Mandatory Labeling: Changes in Consumer and Producer Behavior

By

Henri Rautonen

Dissertation

Submitted to the Faculty of the Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Law and Economics

May, 2016

Nashville, Tennessee

Approved:

Paige Skiba, Ph.D.

W. Kip Viscusi, Ph.D.

Erin O’Hara O’Connor, J.D.

Kelly Haws, Ph.D.

To my parents, without whom I never would have started this,

To my wife, without whom I never would have finished it,

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

FDA: Food and Drug Administration

FTC: Federal Trade Commission

GMO: Genetically Modified Organism

LfL: Like-for-Like

NOP: National Organic Program

rBGH: Recombinant Bovine Growth Hormone

TM: Total Market

USDA: United States Department of Agriculture

WTP: Willingness-to-Pay

Chapter 1: The Preference Shaping Power of Labels: GMO Labeling and the Prevalence Effect

Abstract

This experiment measures the change in valuations, knowledge, and opinions of genetically modified (GMO) food products (potato chips) resulting from different labeling regimes. In the control group 20% of products are labeled as non-GMO, and other products left blank to reflect the current market in the United States. The treatment consists of labeling the other 80% of products as containing GMOs, reflecting the effect of a hypothetical passage of mandatory labeling laws of GMOs in the United States. It is hypothesized that the mere increase in the prevalence of the label will increase consumer willingness to pay (WTP), manifesting as an increase in the percentage of participants in the treatment group opting for the non-GMO product offered across different price premiums. The results show a clear effect on valuations from the labeling change, without a commensurate change in beliefs about GMO products (which are also measured), or knowledge of their prevalence. The average probability increase for choosing the non-GMO alternative was 45% across different price points, across a range from 5% to 55% price premiums being offered.

Introduction

Consumers respond to the information presented to them, and often this happens at the point of sale.[1] This is why it is not surprising to see bitter fights over new labeling laws, whether in tobacco[2], residential mortgages[3], or food products.[4] Both consumer advocacy groups and producers understand the power labels have to shape the market beyond simply acting as a low-cost channel for distributing more or better information to the end-users of products. In stark contrast to this reality in the marketplace, studies on the effects of such labels are often conducted in laboratory conditions with little resemblance to the actual setting in which consumers make such decisions. A

[1] See, e.g., GMA Shopper Marketing Report (2007), *available at* https://www.gmaonline.org/downloads/research-and-reports/shoppermarketing.pdf
[2] See, e.g. Hammond (2011).
[3] See Ben-Shahar (2011).
[4] *Id.*

key ingredient missing from most studies is the prevalence of a label in the broader marketplace: in a real consumer's shopping context organic or non-GMO labels are only one of the things vying for attention, competing with brands, price promotions, flavor options, and colorful packaging. Even within the narrower scope of just labels, there are health labels ("no salt added"), and alternative environmental labels ("natural", or "organic"). The effect of all these factors is probably a reduction in the attention paid to any one attribute in particular, certainly less so than in a laboratory study of GMO products explicitly asking participants again and again to think about the issue of GMOs. Thus what is missing from the discussion of labels' effects on consumers is the role played by the prevalence of the label, and how that can drive valuations.[5]

People recognize labels, but are often unable to articulate the underlying characteristics represented by those very labels. [6] "Organic" to many consumers might be an unarticulated brand-marker like "Nike" or "Apple", rather than primarily an informational tool about specific characteristics. Several labels are disproportionately popular in a narrow product category, but the characteristics they represent, while present in other areas, are not in the consumers' minds. For example, the clout of Fair Trade products in several countries is heavily limited to its coffee and banana product categories, while recombinant bovine somatotropin (rBGH)-free milk, but not rBGH-free cheese or sour cream, is considered an established product category.

---

[5] Note that this study only considers the particular case of GMO labels, where the existing market results in a 20% labeled 80% unlabeled mix, which would change to 100% labeled if mandatory labeling laws were passed. The 80% prevalence rate of GMOs is cited by both proponents and opponents of GMO use, *see* https://factsaboutgmos.org/disclosure-statement as well as http://www.nongmoproject.org/learn-more/ . Further work is needed to establish the exact contours of this prevalence effect for other mixes like 50% labeled/50% unlabeled, or 60%/40%.

[6] A report by the Natural Marketing Institute from 2013 found that only about a third of consumers recognize and *understand* the USDA Organic label. Notably, organic labels are some of the more uniformly regulated (all centralized under the USDA), as well as older labels (>10 years on the market). *See, e.g.*, http://newhope360.com/nfm-market-overview/organic-continues-double-digit-gains?page=2

Rather than coming to the market with a pre-determined list of product characteristics they seek, and the relative valuation for each of these, consumers likely rely on a shortlist of key variables, and can easily take their cue from the advertisers and producers. This possibility is easily demonstrated by a simple example of new products: how do consumers assess products to which they have no reference point or easy comparison, such as the first wave of personal computers, or digital cameras? Megapixels became a heavily advertised aspect of digital cameras, which arguably resulted in them becoming a big part of the consumer's purchasing decision, causing in turn even more producers to start emphasizing this feature.[7]

One can see how similarly consumers have been primed to think about Fair Trade when buying coffee (e.g. 25% of roasted coffee by volume in the UK carries the fair trade label)[8], thus fair trade becomes one of the metrics used to assess this particular product, but fair trade issues are not similarly salient in the market for tea. It is unlikely that consumers simply care less about the potentially exploitative markets facing tea growers than they do for coffee. Nor does it seem plausible that they believe that the prevalence of fair trade tea is so great that no label is needed in that submarket.

Instead of simply informing consumers about underlying characteristics they already care about, labels have the potential to shape these preferences and define the very metrics used by consumers to assess the product. This implies that the prevalence of the label matters beyond the

---

[7] Several years ago the New York Times ran a story on the deceptiveness of megapixels as a shorthand metric for quality. This did not change the market and five years later there were still articles in the news trying to make the same point. See, http://www.nytimes.com/2007/02/08/technology/08pogue.html?pagewanted=all&_r=0 (2007 NYT article), http://gizmodo.com/5888552/reminder-megapixels-dont-matter (2012 Gizmodo article on the same basic issue).

[8] See Fair Trade International Report (2012) http://www.fairtrade.net/fileadmin/user_upload/content/2009/resources/2012_Fairtrade_and_coffee_Briefing.pdf

simple effect of how many consumers are exposed to the information. Labels that are present in a large percentage of the goods offered in a particular category have the power to increase valuations of the labeled characteristic, shaping consumer preferences rather than just fixing a problem of asymmetric information. The classical economist's justification for labeling is a market failure due to asymmetric information (producers know more about the goods than purchasers) and/or costs of acquiring that information as a form of a collective action problem (few consumers care enough about one characteristic in one of the many products they buy to put in the effort to research the issue, but taken as a whole the consumers' collective concern would justify requiring producers to make the information more easily available).

Mirroring this limited view of the important role of category market-share for labels, most studies in economics have focused on the ex-post effects of labels on the marketplace, or the ex-ante estimation of such labels in laboratory conditions. This experiment sets out to show the pitfalls present in both types of studies: shifting preferences ex-post and a lack of sensitivity to the role of salience exhibited in ex-ante studies so far.

A label can drive market change through three different channels: first, through more information to the consumer about the characteristic itself (e.g. "causes cancer 50% of the time"). Second, through more information about the presence and/or prevalence of the characteristics in products (i.e. consumers were already worried about chemical X, but mistakenly thought only 5% of goods have it, while in reality it is 95%). Or third, a shift of preferences unrelated to such new information. The latter channel presents a problem to conventional economic models: if we cannot attach the change in preferences to a change in the first two channels, what exactly has the label done, and what set of consumer preferences should we use to evaluate it, the pre- or post-label?

Traditional models consider the role of information in a consumer's purchasing decision in a vacuum: the attribute Y is factored into the consumer's utility function if they are aware of its presence, and treated as zero otherwise.[9] Contrary to such simplistic models, empirical evidence suggests consumers' utility functions will depend on various aspects of *how* the information is presented.[10] Thus consumers may either ignore or overweigh the information presented, depending on various factors, both internal (such as their Bayesian proclivity to update beliefs in the face of new information) or external (such as the label's type and size). These phenomena still reflect what is at heart an information-driven consumer, simply one that is impaired in their ability to obtain that information under some conditions.

Going beyond this framework, consumers may also draw inferences about the fact that particular information is highlighted: one French study (Nouissar, Robin & Ruffieux 2001) found no initial change in WTP for a genetically modified chocolate bar (<2%), but upon instructor emphasis on the genetically modified attribute the average willingness to pay was reduced by more than 25%. Consumer preferences about the characteristic were malleable, with no change in underlying information. The authors simply assert that the issue was the participants' failure to read the label without the experimenter's additional emphasis in the mix, and do not explore the issue further. However, the result is potentially an exhibit of the issue present in most studies on labeling effects: the prevalence and attention drawn to the label is not explicitly controlled for, nor calibrated to reflect the resulting market under the rule being tested. As discussed above, this can result in valuations that would not occur under market conditions, severely limiting the external validity of such studies.

---

[9] Alternatively, a sophisticated consumer faced with a market with imperfect information may choose to substitute the average value of X in the market when not presented with a label. Such behavior would be consistent with e.g. Akerlof's "Market for Lemons" (1970)

[10] This entails aspects such as the structure, format and context of the labels. (*See* Viscusi& Magat 1987). *See also* Noussair et al. (2002)

This experiment sets out to test the effect the prevalence of a label can have on consumer decision-making using genetically modified organisms (GMOs) as an example. Calibrating the prevalence of the label to underlying market realities (the 20/80 mixture of labeled products discussed above)[11] allows us to more accurately identify the effect of mandatory labeling on consumer behavior. Further, by gauging participants' information about the qualities and prevalence of the GMO characteristics, it sets out to better understand which of the three channels identified is driving the results. The experimental setting allows us to draw from a richer set of data than available in most empirical studies, to further understand how the labeling effect might vary depending on other consumer characteristics, as well as allowing for predictions better tailored to the specific market and characteristic addressed.

LITERATURE REVIEW

The potential effects of labels on consumer goods have been studied extensively. The literature up to this point can be divided into two categories. First are empirical studies using market data that by definition take into account the prevalence of a label, but are unable to isolate this factor and thus estimate the specific role of prevalence. Second, there is a large body of experimental work estimating willingness-to-pay that *could* have controlled for the prevalence of the label, but were not designed to do so. Thus the valuations obtained in these studies are pegged to various implicit prevalence levels (typically high).

[11] See footnote 5

For environmental labels, Teisl et al. (2002) find a market shift caused by dolphin-safe labels using longitudinal data.[12] Survey examples include Loureiro et al. (2001) who find a strong premium for organic labels. Notably the survey design seems to present equal-weighed no label, eco-label and organic label alternatives.[13] Loureiro et al. (2002) considers the GMO label specifically for potatoes, asking consumers "what premium $x would you be willing to pay for characteristic y?". Again due to the structure of the survey, GMO choices reflected 33% of the available products.[14] In an incentive-compatible study, Lusk et al (2001) finds that in an explicitly GMO centered auction, 70% of participants are unwilling to pay *any* premium for GMO products, however the survey elicits the bids as a premium to exchange their given product. This will potentially underestimate the true WTP due to an endowment effect (participants will value goods higher if they are theirs, see Thaler 1980). The experimental design means that *all* rounds involved a GMO choice (the default given to participants).

Noussair et al (2002) find a 30% WTA for GMO products when participants are explicitly told to focus on the characteristic.[15] All relevant rounds included the GMO choice, however only one out of four explicitly drew attention to the characteristic.

In another incentivized study, Huffman et al. (2003) find a 14% WTP for non-GMO products, using an experimental design with 50% of choices identified as containing GMOs (one in each pair of products). Further, the experiment provides significant information about GMOs to

---

[12] Note that this data represents real market behavior of consumers over time, and thus is calibrated to the prevalence of such labels.

[13] An additional issue is the lack of incentives, as the participants were merely asked about their valuations (stated preference) without actually spending any money. The authors suggest running their survey in a grocery store "close" to actual purchasing decisions mitigates this problem.

[14] Organic, GMO and Colorado grown were the three attributes offered, consumers only answered each type of question once. As with the previous study, consumers did not have to actually spend money according to their stated valuations.

[15] Participants are shown a large projected slide with the GMO characteristic highlighted in large font and explicitly discussed by the experimenters.

each group: all six treatment groups are provided with several paragraphs worth of information to read on GMOs. Both choices will result in a higher than current market salience of GMO products in the valuations.

Kaneko & Chern (2005) find a roughly 25% average WTP using a dichotomous choice survey approach. However, here as well the frequency of GMO choices is set at 50% (every pair offers a GMO and non-GMO alternative). Notably, this survey contains some measures for knowledge and attitudes about GMOs in addition to the main WTP measure.

