welfare decrease from these two counterfactuals for most product types is insignificant, likely due to the small number of certified organic products relative to the number of all products observed in the data.

Table 4.4: Equilibrium outcome change (%) from *new price with cost reduction*

	Organic set	*Conventional set*
Price	-34%	0%
Marginal cost	-48%	0%
Sales	-36%	0.5%
Gross profit	-36%	0.5%

For the policy counterfactual of *new equilibrium with cost reduction*, I restrict my analysis to egg purchases made in 2018 for illustrating purposes. Table 4.4 summarizes changes in equilibrium outcomes compared with the *equilibrium with certification*. To make the notation clear, I define *organic collection* as the set of products that are certified organic in *equilibrium with certification* and define *conventional collection* similarly. Under *new equilibrium with cost reduction*, the equilibrium prices, on average, have a 34% decrease for products from *organic collection* but have little change for products from *conventional collection*. In monetary terms, the equilibrium price decrease and cost reduction of products from the *organic collection* are essentially the same, 20.2 and 20.3 cents per egg, respectively. The difference in welfare as a percentage of expenditure is calculated by:

$$\frac{\mathbb{E}[U(p^{(0)})] - \mathbb{E}[U(p^{(3)})]}{p^{(0)}} \tag{4.8}$$

where $p^(0)$ is the price from the data, and $p^{(3)}$ is the new equilibrium price using demand and cost estimates. The calculated difference in welfare is 0.05% of the expenditure, implying a small decrease in consumer welfare when removing organic certification and allowing firms to switch from more costly organic production to conventional production. This number is very small because the reduced production costs lead to lower new equilibrium prices, which then give consumers higher expected utility under counterfactual (3).

According to an industry report by Bizzozero [2020], the total sales of organic eggs amounted to $858 million in 2018. Using the estimate of 0.05%, the welfare loss in the organic egg market alone due to the removal of the certification program is worth $2.9 million. Considering the cost side, the USDA allocated a budget of $8.1 million to maintain the National Organic Program according to the 2018 USDA budget summary [USDA, 2018]. Therefore, the result from the *new equilibrium with cost reduction* counterfactual shows that the additional welfare from organic certification generated by eggs alone can cover more than 1/3 of the maintenance cost of the certification program for all agricultural products.

In addition, the certification program may affect consumer welfare through the change in choice sets. According to the bottom row of table 4.4, the calculated product-level gross profit drops by 35% compared with the *equilibrium with certification*, potentially inhibiting the entry of products from *organic collection* if the organic fixed costs are high enough. As a result, the reduced choice sets are likely to harm consumer welfare.

Moreover, the estimates of welfare change do not consider the social costs imposed by the potential adverse effects of conventional agriculture. Though this paper cannot quantify the externalities generated by organic practices on top of what consumers care about, it provides evidence that consumers' WTP for certified products facilitates the profitable entry of products with otherwise unverifiable attributes. From a social perspective, the existence of the certification program is worthwhile simply by introducing those costly-but-beneficial production processes without harming consumer welfare when government mandates are not an option.

Chapter 5

Conclusion

This paper shows that consumers exhibit positive WTP for organic products whose quality is credibly disclosed by the USDA Organic certification. This result supports the argument that when government mandates are not feasible, certification can be an alternative way to enforce the standards required for the "planet-saving" attributes. In particular, consumers' positive WTP for the certification may enable the profitable entry of products with these attributes.

Moreover, the organic certification program affects consumer welfare in two ways. First, it makes previously imperceptible product attributes "observable" to consumers by inspecting the production process and making credible disclosure. Given the estimated positive WTP for the organic certification, consumers gain direct utility from knowing some products are organic. From a counterfactual analysis of the new equilibrium without the certification program, I find a 0.05% welfare loss, measured in percentage of observed expenditure, evaluated at the new equilibrium prices and reduced costs from shirking the organic standards. Second, the organic certification program may benefit consumers through a potential increase in product selection. The equilibrium result from the counterfactual without organic certification indicates a 35% decrease in gross profit, which leads to smaller product variety when the barrier of entry is high and negatively impacts consumer welfare. Moreover, the estimated welfare effects measure only the benefits directly received by consumers from the certification program. They do not take into consideration the positive externality generated by production processes that improve the environment, treat animals humanely, and promote sustainability. Hence, I interpret the results as a lower bound for the social value of the certification program.