Gifford et al. (2005) found that the self-reported WTP premia for non-GMO products (avg. 4%) were significantly lower than the realized premia in the experiment (avg. 20-30%).[16] The authors find no evidence of a gender effect, or self-reported knowledge about GMO products and attitudes towards them.

This study is the first and only one known to the author beyond the present work that includes a measure related to the prevalence of GMOs in the form of a "GMO content chance", where participants are told there is a x% chance a product may contain GMOs (this results in an asymmetric information dynamic discussed in Section III under "Market for Lemons"). The content chance is not found to systematically influence WTP in their sample, and the authors do not discuss the variable further. However, the authors rely on self-reported estimates of prevalence, rather than actually exposing participants to various prevalence levels.

[16] Note that his goes against conventional models of "cheap talk", where self reported valuations are significantly larger than actual market behavior. If Gifford et al. accurately identify a wider phenomenon, this suggests a significant portion of the work done on GMO labels might be *underestimating* the true market premiums.

MODEL

This paper sets out to test through an experiment whether labeling prevalence affects consumer decision-making. I hypothesize that the treatment condition with more GMO labels presented will result in higher valuations for non-GMO goods. This hypothesis relies primarily on a salience model by Bordalo et al. (2013), but does not require the validity of the behavioral assumptions contained therein. Using only classical economics, the "Market for Lemons" (Akerlof 1970) model with asymmetric information produces comparable predictions.[17]

<u>A Salience Model of Consumer Behavior</u>

What determines the attention given by a consumer to any particular product attribute? Salience theory[18] proposes that qualities that differ the most from the "average" of comparable products are weighed relatively more than those closer to the average: in other words, the most clearly differentiated feature exerts an oversized influence on the consumer's decision. An emerging literature in economics has formally modeled salience in the consumer goods context. Bordalo et al. (2013) constructs a salience function σ(p,q) where p is the price of the good, and q the "quality" (non-price attributes). Values of (p,q) for a good examined are compared against a reference point (p,q) representing the market average n ($\sum p_i / n$) for that attribute. The salience function σ

[17] See Chapter 3 of this dissertation for a detailed discussion of Akerlof's model. In the original paper the asymmetric information causes a total collapse in the market for higher quality goods (in our case: non-GMO products), where an intermediate stage would be a lower average consumer WTP for the characteristic (lack of GMOs) which would be captured in this model. The treatment condition in our experiment solves the asymmetric information problem by introducing complete information into the markets: all products are accurately identified as high quality (non-GMO), or "lemons" (containing GMOs), allowing for the "true" WTP for GMO products to emerge with the removal of uncertainty (all probabilities are 1 after observing the products).

[18] See Bordalo et al. (2013)

overweighs departures further *i*=1 away from that average whether in price or quality. The function σ is ordered: if x'<x<y<y', then σ(x',y')>σ(x,y). Further, σ is homogeneous in degree zero, i.e. σ(αx,αy)=σ(x,y). Thus the consumers' perceived utility from a good can be modeled as:

$$u(q_{\mathrm{i}} - p_{\mathrm{i}}) = \begin{cases} q_{\mathrm{i}} - \delta p_{\mathrm{i}} & if\ \sigma(q_{\mathrm{i}}, \bar{q}) > \sigma(p_{\mathrm{i}}, \bar{p}) \\ \delta q_{\mathrm{i}} - p_{\mathrm{i}} & if\ \sigma(q_{\mathrm{i}}, \bar{q}) < \sigma(p_{\mathrm{i}}, \bar{p}) \\ q_{\mathrm{i}} - p_{\mathrm{i}} & if\ \sigma(q_{\mathrm{i}}, \bar{q}) = \sigma(p_{\mathrm{i}}, \bar{p}) \end{cases}$$

Where δ is the salience distortion (δ=1 implies a rational consumer, 0<δ<1 models the degree of distortion, where δ=0 has salience overpower the other attribute completely.)

The model developed by Bordalo et al. does not discuss the reference level formation further. This paper will suggest two extensions to the model: the formation of the reference level *q* for any particular q based on beliefs (γ), and the extension of q into a **N**x**1** dimensional matrix, with each q_i assigned a weighing factor λ_i.

<u>Formation of a reference point *q*</u>

Instead of an objective level of quality q, consumers' *perception* of quality for a product (γ) is determinative of their decision. Thus the salience function σ operates in a price-belief space where some types of levels of q can be systematically under or overvalued. Further, rather than relying on an arithmetic average of all true q_i in the relevant market, this paper suggests that the consumer relies on the same *perceived* level of the quality (γ) when constructing the reference level for quality.

Thus, a distortion in the salience function σ can be caused even in instances where the individual good's quality is accurately perceived (γ_i=q_i). In the special case with complete information and rational beliefs γ_i=q_i for all i, and thus $q = \gamma$. However, if for example q is the the risk of an allergic reaction to a product, and the potential allergen is not identified on the product (i.e. the consumer does not observe the presence of the risk), γ may be zero, or anchored to a reference point based on a broader set (e.g. all foods, rather foods containing peanuts) q'.

In the context of GMO labeling, let us treat the presence of GMOs as a binary quality where zero indicates no GMO presence, and one the use of GMOs. Having only 20% of products identify the *lack* of GMOs suggests the consumers perceived quality γ_i for an unlabeled product containing GMOs can either be zero (mistakenly believing unlabeled products are non-GMO), or the market average of 0.8 (where the consumer believes the product can either be an unlabeled instance of a non-GMO product (0.2 x 0), or indicative of GMO presence (0.8 x 1) . In either case the perceived quality γ is different from the true quality q, causing a distortion in the salience function. Note that this context is also an example of Akerlof's market for lemons (Akerlof 1970)[19], and thus does not require acceptance of the broader behavioral framework introduced here – in other words even classical economic theory can accommodate a labeling prevalence effect.

Further, assuming the consumer is able to accurately gauge the quality of a particular product k (e.g. one with a prominent "non-GMO" label), the labeling regime can still distort the salience function through the reference level of γ. Although γ_i=q_i in this case, for all i, the quality q can be misperceived, yielding a reference level of γ that may be greater or smaller than the true average. Again, assuming a labeling regime with 20% of goods identified as non-GMO (q=0, γ=0), the consumer may in some, or all, cases treat the lack of a label equivalent to lack of GMOs (q=1, γ=0) resulting in $\gamma < q$ =0.8. Thus the salience σ of a labeled non-GMO product may

[19] See footnote 17 above.

emphasize the price differential between the labeled and unlabeled goods, even though the "true" salience would be in differences of quality.

Assume $\sigma(q_k, q) > \sigma(p_k, p)$, thus $u(q_k,-p_k) = \delta q_k - p_k$, i.e. quality is the salient feature for a fully informed consumer. Further assume that $u(q_k,-p_k)>u(q_i,-p_i)$ for all i=k, so that the fully informed consumer would always choose product k. Then if there is a misperception in the reference level such that $\gamma < q$, it is possible that $\sigma(\gamma_k, \gamma) < \sigma(p_k$, and $u(\gamma_k,-p_k)<u(q_k,-p_k)$ if $\delta<1$.

Thus the chosen labeling regime can influence the salience of quality over price, and switch the market from a quality competition into price competition. In the case of GMO products, labeling only non-GMO products can depress the perception of individual product "risks" (i.e. quality) for unlabeled GMO goods, as well as the reference level of risk in the market.[20] Both phenomena result in weakened demand for the quality "superior" but pricier non-GMO product.

Conversely, merely switching the labeling regime to require labeling of all products containing GMOs (80% of the market) can lead to switching the market to quality-salience, and more demand for the non-GMO products. As the market reference point γ moves closer to the true prevalence q, non-GMO attributes appear to deviate more from the reference level than in the previous labeling regime, causing the "quality" aspect to receive proportionately more weight in the consumer's decision.

Matrix of product attributes weighed in decision

When evaluating a product, the consumer must choose which attributes (q) to consider in addition

[20] Compared to a fully labeled market. Note that even labeling just the non-GMO products (20%) elevates risk beliefs compared to a completely unlabeled market, where the issue of GMOs is not introduced at all.

to price. Bordalo et al. (2013) collapse all non-price characteristics into q, but as this paper will argue, the selection of attributes and their relative weighings are not fixed across time or product categories. Rather than simplifying the product into a two dimensional bundle of (price, other), the non-price qualities can be seen as a **N**x**1** size matrix Q, with each row weighed by a factor λ_i.. The weighing matrix of attributes is not unlike the concept of positioning in marketing: a product is presented to the consumer as a bundle of its most desirable features, attempting to avoid unfavorable comparisons to potential substitutes. With the attribute-matrix, some attributes will be particularly important to the consumer, and the weighings λ_i will be different across products as well as consumers.

Using this model, rather than influencing the reference level directly, the prevalence of a label will affect the attribute's weight λ_i. Even within the narrower subset of attributes relating to product safety and/or environmental characteristics, the consumer can potentially focus on the product's carbon footprint, ethical sourcing, use of chemicals or additives, etc. The weight given to each "risk" can depend on the specific good even within a small subset of products: for example, the use of growth hormones is an important attribute to many consumers when buying liquid milk, but less so in yoghurt or cheese products.[21] The weighing matrix λ then is a pre- determinant of the salience function in the second stage of the consumer's decision- making process. This paper will posit that the attribute's prevalence (whether through labeling, news coverage, frequency and directness of exposure) influences the consumers' a priori weighings, such that λ_I becomes greater the more exposure it has.

The importance of the attribute-matrix is in its ability to explain changes in behavior even

[21] Due to the relative prevalence of labels in each product category.

for fully informed consumers. Building on the GMO example used in the reference level formation section, the attribute-matrix implies that even if δ=1 and $\gamma = q$ (i.e. consumer suffers from no salience bias, and forms an accurate perception of the reference level), the prevalence of labeling can affect choice through λ_i, i.e. which product attributes are considered in the first place. In the case of twenty percent of products labeled non-GMO, the relative prevalence of the "GMO" attribute is relatively smaller than qualities such as price, but also nutritional values, country of origin etc.

DATA AND EXPERIMENT DESIGN & PROCEDURES

To test whether salience affects consumer decision-making through labeling, I conducted an incentivized experiment varying the frequency of GMO labels presented to participants. The experiment was run in a computer laboratory on campus at Vanderbilt University over five consecutive weekdays in October 2015 with participants drawn primarily from the undergraduate population. Participants were recruited by flyers, as well as speaking in classes about the possibility to participate. Compensation was set at a $12 participation award, minus the price of a snack chosen. Participants were unaware of the compensation specifics or the specific purpose of the experiment[22] in an attempt to avoid self-selection into the study. Participants were randomly assigned into the control or treatment group by order of entry.[23] The average time to complete the study was around 10 minutes, with most participants taking between 7-12 minutes.[24] Participants were told one of their snack choices would be chosen as their binding purchasing decision, but were unaware which choice

[22] Referred to as a "Consumer Product Labeling Study" in the recruiting materials.
[23] Every other participant receiving the link to the control study, and the other the treatment.
[24] The software used includes timestamps for the activation of the link (beginning of introduction and qualifying questions) to the final answer in the demographics section, allowing us to observe the exact time taken by participants.

this would be. Pricing variation meant that more than 15% of total compensation was potentially dependent on their choices.

The survey consisted of two practice rounds to familiarize participants with the paired-choices format of the study and ensure understanding of the choice presented. This was followed by thirty rounds of pairwise choices presented one at a time (see Appendix A, Image 1). Half of the choices were between different types of potato chips, and the other half between chocolate bars.[25] Six separate rounds presented participants with the option of choosing a non-GMO product or organic.[26] Further, one round presented participants with the option of choosing an organic product. Each of the non-GMO pairs offered was followed by rounds with no such alternative. The design is thus robust to latent learning effects, where participants slowly update their salience assessment of the non-GMO feature (proof of which would be found in an increasing effect of the treatment in latter pairings, when participants have been exposed to more of the "contains GMOs" labels).[27] At the end of these 30 rounds participants were informed of the snack choice they made that was binding, as well as their final monetary compensation and product to be received. The table below summarizes the design:

TABLE 1. SUMMARY OF NON-GMO AND ORGANIC CHOICES OFFERED

	1st Pair	**2nd Pair**	**3rd Pair**	**4th Pair**	**5th Pair**	**6th Pair**	**7th Pair**
Label	Non-GMO	Non-GMO	Non-GMO	Non-GMO	Non-GMO	Non-GMO	Organic
Price	20 cents	10 cents	40 cents	30 cents	30 cents	70 cents	10 cents
Differential	(~10%)	(~5%)	(~20%)	(~25%)	(~25%)	(~55%)	(~5%)

[25] The product mix was introduced to reduce the risk of decision fatigue arising from a repetitive task.
[26] Note: to increase our ability to observe small variations in valuation changes, all non-GMO rounds consisted of choices concerning potato chips only.
[27] The most direct test for a learning effect is between the 4th and 5th pair, as these offer the same price premium, the only intervening effect being several rounds of GMO labels.

The design was informed both by average WTPs fro non-GMO products observed in pre-existing literature (ranging from 0 to 30%), as well as pre-testing by the author to calibrate the premiums offered to create dispersion in participants' decisions. The range of magnitudes offered is sensitive enough to pick marginal movements among less sensitive participants (5% increments for first two increments above zero valuation), while also capturing shifts for the "average" consumer (with valuations estimated between 10-30%) in the 3^{rd} 4^{th} and 5^{th} pairings. Finally, the 6^{th} pair allows the experiment to capture shifts significantly above the average WTP estimated so far, controlling for a tail-end of extra-sensitive consumers, as well as the possibility of a major shift in the average consumer's valuations.

The design contains both increases (2^{nd} to 3^{rd}, 3^{rd} to 4^{th}, 5^{th} to 6^{th}) as well as decreases (1^{st} to 2^{nd}, 6^{th} to 7^{th}) in the price premiums, which combined with the "noise" of rounds in-between the pairings is designed to minimize participant anchoring on any particular premium (where subsequent offering would be seen as either "raising the price" or a "discount" compared to the original premium offered).

The next section presented to the participants was a brief four-lottery Holt-Laury (2002) measurement of risk preferences through a menu of lotteries (see Appendix A, Figure 3 for details of lottery). The section following that asked participants about their shopping habits, consumption of potato chips/chocolate bars and perceptions about the price and prevalence of non-GMO goods, familiarity with the labels, and their attitudes about mandatory labeling of GMOs.

This was followed by a three part True/False questionnaire testing participants' *objective* knowledge about GMOs.[28] Please see Figure 2 in Appendix A for a full list of the True/False questions. The questionnaire is constructed from questions used in previous literature (Hallman et al.

[28] Note that this is a significant departure from most of the literature which simply relies on a self-reported, subjective knowledge (e.g. "on a scale of 1 to 5, how knowledgeable are you about GMOs").

2003). Participants in this experiment are significantly more informed than those in the previous studies: on the tomato genes question (question #2) 57% of participants responded correctly in the Hallman sample, while in this study 96% answered correctly. Further, on the question about food causing gene modification in humans 68% of Hallman's sample answered correctly, in this study the figure is 92%. This reflects a potentially better educated sample group, as well as increased knowledge of GMOs arising from the passage of time (the Hallman study is over twelve years old, before issues surrounding GMOs became more prevalent in the news).[29] A third question was introduced from the literature on consumer understanding of organic labels[30] (question #1). Here participants only had a 25% accuracy rate, however this was still significantly higher than the population average (10% in Hartman Group survey from 2014). Participants scored an average of 1.8 questions right on the questionnaire used in this experiment, with a standard deviation of 0.82, suggesting that in spite of their superior level of knowledge compared to populations surveyed in the past, the questions successfully created enough variation amongst the sample to be useful for our analysis.

In another question participants were asked to estimate the prevalence of GMO products in packaged goods in the US market today. The true value is approximately 80%, which participants on average estimated fairly accurately.[31] Finally, participants were asked a series of demographic questions.

[29] The participants also scored significantly higher than the pre-test group, suggesting the dispersion of knowledge in the overall population has not happened at an equal pace.
[30] See e.g. "Organic & Natural 2014" survey by the Hartman Group.
[31] Both the control and treatment groups relatively accurately identify the prevalence with average estimates of 70.8% and 71.8% respectively (true value is 80%. Note that the treatment does not seem to have an effect on the perceived prevalence).

TABLE 2. SUMMARY STATISTICS OF PARTICIPANTS AND RELEVANT POPULATIONS

	Sample (n=189)	Vanderbilt Undergraduates	US National
Avg. Age	18.5	N/A	36.8
% Male	71	51	49.2
% White	59.2	66.6	77.4
% Black	7.4	9.6	13.2
% Hispanic	8.4	10.1	17.4
% Asian	25	13.7	5.4

The experiment draws from a very distinct subpopulation of the United States and is not meant to be nationally representative. Even after accounting for the expected differences of college undergraduates to the general population, the experiment seems to have oversampled male and Asian participants. Recruitment efforts were not restricted to any specific classroom or department at the university, as such these variations represent differential *participation* rates for the solicited subpopulations, with a possible self-selection effect. Recruitment efforts for the study were pooled with another experiment that ran concurrently in the same physical location but with a different topic and compensation structure. Further, participants were not aware of the exact nature and subject area of the study. Therefore, subject self-selection should not be causally related to the variable of interest (valuation and attitudes of GMO characteristics), mitigating any concerns of potential bias in the main variable of interest.

To further account for the specific nature of the subpopulation sampled, this experiment used potato chips and chocolate bars rather than goods used in previous experiments, such as canola oil or potatoes.[32] College students are significantly more likely to be active participants in the market for these snack items than grocery shopping in general, allowing us to mimic the overall market setting better in terms of product knowledge, interest and attitudes.

DISCUSSION

GMO label prevalence significantly increases the willingness to pay for non-GMO products

The results provide strong evidence for a labeling prevalence effect. Merely changing the default labeling regime from identifying non-GMO products to flagging products containing GMOs significantly boosted the average willingness-to-pay for both these products, as well as the organic product (risk spillover).

Currently the average estimated WTP premium in the literature for non-GMO products is either close to zero in the older literature, or around 20-30% in newer studies. This study finds that simply drawing more attention to the characteristic through repeated exposure made over 55% of participants choose non-GMO price products at a 25% price premium (as opposed to only 21% in the control), and a further 35% of participants chose the product at a greater than 50% price premium (as opposed to ~20% in the control).

The 80/20 mandatory labeling regime caused participants to choose around 0.9 more non-GMO products over the total of six pairs offering them the chance (see Table 3, Appendix A for OLS regression), and along with gender was the only consistently significant determinant of participants' choices. These results hold even when introducing various controls going beyond demographics: GMO knowledge, attitudes, general price awareness etc. To account for the count-

[32] Huffman et al. relying on their use of an adult/general population sample to allow them to use common household items that participants would typically buy in the marketplace.

nature of the total score variable, a Poisson regression is run as a robustness check (see Table 7, Appendix A). The same variables are statistically significant, with the addition of a positive effect of 0.3 for Hispanic participants that is marginally significant at a 10% level. Under this specification the main effect of the treatment is a more moderate +0.6 non-GMO choices compared to 0.9 under OLS, but the relative magnitude compared to the gender and environmentalist effects remains the same (slightly smaller increase than for females, roughly twice the effect of being an environmentalist).

Studying choices in individual rounds with the probit model (see Table 4, Appendix A), participants were on average around 40% likelier to choose the non-GMO product offered in each round (see Table 4). There is no evidence of a significant learning effect (coefficient increasing with each successive round), or sensitivity of the effect to the actual price premium offered (coefficient being larger/smaller for higher price differences). However, the third pair containing non-GMO goods stands out as the only one with no significant treatment effect. This was the only round during which the product pairing offered had Lays' brand potato chips as the non-GMO alternative pitted against another brand[33], suggesting that brand effects and associations in participants minds might be powerful enough to override the treatment effect.

These results strongly suggest the debate over a mandatory versus voluntary labeling regime has significant implications for the development and size of the non-GMO goods market, but we have not yet discussed *how* the labeling effect is emerging: do consumers update their beliefs on the information, or merely change their preferences? Later sections explore which of the three channels identified in our model is causing the effect.

[33] Several permutations of brand pairings are present in the remaining five pairs ranging from same brand, same flavor to different brand different flavor. Thus the estimated effect is robust, but the 3rd pair offers us an insight into the limitations of extrapolating our results against *all* potential pairings in the market.

Gender is a significant driver of valuations, and thus support for labeling laws

In all of the model specifications used for this experiment, women are significantly more likely to choose the non-GMO product offered both overall and in each individual choice. No similar gender effect is observed for the one organic choice offered. This effect persist even when controlling for various factors beyond demographics such as knowledge of GMOs (true/false scores, where women score the same on average as men), risk attitudes (Holt-Laury lotteries, where women exhibit significantly more risk averse behavior), and even stated support for a mandatory labeling law (at 73%, 20 percentage points more support among women) or different estimates for the market price of chips (so the price awareness effect discussed below does not apply).

While no a priori theory thus justifies the prediction that the interaction effect for treatment and gender should not be zero, the strong empirical evidence suggests considering the possibility. When running the model with such an interaction effect, it turns out that gender on its own drops out (i.e. there is nothing a priori different about women's GMO valuations), however the interaction effect is positive and borderline statistically significant with a p-value ranging from 0.06 to slightly over 0.10 depending on the set of other controls included (see Table 5).[34] None of the other potential interaction effects (race, knowledge of GMOs etc.) come close to this level of explanatory power This suggests it might be possible the salience effect from labeling regimes is stronger for women than men, a possibility that merits further study in future experiments, especially as the absolute magnitude of this gender effect is close to, or even significantly larger than the treatment effect itself (see Table 5). The fact that the effect is not present for the organic product further suggests that i) there is no gender effect in risk spillover and ii) the effect is tied to the treatment and not a priori attitudes about labels or environmental characteristics. Whether the effect is an artefact

[34] Note that the inclusion of the interaction with all the other controls reduces the explanatory power of *all* variables due to the limited number of observations in the model.

of women paying more attention to the information provided and men simply clicking through the experiment[35], or actual sensitivity to the salience effect remains to be addressed by future work.

General price awareness is a determinant of WTP for non-GMO

The results suggest that people who most overestimate the price of chips (whose price guess is higher than both the actual price, and the mean guess of participants) are likelier to buy non-GMO goods when offered, perhaps due to less of a perceived *percentage* price premium. Consumers could value characteristics as GMOs on a logarithmic rather than linear scale, in which case the higher your perceived "base" price, the more of a dollar premium you are willing to pay for the same underlying percentage preference. A significant portion of participants estimated the "real" price of the items offered in the experiment at close to $1, slightly higher than the actual price (~70c), but noticeably lower than another cluster of participants centered in the $1.80-$2.20 range. For the latter group, the highest price differential offered represents around 30% of their perceived market price, while for the others it is closer to 70%.

While the prices offered in the experiment were fixed, it is conceivable that participants are influenced by their knowledge of a market for the desired characteristic just outside the experiment, in the form of various shops and cafes selling the same products on campus. Participants realize they can exchange the higher cash award for the desired products, and those with lower price estimates perceive less of a cost for doing so, and will thus value cash more highly than the group with higher estimates of prices.

The data supports this notion as the effect is linearly increasing in price guesses: i.e. the more you think chips usually cost, the likelier you are to buy the more expensive non-GMO alternative when offered. Similarly, "value buyers" (people whose estimate of chip prices are below

35 However, the time taken to complete the survey does not differ on average between men and women.

the group average) are significantly less likely to buy the non-GMO goods (-0.4 total score, p value 0.02, see Table 8 in Appendix A).

The valuation changes are not driven by differences in information

Further, it does not seem that the effect depends on the level of information participants have coming in to the experiment, i.e. the inclusion of the label did not change participant understanding: True/False scores for the questionnaire in Appendix A, Figure 2 are the same for both groups. This is not surprising as the label used here does not contain information beyond alerting the participant to the presence/absence of GMOs, or their estimate of the prevalence of the characteristic ("what percentage of packaged foods, e.g. potato chips, frozen meals, do you believe currently contain GMOs?").[36] The effect from labeling materializes through something other than better information about the underlying characteristic or its prevalence in the marketplace.

Thus the reference points posited by our salience model (treating a good with no information on the characteristic as having a probabilistic market "average" of that feature) do not seem to be influenced, but rather the change occurs in λ, our weighing parameter for that characteristic. Further research into what exactly is driving λ is necessary to distinguish between the mere presence of a label, and the potential effect the source can have.[37] Further, isolating a possible tipping point in the prevalence and generally studying the shape of the labeling effect across a range of probabilities will allow us to generalize the results of this experiment beyond the GMO context.

[36] See footnote 31 at page 17 for details.

[37] E.g. Shogren (2004) looks at how explicitly pro-industry, pro-regulation and third party labels have different effects. Note that in this study the non-GMO labels in question are a *private* label, and we would not expect to see any confounding deference-to-authority effect causing the salience effect.

The treatment does not influence opinions on a mandatory labeling law

It does not seem that the increased willingness to pay to avoid non-GMO products has any effect on the participant's desire to see these products labeled (see Table 6 in Appendix A). While summary statistics show a large jump for the treatment group in support for such a law, the effect is not statistically significant when controlling for other factors. The implications of this lack of change are surprising: the presence of GMOs has become a significantly larger determinant of purchasing intent for these participants, yet their desire to see that information readily provided does not increase.