The certification provides a solution to the asymmetric information problem between consumers and producers and potentially enables more entry of products that possess socially desired but costly-to-produce attributes. While this paper focuses on the USDA Organic certification, the implications extend to broader questions and may inspire further research. In particular, it is worth studying the welfare effects of certification in fields related to other costly but beneficial practices, such as the MSC (Marine Stewardship Council) label for sustainable fishing, carbon neutral certification for the 1.5 degrees Celsius goal of future warming, and fair-trade certification for ensuring farmers from developing areas a living wage. It would be of great interest to learn how consumers' taste or care for the "planet-saving" practices can shape the composition of products and firms in the market and contribute to overall environmental friendliness and sustainability in production.

References

George A. Akerlof. The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3):488–500, 1970. ISSN 00335533, 15314650.

Marvin T Batte, Neal H Hooker, Timothy C Haab, and Jeremy Beaverson. Putting their money where their mouths are: Consumer willingness to pay for multi-ingredient, processed organic food products. *Food policy*, 32(2):145–159, 2007.

Steven Berry, James Levinsohn, and Ariel Pakes. Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, pages 841–890, 1995.

Steven T Berry. Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, pages 242–262, 1994.

Judie Bizzozero. Us organic sales ring up record $52.5 billion in 2018, Feb 2020. URL https://www.foodbeverageinsider.com/ingredients/us-organic-sales-ring-record-525-billion-2018#:~:text=Dairy%20and%20egg%20sales%20were,up%209.3%20percent%20from%202017.

Andy S Choi and Brent W Ritchie. Willingness to pay for flying carbon neutral in australia: an exploratory study of offsetter profiles. *Journal of Sustainable Tourism*, 22(8):1236–1256, 2014.

Randi Dalgaard, Niels Halberg, Ib S Kristensen, and Inger Larsen. Modelling representative and coherent danish farm types based on farm accountancy data for use in environmental assessments. *Agriculture, Ecosystems & Environment*, 117 (4):223–237, 2006.

Patrick De Pelsmacker, Liesbeth Driesen, and Glenn Rayp. Do consumers care about ethics? willingness to pay for fair-trade coffee. *Journal of consumer affairs*, 39(2): 363–385, 2005.

David Dranove and Ginger Zhe Jin. Quality disclosure and certification: Theory and practice. *Journal of economic literature*, 48(4):935–63, 2010.

David Dranove, Daniel Kessler, Mark McClellan, and Mark Satterthwaite. Is more information better? the effects of "report cards" on health care providers. *Journal of political Economy*, 111(3):555–588, 2003.

R-A Gerhardt. A comparative analysis of the effects of organic and conventional farming systems on soil structure. *Biological Agriculture & Horticulture*, 14(2): 139–157, 1997.

Austan Goolsbee and Amil Petrin. The consumer gains from direct broadcast satellites and the competition with cable tv. *Econometrica*, 72(2):351–381, 2004.

Justine S Hastings and Jeffrey M Weinstein. Information, school choice, and academic achievement: Evidence from two experiments. *The Quarterly journal of economics*, 123(4):1373–1414, 2008.

Jerry Hausman. Contingent valuation: from dubious to hopeless. *Journal of economic perspectives*, 26(4):43–56, 2012.

Ginger Zhe Jin and Phillip Leslie. The effect of information on product quality: Evidence from restaurant hygiene grade cards. *The Quarterly Journal of Economics*, 118(2):409–451, 2003.

Ginger Zhe Jin and Alan T Sorensen. Information and consumer choice: the value of publicized health plan ratings. *Journal of health economics*, 25(2):248–275, 2006.

Sia Bik Kai, Ooi Bee Chen, Chong Shyue Chuan, Lim Chee Seong, and Low Lock Teng Kevin. Determinants of willingness to pay of organic products. *Middle-East Journal of Scientific Research*, 14(9):1171–1179, 2013.

Athanasios Krystallis and George Chryssohoidis. Consumers' willingness to pay for organic food: Factors that affect it and variation per organic product type. *British food journal*, 2005.

Donald W Lotter. Organic agriculture. *Journal of sustainable agriculture*, 21(4): 59–128, 2003.

Maria L Loureiro, Jill J McCluskey, and Ron C Mittelhammer. Will consumers pay a premium for eco-labeled apples? *Journal of Consumer Affairs*, 36(2):203–219, 2002.