A potential explanation is the participants' naïveté: both the treatment and the control group might believe they have access to all the information needed through the voluntary labeling of non-GMO products, unaware of the change in their valuations' caused by the change in labels presented. This would result in the disconnect between the actual change in behavior, and the perceived lack of value in the chosen information regime.

Alternatively if the participants are sophisticated naïfs[38] they would realize the labeling regime has the power to change their behavior, and in this case find such interference undesirable.[39] Given the large difference in purchasers of non-GMO goods between the groups (approximately 20 percent of all consumers in several rounds) this is not a fringe phenomenon, but a large share of the market changing their behavior without a commensurate change in their beliefs or attitudes.

[38] See O'Donoghue & Rabin (1999)

[39] As opposed to commonly presented scenarios of such self-awareness, where an intervention is sought out, e.g. in the form of a commitment device (automatic deductions for savings plans).

CONCLUSION

The debate over labeling often assumes that labels merely enable consumers to have access to more or better information, treating preferences over the actual characteristic itself as static. However, the results of this experiment demonstrate the preference shaping power of labels: participants did not change their beliefs about GMOs, or the prevalence of GMO products, but were still ready to pay much more for non-GMO alternatives. Previous experimental research has shown an incredible impact on valuations from emphasizing the GMO characteristic as part of the experiment, an issue most studies do not explicitly address at all. Whether the GMO premium is elicited as one choice in a series of unrelated questions (e.g. brand preference, fair trade) or in a battery of all-GMO questions will drive a large part of the obtained willingness to pay. This experiment has shown that i) labeling prevalence influences valuations ii) there is a potential gender effect iii) neither one of the previous two phenomena can be linked to changes in information available to consumers, or beliefs about the characteristic.

The shift in valuations observed here presents a challenge to policymakers, as it is accompanied by a commensurate change in beliefs or information about the now-labeled characteristic. Consumers who exhibit a higher willingness-to-pay as a result of the treatment do not fit the modeling of labels in valuation studies so far. As it is the preferences themselves that seem to be changing through the decision weight given to a particular attribute, which set of preferences should inform the regulator considering a labeling program, ex ante or ex post? How can we make meaningful comparisons of consumer welfare between these two states if the policy is changing the very parameters it is seeking to optimize? Finally, the role of brands should not be ignored, as it seems some brands are more easily granted premium valuations than others when presented with new information. Future work on labeling needs to address the potentially distinct effects arising from information, and the preference shift obtained here through the prevalence treatment.

REFERENCES

Akerlof, G. A. (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, *84*(3), 488-500.

Ben-Shahar, O., & Schneider, C. E. (2011). The failure of mandated disclosure. *University of Pennsylvania Law Review*, 647-749.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2013). “Salience and Consumer Choice” *Journal of Political Economy*, 121(5):803-843.

Gifford, K., Bernard, J. C., Toensmeyer, U. C., & Bacon, R. (2005, July). An experimental investigation of willingness to pay for non-GM and organic food products. In *American Agricultural Economics Association Annual Meeting, Providence, Rhode Island*.

Hammond, D. (2011). Health warning messages on tobacco products: a review. *Tobacco control*, 327-337.

Hebden, W. C., Aquino, H. L., Cuite, C. L., & Lang, J. T. (2003). *Public perceptions of genetically modified foods: A national study of American knowledge and opinion*. Food Policy Institute, (Publication number RR- 1003-004).

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American economic review*, *92*(5), 1644-1655.

Huffman W., Shogren J., Rousu M., Tegene A. (2003), “Consumer Willingness to Pay for Genetically Modified Food Labels in a Market with Diverse information: Evidence from Experimental Auctions”, *Journal of Agricultural and Resource Economics* 28, 481-502

Kaneko, N., & Chern, W. S. (2005). Willingness to pay for genetically modified oil, cornflakes, and salmon: Evidence from a US telephone survey. *Journal of Agricultural and Applied Economics*, *37*(03), 701-719.

Loureiro, M. L., McCluskey, J. J., & Mittelhammer, R. C. (2001). Assessing consumer preferences for organic, eco-labeled, and regular apples. *Journal of agricultural and resource economics*, 404-416.

Loureiro, M. L., & Hine, S. (2002). Discovering niche markets: A comparison of consumer willingness to pay for local (Colorado grown), organic, and GMO-free products. *Journal of Agricultural and Applied Economics*, *34*(3), 477-488.

Lusk, J. L., Daniel, M. S., Mark, D. R., & Lusk, C. L. (2001). Alternative calibration and auction institutions for predicting consumer willingness to pay for nongenetically modified corn chips. *Journal of Agricultural and Resource Economics*, 40-57.

Noussair, C., Robin, S., & Ruffieux, B. (2002). Do consumers not care about biotech foods or do they just not read the labels?. *Economics letters*, *75*(1), 47-53.

O'Donoghue, T., & Rabin, M. (1999). Doing it now or later. *American Economic Review*, 103-124.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, *1*(1), 39-60.

Viscusi, W.K, & Magat, W.A. (1987). Learning about Risk: Consumer and Worker Responses to Hazard Information. *Harvard University Press*

APPENDIX A

TABLE 3. OLS NON-GMO CHOICES REGRESSION

Total Non-GMO Choices (0-6)

	(1)	**(2)**
Treatment (1=GMOs labeled)	0.9243**	0.8465*
Sex (1=Female)	1.1284**	1.0378*
Black	-0.5139	-0.4049
Hispanic	0.6124	0.7793
Asian	-0.1551	-0.1106
Democrat	0.0698	0.0272
Republican	-0.3781	-0.4157
Citizen (1=Yes)	0.4592	0.3846
Environmentalist (1=Yes)	0.5404 †	0.4579
Lottery Score (0-4, 4=max. risk seeking)	-0.1720	-0.1689
GMO Knowledge Score (0-3, 3= max. knowledge)	0.0779	0.0584
Constant	0.3652	0.2606
Frequency of Buying Chips (0-4, 4=most often)		-0.0849
Seen non-GMO labels before		0.3609
Seen GMO labels before		0.5880 †
Demeaned Price Guess		0.5528 †

** Significant at 1% level *Significant at 5% level † Significant at 10% level

Notes: The total score reflects the number of times participants chose the non-GMO alternative when offered (up to a maximum of six times). Demeaned Price Guess is constructed by subtracting the group average estimate for the price of a bag of potato chips from the individual's guess (i.e. positive values indicate a higher than average estimate).

Table 4. Probit Individual Pair Choices of non-GMO (1=Non-GMO Product Chosen)

	Choice 1	Choice 2	Choice 3	Choice 4	Choice 5	Choice 6	Choice 7
	+20 cents	**+10 cents**	**+40 cents**	**+30 cents**	**+30 cents**	**+70 cents**	**+10 cents**
Treatment (1=GMOs labeled)	0.4796*	0.7607**	-0.2254	0.6318*	0.4949*	0.5422*	0.4453*
Sex (1=Female)	0.4651 †	0.5572*	0.5672*	0.7766*	0.4522 †	0.5936*	0.0622
Black	-0.4416	-0.6361	0.3875	-0.3229	-0.1732	-0.1164	-0.9969*
Hispanic	0.2414	0.2261	0.8902*	0.5359	0.1721	0.7975*	-0.1746
Asian	-0.1676	-0.2254	0.0807	0.0492	0.0223	-0.1447	-0.1918
Democrat	-0.0891	0.0123	-0.1909	-0.2838	0.2058	0.2069	0.0234
Republican	-0.1518	-0.2772	-0.5255	-0.2637	-0.1483	-0.0924	0.01658
Citizen (1=Yes)	0.0207	-0.1027	0.0821	0.1472	0.1936	0.2827	0.2416
Environmentalist (1=Yes)	0.3889 †	0.2751	0.3909	0.2727	0.3589	0.0442	0.4190*
Lottery Score (0-4, 4=max. risk seeking)	-0.1798 †	-0.1656 †	0.0346	-0.1142	-0.1206	0.0495	-0.1994*
GMO Knowledge Score (0-3, 3=max.)	-0.3131	-0.0100	0.2752	0.1657	0.0839	-0.0208	-0.0234
Constant	-0.2818*	-0.6653	-2.1406	-1.3533	-1.5386	-1.8769	-0.3343
Freq. of Buying Chips (0-4, 4=most freq.)	0.2180	-0.9193 †	0.0139	-0.0177	-0.1679	-0.0038	-0.0566
Seen non-GMO labels before	0.7285*	0.3323	-0.1979	0.1096	0.2499	0.4353 †	0.3640
Seen GMO labels before	-0.3515	0.0896	0.5759 †	0.2139	0.2947	0.3202	-0.1913

** Significant at 1% level *Significant at 5% level † Significant at 10% level

Notes: Each round is individually modeled with a probit regression (outcome 1=non-GMO product chosen) Omitted category for race is White, and for political affiliation Independent

TABLE 5. OLS GENDER EFFECT ON NON-GMO CHOICES

Total # of Non-GMO Choices (0-6)		
	(1)	(2)
Treatment	0.569 †	0.515 †
Sex (1=Female)	0.6444	0.468
Black	-0.479	-0.412
Hispanic	0.551	0.523
Asian	-0.132	-0.066
Democrat	0.205	0.105
Republican	-4.291	-0.371
GMO Knowledge Score		-0.089
Lottery Score		-0.117
Support Labeling Law (1=Yes)		1.195
Treatment **x** Sex	1.131 †	0.954

* Significant at 1% ** Significant at 5% † Significant at 10%

Notes: Omitted category for race is White, for political affiliation Independent. Total # of non-GMO choices modeled through OLS. Each non-GMO product chosen represented as 1 point, up to a maximum of 6.

TABLE 6. PROBIT LIKELIHOOD OF SUPPORTING LABELING LAW

(1= SUPPORT MANDATORY LABELING)

Treatment	0.254
Female	0.477**
Black	-0.213
Hispanic	0.044
Asian	-0.149
Democrat	0.164
Republican	-0.137

* Significant at 1% ** Significant at 5% † Significant at 10%

Notes: Omitted category for race is White, for political affiliation Independent.

Table 7. Poisson Regression of Total Number of Non-GMO Choices

Treatment	0.569*
Female	0.599*
Black	-0.353
Hispanic	0.388 †
Asian	-0.107
Citizen	0.279
Democrat	0.049
Republican	-0.271 †
Environmentalist	0.329*
GMO Knowledge Score (max 3)	0.019
Lottery Score (max4., most risk seeking)	-0.108
Constant	-0.306

* Significant at 1% ** Significant at 5% † Significant at 10%

Notes: Omitted category for race is White, for political affiliation Independent

TABLE 8. POISSON REGRESSION OF TOTAL NUMBER OF NON-GMO CHOICES, LOW PRICE ESTIMATE SPECIFICATION

Treatment	0.601*
Female	0.619*
Black	-0.367
Hispanic	0.481**
Asian	-0.058
Citizen	0.352
Democrat	0.063
Republican	-0.237
Environmentalist	0.384*
GMO Knowledge Score (max. 3)	0.011
Lottery Score (max. 4 , most risk seeking)	-0.107 †
Estimated Price Below Group Average	-0.407*
Constant	-0.256

* Significant at 1% ** Significant at 5% † Significant at 10%

Notes: Omitted category for race is White, for political affiliation Independent

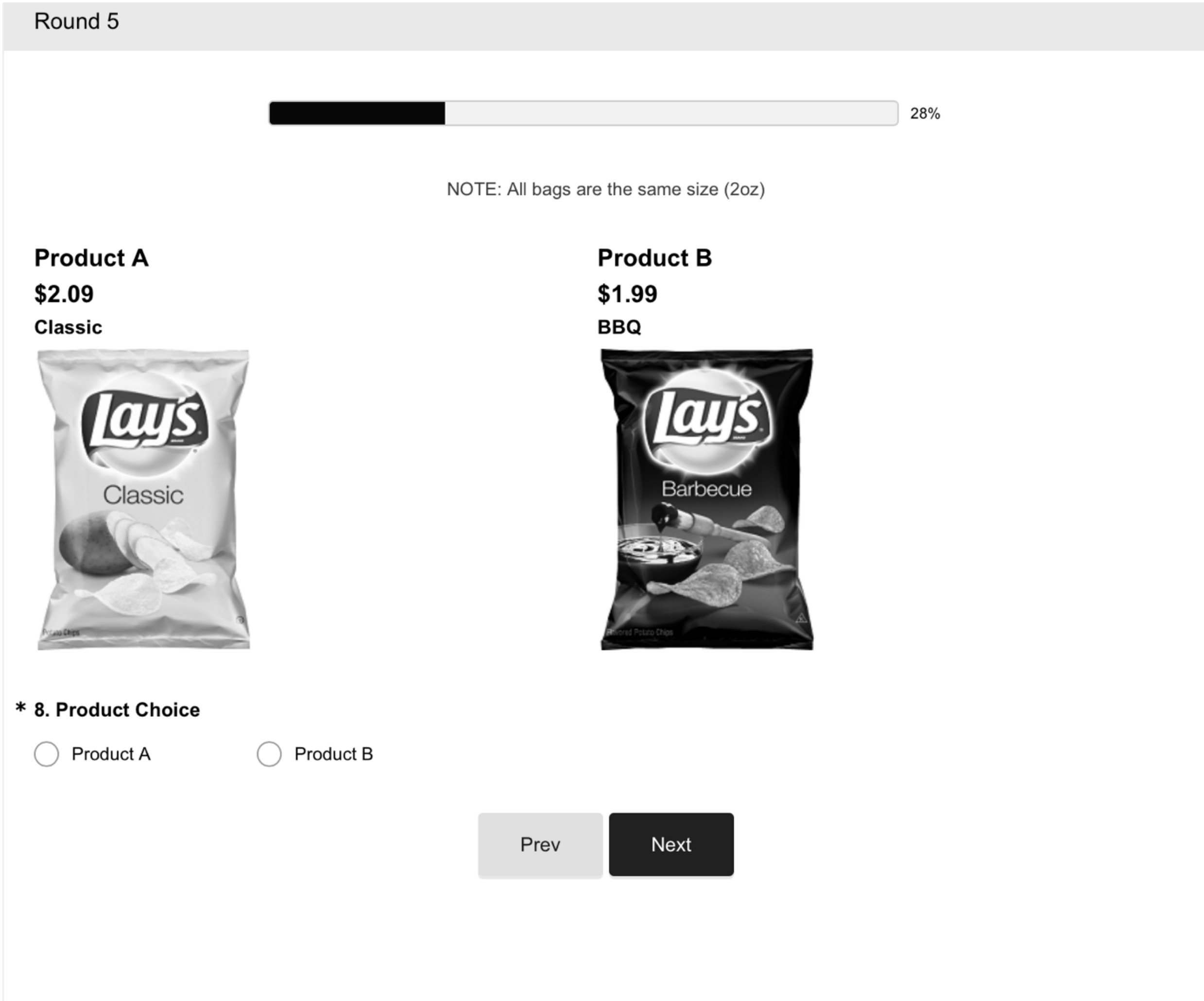


FIGURE 1. EXAMPLE OF DECISION SCREEN SHOWN TO PARTICIPANTS

44. True of False: All organic products are also free from genetically modified organisms (GMOs)?

- ◯ True
- ◯ False

45. True of False: Regular tomatoes do not have genes, but GMO tomatoes do?

- ◯ True
- ◯ False

46. True or False: Eating genetically modified food (GMOs) will change your own genes?

- ◯ True
- ◯ False

FIGURE 2. GMO KNOWLEDGE QUESTIONNAIRE

In this section you are presented with a series of choices between two lotteries. Please choose which lottery you would prefer for each set of odds.

34. Lottery A	**Lottery B**
◯ 40% chance of $2.00, 60% chance of $1.60	◯ 40% chance of $3.85, 60% chance of $0.10
35. Lottery A	**Lottery B**
◯ 50% chance of $2.00, 50% chance of $1.60	◯ 50% chance of $3.85, 50% chance of $0.10
36. Lottery A	**Lottery B**
◯ 60% chance of $2.00, 40% chance of $1.60	◯ 60% chance of $3.85, 40% chance of $0.10
37. Lottery A	**Lottery B**
◯ 80% chance of $2.00, 20% chance of $1.60	◯ 80% chance of $3.85, 20% chance of $0.10

FIGURE 3. HOLT-LAURY LOTTERY

Chapter 2: Demand Effects and Producer Response to rBGH Free Milk De-Labeling

Abstract
In 2008 Ohio introduced a regulation banning the advertising of milk based on the lack of bovine growth hormone (rBGH). From 2008 milk in Ohio was not allowed to have a "non-rBGH" label. These labels were only reintroduced after a federal court overturned the ban in 2010. Using a differences-in-differences model, this chapter examines the effect on the retail milk market from the removal of these labels. Compared to previous literature this chapter will examine a broader set of measures beyond the average price, including temporary price discounts and the use of in-store displays. The results show that producers of rBGH-free milk responded to the labeling ban primarily through the aggressive use of in-store displays (the use of which increased ten-fold during the ban), rather than engaging in any price discounting. The differences-in-differences model shows a sustained price premium of over $1 through the ban. The results highlight the need for a broader scope of inquiry, as traditional willingness-to-pay studies would have missed this effect.

INTRODUCTION

There is currently a lack of scientific consensus on the health effects to end-consumers of milk from cows treated with bovine growth hormones (rBGH or rBST). The Food and Drug Administration (FDA) has not found any difference in the milk from cows treated with the hormone in comparison to milk from untreated cows.[1] However, despite federal non-intervention the issue has garnered widespread consumer and regulator interest. On the one hand, rBGH-free products are becoming more popular every year - large retail chains such as Kroger and Publix have moved to only selling products from untreated cows[2], Starbucks has announced its intention to only use such milk,[3] and entire states have enacted policies to discontinue the use of hormones in dairy production.[4]

---

[1] "Report on the Food and Drug Administration's Review of the Safety of Recombinant Bovine Somatotropin", *available at* http://www.fda.gov/AnimalVeterinary/SafetyHealth/ProductSafetyInformation/ucm130321.htm
[2] See https://www.avma.org/News/JAVMANews/Pages/071001n.aspx
[3] See, e.g., http://www.nbcnews.com/id/16655614/ns/business-us_business/t/starbucks-moving-hormone-free-milk/
[4] https://www.organicconsumers.org/news/krogers-ban-monsantos-controversial-bovine-growth-hormone-tipping-point

On the other hand, there has been a backlash from some sources, including farmers and dairies, highlighting the supposed lack of compositional difference in the milk itself and the attractive economics behind using rBGH. In 2008 Ohio's Department of Agriculture found that the widespread "rBGH-free" labels were prohibited *compositional* claims (rather than process claims, which focus on characteristics unrelated to the final product, such as "family owned", "fair trade" etc.), and were potentially misleading consumers due to a lack of verifiable difference in milk from treated versus untreated cows. At the time of this regulation several other states were debating the introduction of similar labeling ban, and two states (Kansas, Indiana) came very close to passing such regulations before either backing down or being voted down. Ohio's ban was challenged in federal court, and in October 2010 the 6th Circuit Federal Court of Appeals struck down the ban, leading to the reintroduction of the hormone-free labels in Ohio. Notably, the court also stated that it believed there are in fact compositional differences between milk from untreated cows and from those treated with rBST[5], leaving the door open for further action by the FDA (or in its absence, to compel the FDA to act). The conflict over labeling rBGH-use mirrors a more widespread trend in consumer goods calling for greater transparency, often via mandatory labeling laws for issues like the use of genetically modified organisms[6], or a product's carbon footprint.[7]

This chapter studies the effects of this exogenous shock on the milk market, looking among other things at the absorption of information (how fast do people forget about rBGH-use in certain brands, and how fast do they "re-notice" the labels)[8], the milk producers' response to the changed law, and the relative importance of the label itself as opposed to broader media coverage and awareness of the issue (contrasting the market in Ohio to those in Kansas and Indiana). Since the

---

[5] *Int'l Dairy Foods Ass'n v. Boggs*, 622 F.3d 628 (2010)
[6] See Bracken (2012) for an overview of recent legal developments in GMO labeling.
[7] See, e.g., Vandenbergh (2011) for an argument supporting such a labeling regime.
[8] Teisl (2002) studies these absorption rates in the canned tuna market following dolphin-safe labeling.

data available extend to years before and after the "event" of interest, and includes "no-label", "media" and "control" states, the paper is able to move beyond the pre-existing literature's norm of either a time-series or cross-sectional analysis by constructing differences-in-differences, and triple-differences estimates. Further, the model used in this paper goes beyond simple price/volume observations to observe changes in other dimensions, such as temporary price discounts and the use of in-store display spaces, allowing this work to capture a broader set of market reactions to the ban.

In-store display use is often overlooked by consumer goods literature, even when such work is using the detailed IRi[9] data (used by this paper) including this information.[10] The physical placement of products in stores, however, is a key determinant of consumer decisions. As much as 68% of purchasing decisions in a typical grocery store happen at the point-of-sale, rather than being pre-planned and researched.[11] Producers have taken note, paying billions of dollars annually to retailers for shelf-space, aisle-end displays and similar display features.

Large producers such as Procter & Gamble and Coca Cola have scores of employees focusing full-time on merchandising activities like designing and pitching new planograms (the physical layout, or "shelving" of the products) and flashy displays to retailers. Bitter legal battles are fought over allegedly anti-competitive sale practices, such as Philip Morris' alleged use of its dominant market position to force retailers to stock competitors' products in less desirable store-areas far from the center and on lower shelves.[12] Similarly, McCormick has faced lawsuits for

[9] IRi is a leading market research company providing consumer and retailer analysis to its clients, while also offering academic researchers access to some of its store-level scanner data.
[10] IRi Consumer Goods Academic Dataset. See Bronnenberg et al. (2008)
[11] See, e.g., GMA Shopper Marketing Report (2007) , *available at* https://www.gmaonline.org/downloads/research-and-reports/shoppermarketing.pdf
[12] See, e.g. Forbes article "Shelf Determination" by Copple (2002), *available at* http://www.forbes.com/forbes/2002/0415/130.html

payments made to retailers keeping competitors' products away from "prime" space on the spice-shelves.[13]

Perhaps the researchers' reluctance to include display features in their analysis is due to the inherent difficulty of quantifying such display practices in comparison to clean-cut price change data.[14] However, the immense resources spent by producers in this area suggest it is of more than passing importance to the consumers decision-making process.

The use of in-store display space is especially important when considering product labels, as they are simply another point-of-sale feature. If in-store displays are meant to draw as many customers as possible to examine the product, labels are a tool for converting such views into purchases (customer views x conversion rate[15] = sales). As such, one would expect there to be some sort of relationship not only between prices and labels (co-determinants of the conversion rate), but also the use of display-space employed to funnel customers into the conversion stage. By ignoring this first stage, researchers risk misunderstanding the overall market positioning and response of producers to a label, and potentially mischaracterizing any subsequent consumer behavior.

The problem is exacerbated by the current state of experimental literature on the effects of labels. Most experiments are not designed to gauge the true "market" effect of a label, as they fail to control for the salience of the label to consumers (ensuring attention by, often explicitly, forcing participants to evaluate the studied characteristic), skewing the resulting valuations.[16] The experimental literature has focused on isolating the label from surrounding "noise" in markets, but

---

[13] See Schoenberger (2000) article in Forbes, *available at* http://www.forbes.com/forbes/2000/0612/6514084a.html

[14] Following the coding of the data, this paper dramatically simplifies the dimension into three possible states in ascending order: no features, some feature, aisle-end feature

[15] The probability that a consumer viewing the product buys the product.

[16] This issue is explored in depth in Chapter 2 of this dissertation, starting in section III.

this artificially inflates the focus given by participants to the characteristic in question.[17] This chapter's broader look at the actual market effects of label changes allows us to draw important policy implications and account for the dynamic behavior of producers in determining the final outcome of a labeling scheme, potentially muted by producers shifting their marketing resources into other, less regulated areas including the use of in-store displays.

LITERATURE REVIEW

A. The history and economics of bovine growth hormones

Bauman (1992) outlines the early history and biochemistry behind bovine growth hormones. Somatotropin (the "growth hormone") had already been discovered prior to the Second World War, but it was not until the 1940s that a link between the hormone and increased lactation (in rats at first) was discovered. The commercial breakthrough started decades later in the 1980s when advances in DNA-technology allowed for the isolation of the gene controlling bovine somatotropin (BST) production in cows[18], and subsequent laboratory mass production of the hormone, thus creating recombinant BST, or rBST/rBGH.

Using rBGH allows dairies to increase their milk production by up to 20-25%, as well as improving the cows' efficiency in converting energy from feedstock to milk.[19] Further, the cost of rBGH is less than $0.45/cow/day, while yielding an additional 8 pounds of milk every day. Conceivably, a high proportion of the costs incurred by dairies are fixed at a farm level (e.g. processing technology like sterilization equipment etc.), and a high proportion of the remaining

[17] See section IV, Chapter 2 of this dissertation, showing a link between prevalence of labeling and implicit valuations of the labeled characteristic.
[18] Discovered by the biotech company Genentech in the 1970s.
[19] Untreated cows in the study used 10% more of their energy on milk production

costs are incurred on a per cow basis (room for each animal, milking equipment etc.). This means that the profit margin on each incremental gallon of milk *per cow* has a significantly higher profit margin than the dairies' pre-existing average.[20] Butler (1999) arrives at a conservative estimate of an 11% marginal profit increase per cow from using rBGH, which is not taking into account any possible increases in feed efficiency or reduction in the overall share of fixed costs per gallon of milk produced, leveraging any initial profit increase.