John P Reganold, Lloyd F Elliott, and Yvonne L Unger. Long-term effects of organic and conventional farming on soil erosion. *Nature*, 330(6146):370–372, 1987.

Imke Reimers and Joel Waldfogel. Digitization and pre-purchase information: the causal and welfare impacts of reviews and crowd ratings. *American Economic Review*, 111(6):1944–71, 2021.

Elsa Rodríguez, Victoria Lacaze, Beatriz Lupín, et al. Willingness to pay for organic food in argentina: Evidence from a consumer survey. *International marketing and trade of quality food products*, 297, 2007.

Maryam Saeedi. Reputation and adverse selection: theory and evidence from ebay. *The RAND Journal of Economics*, 50(4):822–853, 2019.

Yaowarat Sriwaranun, Christopher Gan, Minsoo Lee, and David A Cohen. Consumers' willingness to pay for organic products in thailand. *International Journal of Social Economics*, 2015.

The USDA Economic Research Service. New products, 2021. URL https://www.ers.usda.gov/topics/food-markets-prices/processing-marketing/new-products/.

USDA. 2018 usda budget summary, 2018. URL https://www.usda.gov/sites/default/files/documents/fy18-budget-summary.pdf.

Jenny Van Doorn and Peter C Verhoef. Willingness to pay for organic products: Differences between virtue and vice foods. *International Journal of Research in Marketing*, 28(3):167–180, 2011.

Appendix A

Appendix to Chapter 2

Figure A.1: Correlation: Price Premium and Predictability for Organic Certification by Product Type