Due to its highly desirable economics, rBGH use became widespread in the United States under largely Monsanto's distribution. Butler (1999) estimates that at least 20% of cows nationally were treated with rBGH, and Monsanto reported 35-40% sales growth for the hormone through most of the 1990s. Lesser et al. (1999) reported adoption rates of 39% in New York by 1996, while Butler (1999) finds that some 30% of dairies in California used the hormone in 1998.

The fact that milk sells on the global commodity markets and can be "blended" from several dairies at a larger processing plant means that even if less than half of all cows in the US were treated with rBGH, significantly more than half of the milk sold in retail markets by volume *may* contain milk from a treated cow. Indeed, between 2002-2011 in the retail data used by this study, only about 1% of the milk sold was *labeled* as rBGH-free.[21]

Since the use of rBGH is not immediately chemically discernible from the milk itself, being able to label the milk in this way requires producers to be able to track down each gallon of milk down to the farm/cow level, and have some certification system in place to verify that the cows are not being treated with rBGH. The complexities and potential liability associated with a quality control slip partially explain the disconnect between the percentage of cows never treated with rBGH, and the percentage of retail milk labeled as coming from such cows.

---

[20] Note that of the gallons per cow marginal costs the highest is likely in feed consumption, a ratio which rBST also improves.

[21] A significantly higher proportion could also *be* rBGH-free, but not labeled as such.

B. Effects of consumer goods labels on the retail market

This chapter draws on an established literature of label-effects on consumer demand generally (Teisl et al. 2002), as well as studies focused on the demand for milk specifically (e.g. Heien 1988, Kiesel 2007). This paper also draws on marketing literature to establish the case for looking beyond a simple price v. volume optimization model for producer responses. For example, Yoo et al. (2000) find the frequent use of price promotions erodes the consumers' perception of brand quality. This suggests the producers' likely response to a demand shock might not play out in just these two dimensions (unit sales, price per unit), but includes the use of several other marketing methods as well (e.g. advertising and in-store product displays).

Teisl (2002) finds that with the introduction of dolphin-safe labels in the canned tuna market there was a noticeable effect on sales, with a s-shaped time lag. In other words, from the introduction of the labels there was a small, slowly growing effect on the market that accelerated into exponential growth, before plateauing at a higher level. This study has a chance to observe the information-absorption dynamic at a granular level (weekly data), for both the de-labeling (how fast do the effects of a previously featured label fade), and subsequent relabeling (how fast does the reintroduced label start influencing consumers again). In this case the rBGH-free labels had been in widespread use for more than a decade, and the duration of the labeling ban spans several years allowing us to adequately capture the full effect of removing a "mature" label from the marketplace.

The impact of rBGH-free labeling in milk specifically has been studied primarily by Kiesel (2005, 2007). While using data provided by IRi, as in this chapter, Kiesel models consumption solely on a national level with time-series data. Notably absent from the studies are changes in labeling, either in geographic areas or individual brands of milk. Further, the paper only examines sales volume, and does not control for the possibility of changing price premiums over the study period (on average the unit price difference during this time was 15-30% depending on the container size). As rBGH labels gain popularity it is conceivable that high-cost niche producers are joined by "core value" brands, such as Kroger's store brand, which started carrying the rBGH-free labels at the start of 2008. Thus increasing sales over time might result not only from increased willingness-to-pay (WTP) for the label, but lower cost alternatives. The data supports this hypothesis, as overall the per gallon price premium for all rBGH-free products on average is only 2-3%, much smaller than the premiums observed by Kiesel in prior data.

This chapter builds on the work by Kiesel by focusing on the "shock" of labeling (or removal thereof), and attempts to isolate the labeling effect from any underlying national time-trends and shifts in preferences. Further, this study utilizes the available data more fully by examining the price premium commanded by rBGH-free products before and during the labeling ban, as well as the potential role of in-store display space in the producers' response.

Finally, it is unclear whether the removal of a label should have any effect at all, especially in the short run. Small purchases such as a gallon of milk are likely low-involvement habits, where the consumer does not assess the options available to them again each time they visit the grocery store. Aggressive price promotions are used by producers to overcome and capitalize on this stickiness (discounting enough to make consumers switch during the promotion, and hopefully convert a fraction of these into regular customers). Pesendorfer (2002) finds pervasive use of such frequent, sharp discounting practices in the retail market for ketchup in the United States, although he assigns the prevalence of the practice to simpler price-discrimination motives.[22]

With respect to product "stickiness", the question this paper sets out to answer is whether the removal of the label constitutes a sufficient "shock" analogous to a competitor's price discount[23]

[22] Producers keep prices at regular levels for extended periods, serving the less price sensitive portions of the market, then engage in temporary promotions to "clear" built-up demand from more value oriented consumers.
[23] As the now unlabeled product has lost its visual differentiator, and is still asking for a higher price

to induce consumers to switch[24], and whether this results in pricing pressure for the rBGH-free brands. This paper also observes the discounting behavior of "regular" milk producers to capture the possibility of price discounting of *ordinary* milk to overcome consumer's sticky habits (due to e.g. search costs) and induce product switching.

DATA

This paper uses the IRi Consumer Packaged Goods Dataset made available for academic research. The data consists of weekly, SKU[25]-level panel data on sales of a variety of items per grocery store (and drug stores) extending from 2001 through 2011, including sales of all liquid milk products. In addition to the number and price of the goods sold, the use of temporary price discounts (>5% off of base price) and in-store displays (e.g. aisle-end units, special stands) is recorded for each item. While not necessarily a representative sample of all products[26], consumers and store types by design (e.g. the data seems to skew in favor of traditional grocery stores, not including as many "dollar stores" and other alternative sources for groceries), several grocery stores from most states are present –including stores in Ohio, Kansas and Indiana.

Individual stores are located for the model based on their reported metropolitan statistical areas (MSAs). Within Ohio the study will look at stores in Cleveland (~28 grocery stores) and Toledo (~21 grocery stores)[27] and for the "media only"[28] states Kansas City[29]~22 and Indianapolis ~

[24] See, e.g. Varian (1980) for a theoretical model of sales, search costs and consumer product switching.
[25] Stock Keeping Unit
[26] Very small local brands would not be present, nor all private labels. However, several small regional producers are represented, as well as private labels of large retailers.
[27] These are the only two MSAs present for Ohio.
[28] States that considered passing such legislation but did not finalize such rules
[29] Note that due to the multi-state span of Kansas City and lack of precise location data in the iRI dataset this potentially includes stores in Missouri as well as Kansas. However, since there was no actual labeling change

16). The primary model is applied only to grocery stores, as drug stores represent only minimal volume for milk sales[30], and presumably these sales occur in a very different context from grocery stores. While grocery stores are not identified by chain name, the product assortment suggests a range of price points from value stores to high-end organic/imported goods retailers are present.

On the product side, starting from the master dataset, this paper filters out all observations outside the liquid milk category (iRI supplies the data pre-grouped by product category, and the observations were further verified by the author against an exhaustive SKU-list provided with the data). Each milk product is then coded as either rBGH-free, or unlabeled. The classification is done on the basis of the iRI data itself, consumer reports from the time[31], and verified with the individual dairies where possible.[32]

Nationally there are approximately five million unique observations of milk sales for each year (see Table 1 below), of which roughly 50,000 (one per cent) are for rBGH-free milk. For Ohio there are some 200,000 annual observations of which more than 2000 are rBGH-free milk, yielding plenty of data even for the most sparsely populated cells.

---

in one state or the other, it is reasonable to assume the local news coverage for the metropolitan area was roughly equivalent on either side of the state border.

[30] Less than 2% of milk volume in the data.

[31] Using reports by *Food Democracy*, *Center for Food Safety,* and *Organic Consumer Association*.

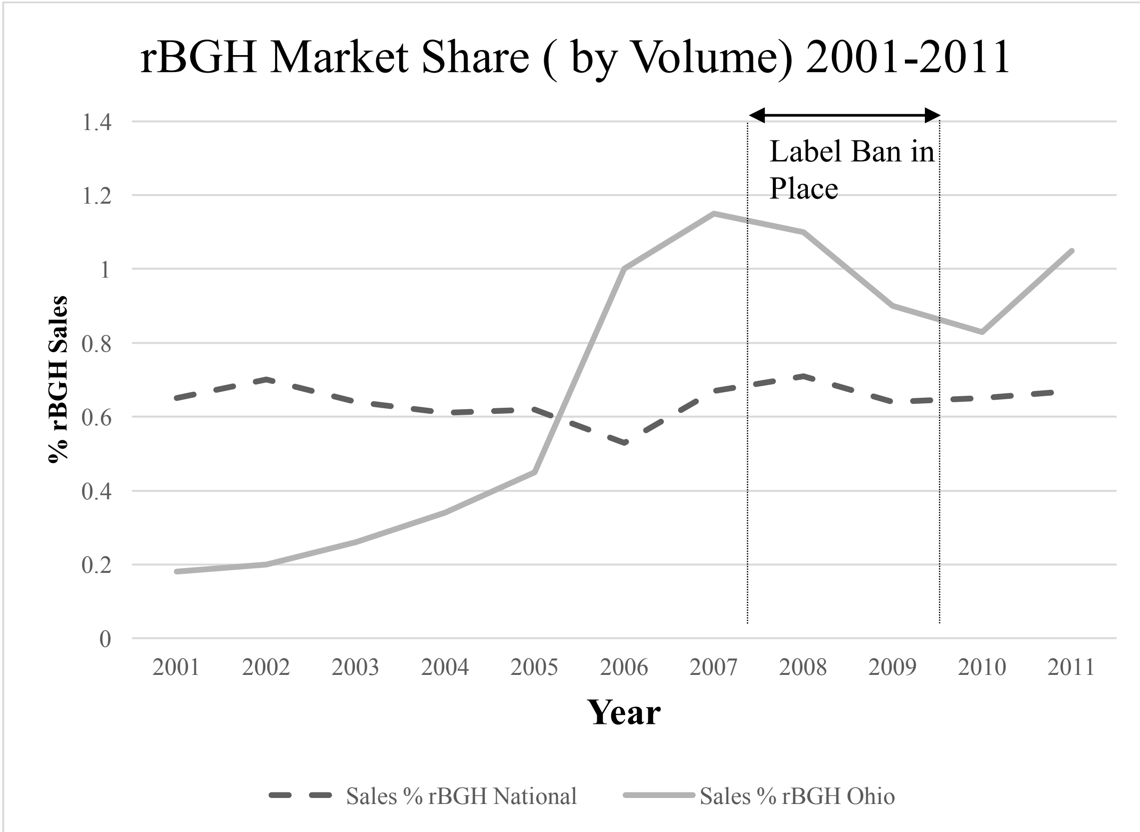
[32] All dairies whose products were identified by consumer reports as rBGH-free were contacted by the author to verify this status for products sold in Ohio between 2001-2011. Of these 10+ dairies only three responded (Dean Foods, Prairie Farms and Oakhurst Dairies), however at least for this subset the information from consumer reports was accurate, suggesting it is reasonable to rely on the accuracy of the consumer reports.

TABLE 1. MILK SALES SUMMARY STATISTICS

	Mean	Std. Dev.
rBGH-free	0.086	0.28
Price ($)	$2.78	$1.22
Display	0.023	0.2
Ohio	0.013	0.16
Kansas	0.017	0.131
Indiana	0.005	0.074

Notes: Total number of observations 65,314,709. This represents the entire IRi-dataset on liquid milk sales from 2001 to 2011.

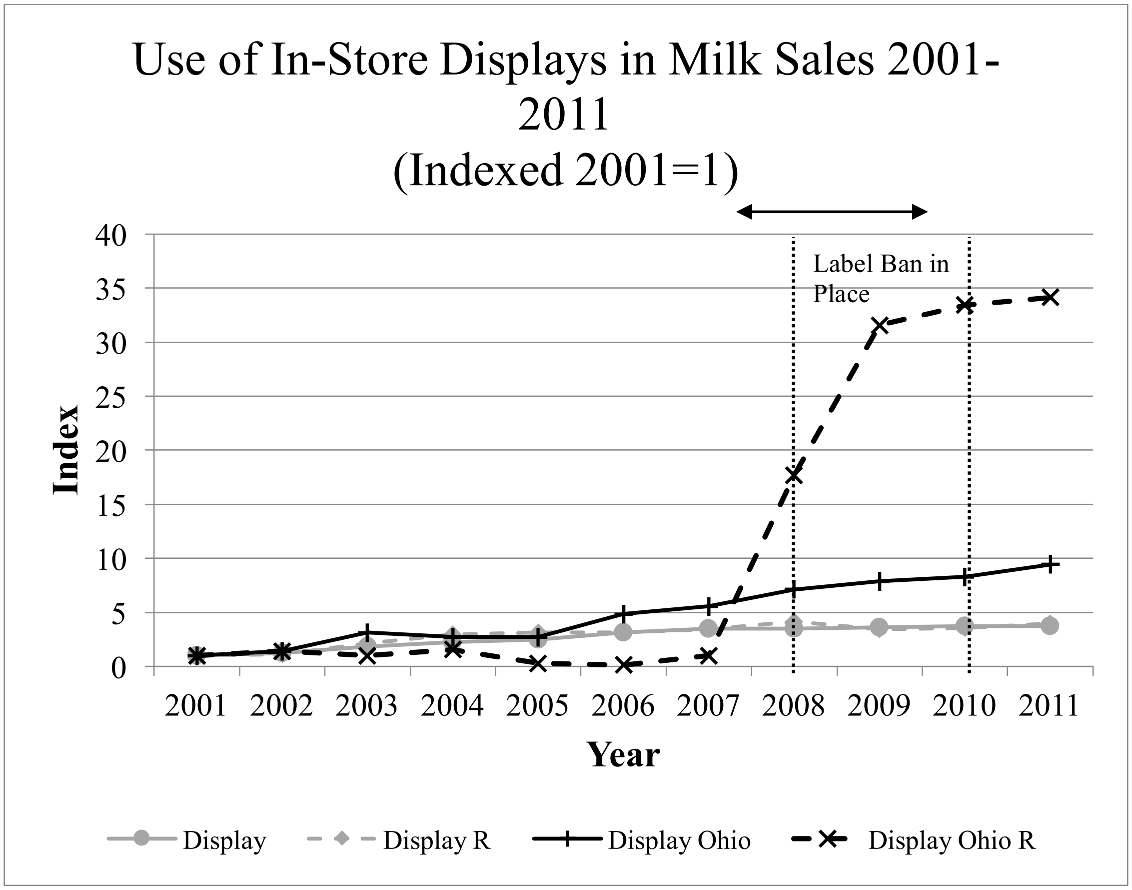
Immediately visible from the data is the effect of the ban on milk sales in Ohio (see Figure 1 below). Prior to 2008, rBGH-free milk was rapidly gaining share even beyond the national average, but from 2008 to 2010 this trend is sharply reversed, only to start growing again in 2011 after the removal of the ban.



Notes: The Sales % rBGH (black line) measures the market share of rBGH-free produts by volume (oz) sold nationally, while the Ohio Sales% measure tracks the market share within Ohio. This and all following graphics are drawn from the IRi consumer data grouped and edited by the author for the regressions in Appendix B.

FIGURE 1. SALES VOLUME

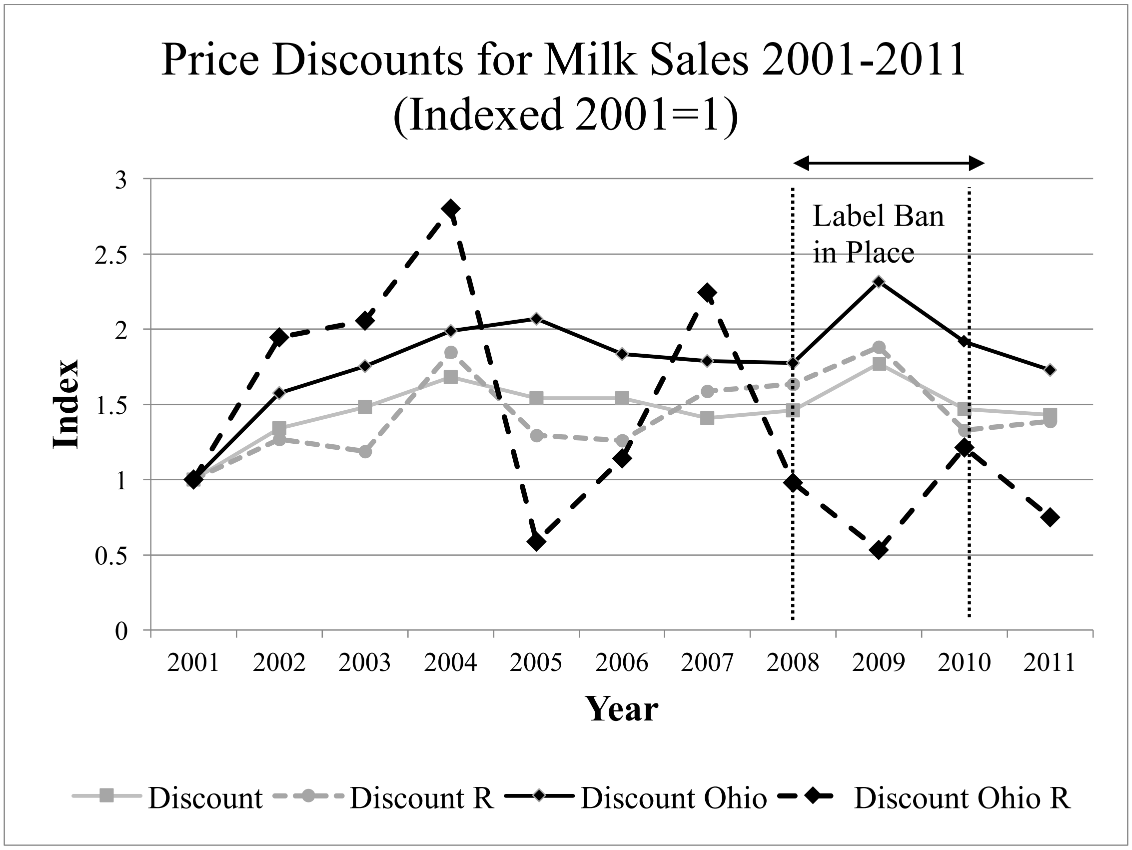
The raw data also points to the significance of in-store displays (see Figure 2 below): during the ban from 2008-2010, display use for rBGH-free milk in Ohio increased more than thirty-fold (from a base of less than one percent to more than twenty percent of all rBGH-free milk sold). Nationally only about two percent of milk sold uses a display, so the pattern in Ohio is a significant deviation from broader market practice. Notably, display use in Ohio during the ban was elevated compared to the national average even for “conventional” milk, further suggesting that displays played a significant role in the market dynamics during that time.



Notes: The Display R and Display Ohio R track the use of in-store displays for rBGH free milk products nationally (ex. Ohio), and within Ohio respectively. The baseline for the index in 2001 is approximately 0.7 percentage points in Ohio, and 1.2 percent nationally.

FIGURE 2. USE OF IN STORE DISPLAYS

Further highlighting the importance of controlling for display use in models is the fact that no dramatic changes are visible during the ban in the raw price data (see Figure 3 below). In fact, the data suggest an increase in discounting by conventional producers (perhaps to lure consumers of the now de-labeled milk to switch), but also a decrease in the amount of discounts offered by rBGH-free producers. This behavior goes against conventional wisdom as rBGH-free milk should have become *less* attractive to consumers with the removal of the identifying labels, leading us to expect an increase not a decrease in the discounts offered.



Notes: The Discount R and Discount Ohio R track the price discounts of rBGH free milk products nationally (ex. Ohio), and within Ohio respectively. Discounts are defined as temporary price decreases of 5% or more.

FIGURE 3. PRICE DISCOUNTING

METHODOLOGY

This paper will focus on two primary measures of demand for rBGH-free milk: the actual amounts sold, and as a determinant of sales, the prevalence of price reductions for non-rBGH products. An additional variable to be examined is the visual prevalence of non-rBGH products on store shelves (e.g. “displays” in stores, either at aisle-ends, or “windows” on the aisle itself). To develop the differences-in-differences model, this chapter will use additional controls including, but not limited to, regional demographics and unemployment rates. This will allow for more comparable “control” cities to be used in addition to the national (excluding cities with the ban) averages. The comparability of cities, and the use of macroeconomic indicators is warranted due to the unusual nature of the time period of the labeling ban, coinciding almost perfectly with the financial crisis of 2008, and subsequent recession. By accounting for such macroeconomic trends we remove any noise caused by a potentially disproportionate impact of the crisis on places like Cleveland and Ohio.

Model

The chapter will primarily rely on a differences-in-differences model (as well as triple-differences). Potential areas of concern include a simultaneity problem (sales and price combine to create a single observed level), unobserved store variation (beyond store type and demographics), and the lack of any advertising data for liquid milk producers from the time period. To account for the nested nature of the data (potentially serially correlated standard errors between observations) the standard errors are clustered at the individual store level.

The analysis will focus on three dependent variables: i) sales volume (measured in gallons), ii) average sales price ($/gallon) and iii) the probability of observing in-store display features for a sale. Within each model several sets of explanatory variables will be modeled (mostly binary “dummy”-variables signified by a *d* in the models below), including time-trends and layers of controls for local macroeconomic indicators (unemployment rate). The models will all follow a similar format:

Model 1.1 (Volume-OLS)

Y(Sales volume for milk, gallons)=

$\alpha + \beta_1(d[rBGHfree]) + \beta_2(d[Ohio]) + \beta_3(d[Kansas]) + \beta_4(d[Indiana]) + \beta_x(d[Year]) \ldots +$

$\gamma_1(d[rBGHfree] * d[Ohio]) + \gamma_2(d[rBGHfree] * d[Kansas]) + \gamma_3(d[rBGHfree] *$

$d[Indiana]) + \delta_x(d[rBGHfree] * d[Ohio] * d[Year]) \ldots$

Model 1.2 (Price-OLS)

Y(Sales price for milk, \$/gallon)= $\alpha + \beta_1(d[rBGHfree]) + \beta_2(d[Ohio]) + \beta_3(d[Kansas]) +$

$\beta_4(d[Indiana]) + \beta_x(d[Year]) \ldots + \gamma_1(d[rBGHfree] * d[Ohio]) + \gamma_2(d[rBGHfree] *$

$d[Kansas]) + \gamma_3(d[rBGHfree] * d[Indiana]) + \delta_x(d[rBGHfree] * d[Ohio] * d[Year]) \ldots$

Model 1.3 (Display-Probit)

Y(Display feature [0/1])= $\alpha + \beta_1(d[rBGHfree]) + \beta_2(d[Ohio]) + \beta_3(d[Kansas]) +$

$\beta_4(d[Indiana]) + \beta_x(d[Year]) \ldots + \gamma_1(d[rBGHfree] * d[Ohio]) + \gamma_2(d[rBGHfree] *$

$d[Kansas]) + \gamma_3(d[rBGHfree] * d[Indiana]) + \delta_x(d[rBGHfree] * d[Ohio] * d[Year]) \ldots$

Note that while the first two models are potentially problematic due to a simultaneity problem with price and quantity (as discussed in the sections above), the display model does not have this problem. Both the volume and price models use ordinary least squares, while the display use is modeled through probit to account for the binary nature of the display variable being studied.

The primary variables of interest are the three-way interaction terms of rBGH-free x Ohio x Year, which allow us to isolate the effect of the ban in Ohio on the rBGH free products sold there. Similarly, the three-way interaction terms for rBGH-free x Year x Kansas and Indiana allow us to further differentiate the effect of the actual labeling ban from a potential separate effect created by the publicity surrounding the issue (e.g. increased awareness of hormone use in milk, shifting demand independently of the actual labeling regime chosen). Comparing these terms for years during which the ban was in place to years before and after that further allows us to control for any underlying time-trend (e.g. decreasing price premiums), while still detecting any discontinuous movement in that trend during the ban.

Hypotheses

<u>H1 (Decreased Price)</u>

In the price model, the interaction-term rBGH-free x Ohio x Year should be (disproportionately compared to other years) negative for the years 2008-2010, while the labeling ban is place. This reflects the lost competitive advantage of the physical labels, and subsequent pricing pressure. One would expect producers to be fairly short-run inelastic in their milk supply, as the investments in livestock, distribution etc. all represent fixed costs in the time period considered.

For the media-only states (Kansas, Indiana) one would expect these interaction terms to be either insignificant (no effect from media), or even positive (increased awareness of issue combined with continuing presence of labels allowing consumers to differentiate between products). It is also possible to observe a negative effect for the media-only states as producers attempt to capitalize on the temporarily increased consumer attention with aggressive price promotions designed to entice more consumers to “switch” their purchasing habits.

<u>H1</u>

:rBGH-free x Ohio x (2008-2010) <0 and

rBGH-free x Ohio x (2008-2010) < rBGH-free x Ohio x (pre-2008, post 2010)

H2 (Decreased Sales)

One would expect the absolute sales volume to either fall during the ban, or assuming a perfectly compensating strong re-pricing/promotion response from producers it could stay the same (H2A). However, it is possible that an overlying trend of increased popularity for such products continued regardless of the ban. However, the overall sales volume should at least satisfy H2B (weaker hypothesis), where any such *absolute* growth is still weaker during the ban than in the years before and after. Similar to H1, the predicted value for Kansas and Indiana would either be insignificant (no effect from media), or even positive (increased awareness and continued presence of labels lead to increased demand).