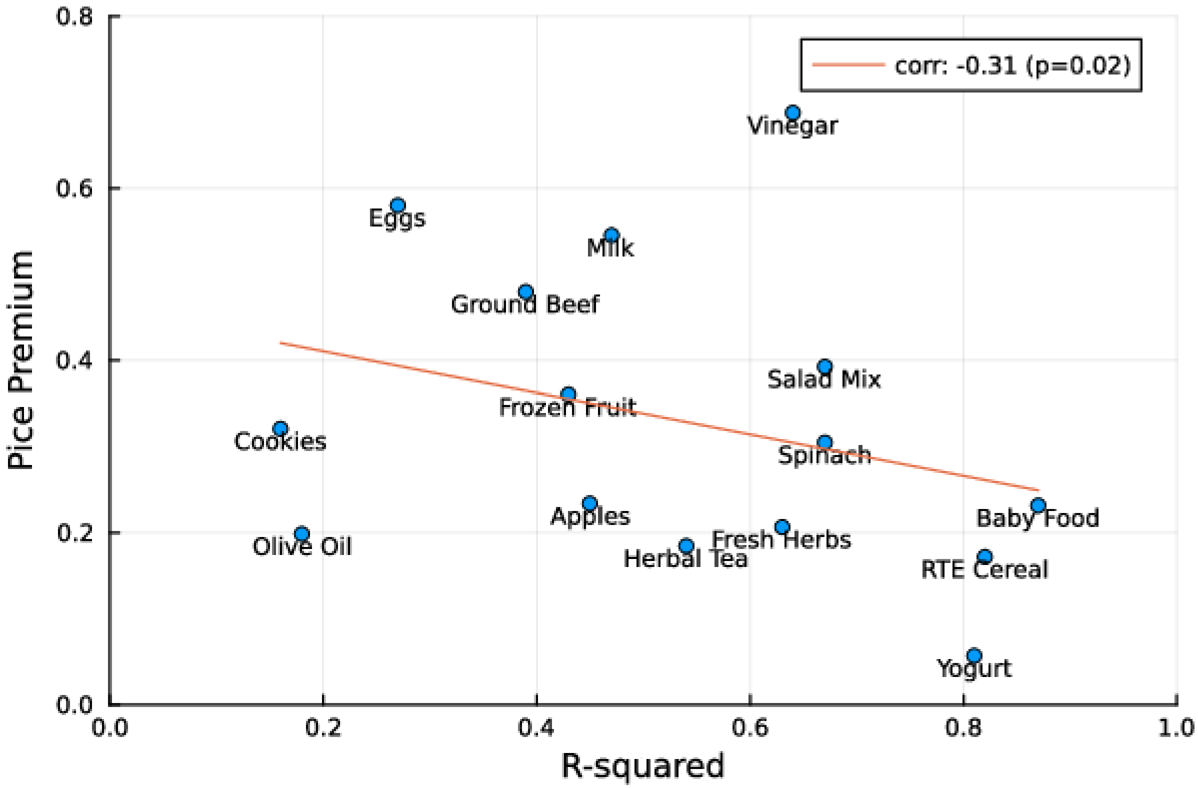


Table A.1: Product Characteristics Variables

Name	Variables
Vinegar	Balsam, mild, distill, cider, import, wine, aged, rice, pickling, glass container, and package size in ounces.
Olive Oil	Glass container, tin container, extra virgin, light, pure, and package size in ounces.
Baby food	Jar container, tub container, fruit, vegetable, dinner, apple, blueberry, mango, banana, beef, carrot, chicken, sweet potato, green bean, ham, peach, pear, prune, squash, turkey, rice, non-shelf-stable, and package size in ounces.
RTE cereal	Cinnamon, nuts, fruit, vanilla, chocolate, honey, original, and package size in ounces.
Cookies	Almond, chocolate chip, oat, peanut butter, sugar, vanilla, duplex, fig bar, sandwich, sugar wafer, wafer, chocolate enrobed, other enrobed, and package size in ounces.
Herbal tea	Caffeinated, cinnamon, apple, ginger, honey, lemon, mint, orange, passion flower, peach, sweet, berry, vanilla, and package size in bags.
Frozen fruit	Mixed fruit, mango, blueberry, blackberry, raspberry, strawberry, cherry, berry blend, peach, pineapple, triple mix, and package size in ounces.
Salad mix	Chopped, cut, shredded, baby leaves, slaw, Caesar, kale, lettuce, spring mix, spinach, and package size in ounces.
Ground beef	Extra lean, lean, chuck, sirloin, round, and package size in ounces.
Yogurt	Non-dairy, Greek, natural, nonfat, low fat, whole milk, lactose-free, blueberry, blackberry, vanilla, strawberry, cherry, mixed berries, orange, key lime, high protein, peach, pain, raspberry, banana, cream pie, chocolate, lemon, pineapple, and package size in ounces.
Milk	Carton package, low fat, skim/nonfat, whole milk, and package size in ounces.
Apple	Extra fancy, and package size in ounces.
Fresh herb	Rosemary, mint, thyme, parsley, cilantro, chive, basil, sage, dill, bunch, and package size in ounces.
Egg	Grade AA, brown, size (medium, extra large, jumbo), and package size.
Spinach	Baby spinach, whole spinach, and package size in ounces.

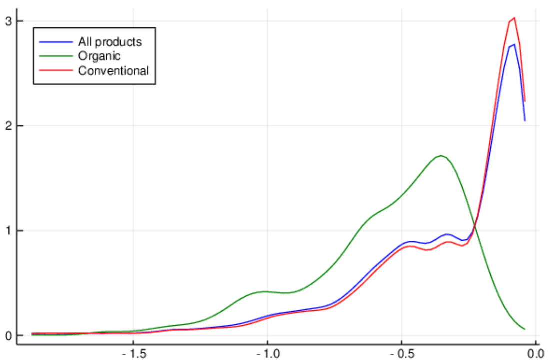
Table A.2: Examples of Occupation by Group

Group	Examples of occupations
1	Scientist, accountant, doctor, lawyer, banker, etc.
2	Baker, carpenter, truck driver, construction worker, etc.
3	Salesman in wholesale, real estate, insurance, retailer, etc.

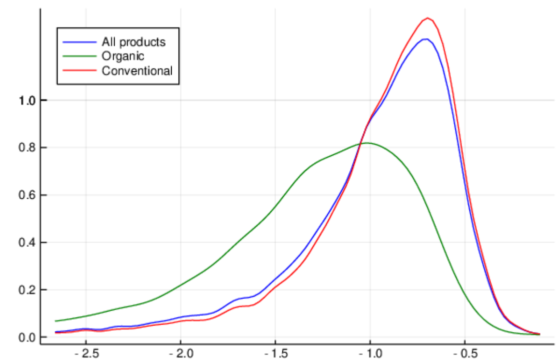
Appendix B

Appendix to Chapter 3

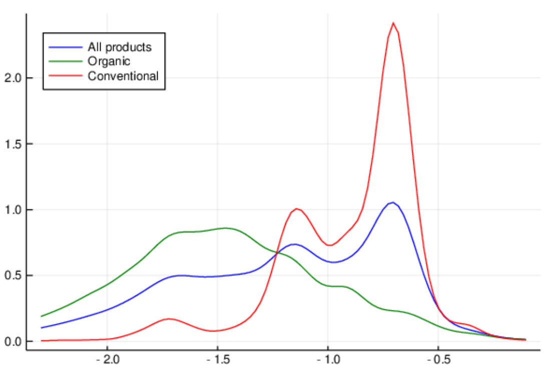
Figure B.1: Distribution of Own-Price Elasticity by Product Type (1)



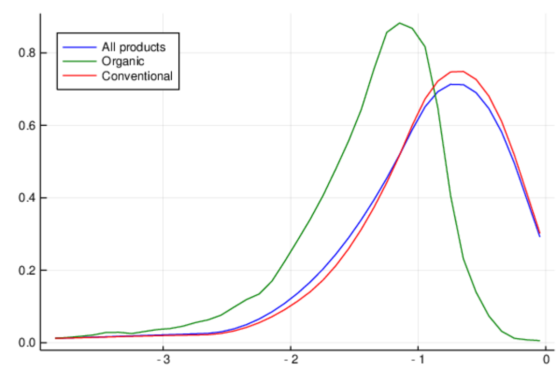
(a) Vinegar



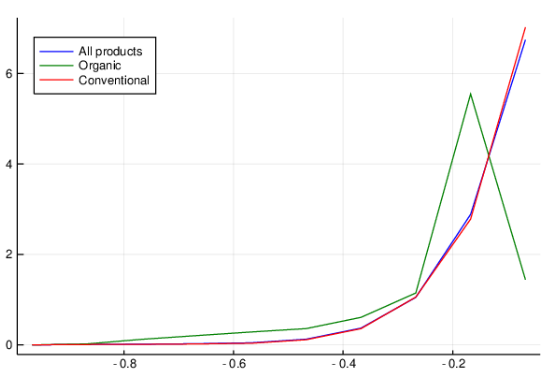
(b) Olive Oil



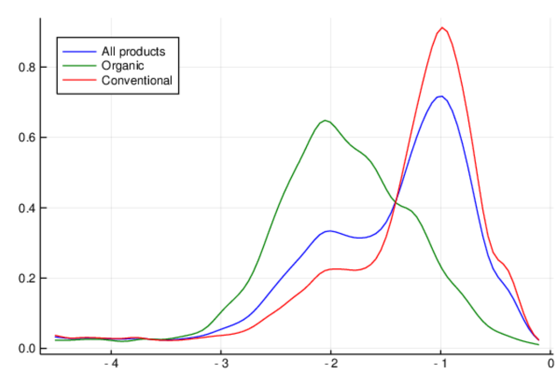
(c) Baby Food



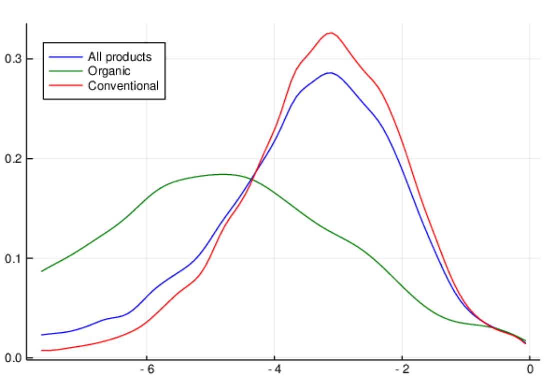
(d) RTE Cereal



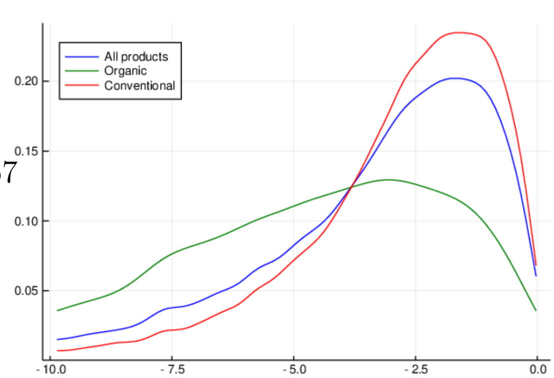
(e) Cookies



(f) Herbal Tea

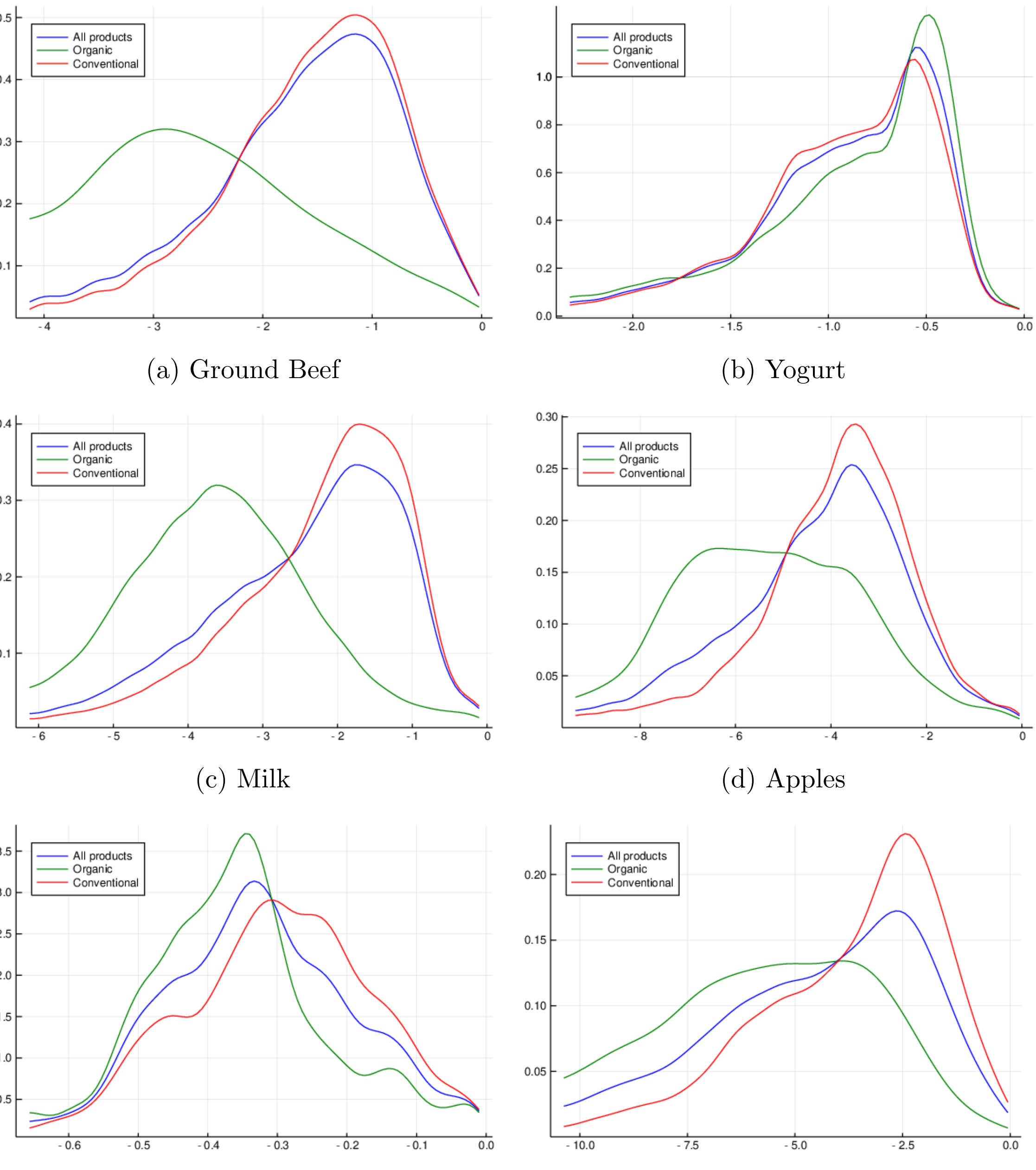


(g) Frozen Fruit



(h) Salad Mix

Figure B.2: Distribution of Own-Price Elasticity by Product Type (2)



(a) Ground Beef

(b) Yogurt

(c) Milk

(d) Apples

(e) Fresh Herbs

(f) Spinach

Figure B.3: Correlation: WTP and Predictability for Organic Certification by Product Type