H2A: rBGH-free x Ohio x (2008-2010) ≤ 0 and/or

H2B: rBGH-free x Ohio x (2008-2010) < rBGH-free x Ohio x (pre 2008, post 2010)

H3 (Difference in use of display space)

The final hypothesis tested is the possibility of producers reacting to the ban by increased use of in-store displays. As the label being forcefully removed from the products was primarily a *visual* differentiator of the products at the point of sale, the closest substitute available to rBGH-free producers is arguably the use of display space to re-capture the lost consumer attention. Admittedly, the label influences more the likelihood of purchase *upon* inspection of the product, while in-store displays are more geared towards getting more consumers to inspect the product at all. Thus, it is unclear whether the primary effect observed from the labeling ban should be H3 (display) or H1 (price, a purchasing determinant once product is being inspected).

Alternatively, if the conversion rate upon visual inspection of the product decreases significantly due to the removal of the label, it is possible the economic rationale for spending money to funnel consumers into this second stage is weakened, and we actually observe a *decrease* in the use of display space.

Similar to H1, here the effect in the media-only states could cut either way: increased demand could allow rBGH-free producers to cut back on their promotional budgets, but they could also attempt to capture market share with aggressive display campaigns (with an elevated likelihood of purchase upon inspection, the numbers might justify a larger spend to funnel consumers into noticing the products in the first stage).

<u>H3:</u>

rBGH-free x Ohio x (2008-2010) >0 or

rBGH-free x Ohio x (2008-2010) <0

RESULTS

TABLE 2. MAIN FINDINGS FOR rBGH-FREE MILK MARKET IN OHIO

		Result
Hypothesis 1	Price Decrease	**X**
Hypothesis 2	Sales Decrease	**X**
Hypothesis 3	Change in use of Displays	√ **(increased)**

The summary statistics (see Table 1 above) suggest total sales and market share of rBGH-free milk in Ohio declined significantly during the labeling ban, and resumed growth after the reinstitution of the labels. No similar time-trend was present in the national averages. However, in the regression model (Appendix B, Table 3) the only statistically significant dip in volume occurs in the first year of the ban. In stark contrast, the first year saw significant *increases* in rBGH-free sales for Kansas and Indiana, as hypothesized earlier (the increased media attention translating to higher consumer demand).

The primary model captures little price discounting due to the de-labeling, despite falling sales volumes, suggesting that producers did not seek to protect their market share via more aggressive pricing. On the contrary, there is a significant price *increase* coinciding with the years the ban is in place. Starting at roughly 70 cents in 2008, the price increases escalate through the bam reaching levels higher than $1/gallon, or over 30% of the base price for 2009 and 2010. Even after the ban is lifted, the price premium increases a further 7 cents to $1.24 (well above 40% of the base price) in 2011. These large premiums run against the idea that de-labeling weakened rBGH-free milk's market position. A potential explanation is in repositioning by producers into a more premium price, low volume niche. A later specification will attempt to control for a potential shift in product mix by considering the premium priced products separately.

The price effect is slightly dampened when we control for the overall time-trend in prices for rBGH-free milk in Ohio: the interaction terms for the years before the ban is around 80 cents/gallon, compared to the $1.03 average during the ban (see Figure 4 in Appendix A). This still suggests over 20%, or roughly 20 cents of the observed price premium is a direct result of the ban.

Note that there is no similar increase in the national prices, or in the media-only states. Kansas experienced significant discounting (coupled with increasing sales volumes), while there was a slight increase in prices in Indiana (also coupled with increasing sales volumes, suggesting an overall increase in consumer demand). For Kansas the average discount of rBGH-free milk compared to all other milk hovers around 40 cents/gallon, peaking at nearly 60 cents in 2009 (see Table 3 in appendix B). In percentage terms this is a significant 20% of the base price. For Indiana there is a small, but statistically significant price premium of roughly 10 cents/ gallon through the year 2008 to 2011 (see Table 3 in Appendix B). The effects in both of these states are significantly smaller than the Ohio de-labeling effect. Further, in 2011 rBGH-free milk in both Kansas and Indiana either held prices steady at a discount (Kansas) or decreased prices to on par with conventional milk (Indiana). This makes the post-ban increase in Ohio's price premiums all the more significant as it occurred against a backdrop of falling price premiums elsewhere.

This price behavior is somewhat surprising as one would expect liquid milk production to have fairly high fixed costs/output in the short run (i.e. producers cannot easily dial-down production volumes in response to the now lower demand). One possible explanation is that more of the milk is simply sold out-of-state (Ohio's share of all rBGH- free sales is <5% nationally, and so the demand shortfall could represent an undetectably small shift into national sales). Further, as noted earlier the overall average price premium for rBGH-free milk has all but vanished during the time period observed here (<5% premium for rBGH-free products).

To account for this phenomenon this paper uses an alternative segmentation, running the model and separating rBGH-free milk with an initial price premium >5% to the overall milk average (see Table 4 in Appendix B). This way the paper captures the reaction of *premium* producers who arguably have a significantly greater capacity to engage in price cutting as needed. Controlling for this we see mixed evidence for the non-premium rBGH-free milk producers, with prices swinging from a discount of 30 cents at the start of the ban in 2008 to only a 10 cent discount by 2010 and finally a 30 cent premium with the removal of bans in 2011. While not conclusive, these results suggest that non-premium producers were not engaged in protracted, aggressive discounting and might even have raised prices during the ban. For premium producers, none of the interaction terms for premium products' price during the ban are statistically significant. Taken together the results suggest that the refusal to cut prices was universal among rBGH-free milk brands, rather than caused by just one segment of producers.

One possible explanation is that producers were simply repositioning their products in the short-run to emphasize profit per consumer, rather than volume sold. While excess capacity is expensive, it is possible that the removal of the label primarily hurt the ability of rBGH-free producers to attract new consumers on the margin. Conceivably a large segment of purchasers of rbGH-free milk could be very price inelastic (willing to pay a lot to avoid rBGH products) as well as well-informed (know about rBGH-free characteristic even without label, e.g. through consumer reports). In this case, the profit maximizing move for newly de-labeled producers could be to abandon efforts to attract marginal consumers (through price discounts), and instead extract more margin from their core group of consumers. While credible, as highlighted below the producers' actions in the use of display space run counter to this narrative, as we see significant investments particularly into increasing product visibility and mass appeal.

It is important to note that most of the existing literature would end the inquiry here. Typical time-series studies of the price/volume mix have not sought to examine the market response in more detail. For this paper, however, the rich IRi dataset allows us to also control for temporary price reductions and the use of in-store displays. In fact, it seems most of the effect of the labeling ban occurs in in-store display usage which increased tenfold during the ban (see Figure 3 above).

In the differences-in-differences model (see Table 3 in Appendix B), the likelihood of using in-store displays increases each year the ban is still in place, rising from 1% at the start of the ban to 3% at the end, and 6% in the year after. In estimating the use of displays via a probit model, the three-way interaction term of Ohio x rBGH free milk x Year is not significant for years before the ban, but is significant and even higher in magnitude for the year after the ban. Further study of this phenomenon is warranted to see if the term would still be significant in subsequent years after the removal of the labeling ban. As it stands, the model could be picking up an aggressive campaign to regain lost market share by exposing as many customers as possible to the newly *relabeled* milk products (as supported by the increased sales volumes in 2011).

The focus on in-store displays makes intuitive sense: the labeling ban removed a key *visual* brand differentiator for rBGH-free producers, therefore boosting their products' visibility in retail locations through other means (i.e. aisle-end displays and such) is the most direct measure available to compensate for the lost feature). While we are unable to observe the specific characteristics of the displays being used (e.g. do they advertise features relating to the lack of growth hormones?), we know that by their very nature they are designed to attract more consumers to view the product. Despite this effect, the disproportionate use of display space is still surprising as the fees tend to be high: estimates suggest typical costs of over $200/day/store/aisle, or off the shelf-space pricing for new products in excess of $6500/month/retail chain/metropolitan area.[33]

While not all in-store displays captured by the dataset would reflect such "prime space" costs, given the low sales volumes at the stores observed (roughly 2100 units per store annually, or ~6 units/day in Toledo, ~35/day in Cleveland) any fixed costs born for the space likely reflect a significant increase in the producers' costs per unit sold, certainly in excess of several percentage points of total sales.

Notably, even with the massive increase in the in-store placement (and thus costs), overall sales were reduced for rBGH-free milk in Ohio. Combined with the near total lack of price cutting or temporary price promotions (and the apparent price increases), the reliance on in-store displays indicates a strong willingness by producers to spend promotion dollars rather than risk any brand dilution by adjusting prices. It is unclear how permanent the forced de-labeling was seen as by producers, as almost immediately after its passing there was a legal challenge filed.[34] However, the ban lasted more than two years and during that time there was no legal support for the notion that

[33] See Schoenberger (Dec. 2000) Forbes article "Ca-Ching!"

[34] The 6th court of Appeals notes in its opinion (622 F. 3d 628) that both the International Dairy Foods Association and the Organic Trade Association filed almost immediately after the passing of the final rule.

the ban would be lifted, as the lower court sided with the state on all the claims.[35]

CONCLUSION

The sudden removal of a well-established product label in Ohio seems to have had a noticeable impact on the market demand for rBGH-free milk. However, the producer response seems to have focused primarily on the aggressive use of in-store display space, a dimension of analysis often missing from the labeling literature with its primary focus on price. The results from Kansas and Indiana show that the impact arose from the actual removal of the label, rather than any surrounding media attention to the issue.

There is some evidence that the effect of the ban in Ohio was cumulative over time, with both sales volumes falling and the use of displays increasing over each year the ban was in place. However, unlike Teisl's (2002) finding for dolphin safe tuna, here there is no observable "S' shaped information absorption dynamic, but rather a steady increase. It is possible that the gradual deterioration in rBGH-free markets was a result of a strong habit (low propensity to switch) among milk purchasers.

Given that liquid milk is often positioned as a "loss leader" for retailers (sold at low or negative margins to attract more footfall[36] into the store), low-loyalty shoppers likely would have self-sorted away from rBGH-free milk even before the ban. Further research leveraging customer-level panel data (including demographic information) could map out the ban's effect on such

[35] The de-labeling rule was challenged on several constitutional grounds, including infringement of First Amendment rights, violation of the dormant commerce clause and unconstitutional vagueness. The district court granted summary judgment for the state on all claims.
[36] Incresing the number of customers under the theory that once there, they will also purchase goods sold at a higher profit margin.

different customer segments. The current paper is only able to approximate for this dynamic by splitting the sales by *product* type into high or low price.

Given the continuing conflict between some industry actors and consumer advocates around issues of labeling, such as genetically modified organisms, carbon labels etc., a better understanding of the dynamic market response to such regulations is crucial. Economists must look at measures going beyond traditional demand models, or run the risk of severely biased estimates of the labeling effect.

REFERENCES

Bauman, D. E. (1992). Bovine somatotropin: review of an emerging animal technology. *Journal of Dairy Science*, *75*(12), 3432-3451.

Bracken, M. (2012). GMO labeling: California's Proposition 37 and federal constitutional implications. *The Arizona Journal of Environmental Law & Policy* 3(1) 1023-1027.

Bronnenberg, B., Kruger, M., Mela, C. (2008). The IRI Marketing Data Set. *Marketing Science* 27(4), 745-748

Butler, L. J. (1999). The profitability of rBST on US dairy farms. *AgBioForum* 2(2), 111-117

Grocery Manufacturers Association (2007). Shopper Marketing: Capturing a Shopper's Mind, Heart and Wallet.

Heien, D. M., & Wessells, C. R. (1988). The demand for dairy products: structure, prediction, and decomposition. *American Journal of Agricultural Economics*, *70*(2), 219-228.

Kiesel, K., Buschena, D., & Smith, V. (2005). Do voluntary biotechnology labels matter to the consumer? Evidence from the fluid milk market. *American Journal of Agricultural Economics*, *87*(2), 378-392.

Kiesel, K., & Villas-Boas, S. B. (2007). Got organic milk? Consumer valuations of milk labels after the implementation of the USDA organic seal. *Journal of agricultural & food industrial organization*, *5*(1).

Lesser, W., Bernard, J., Billah, K. (1999). Methodologies for Ex Ante projections of adoption rates for agbiotech products: Lessons learned from rBST. *Agribusiness*, *15*(2), 149-162.

Pesendorfer, M. (2002). Retail sales: A study of pricing behavior in supermarkets. *The Journal of Business*, *75*(1), 33-66.

Teisl, M. F., Roe, B., & Hicks, R. L. (2002). Can eco-labels tune a market? Evidence from dolphin-safe labeling. *Journal of Environmental Economics and Management*, *43*(3), 339-359.

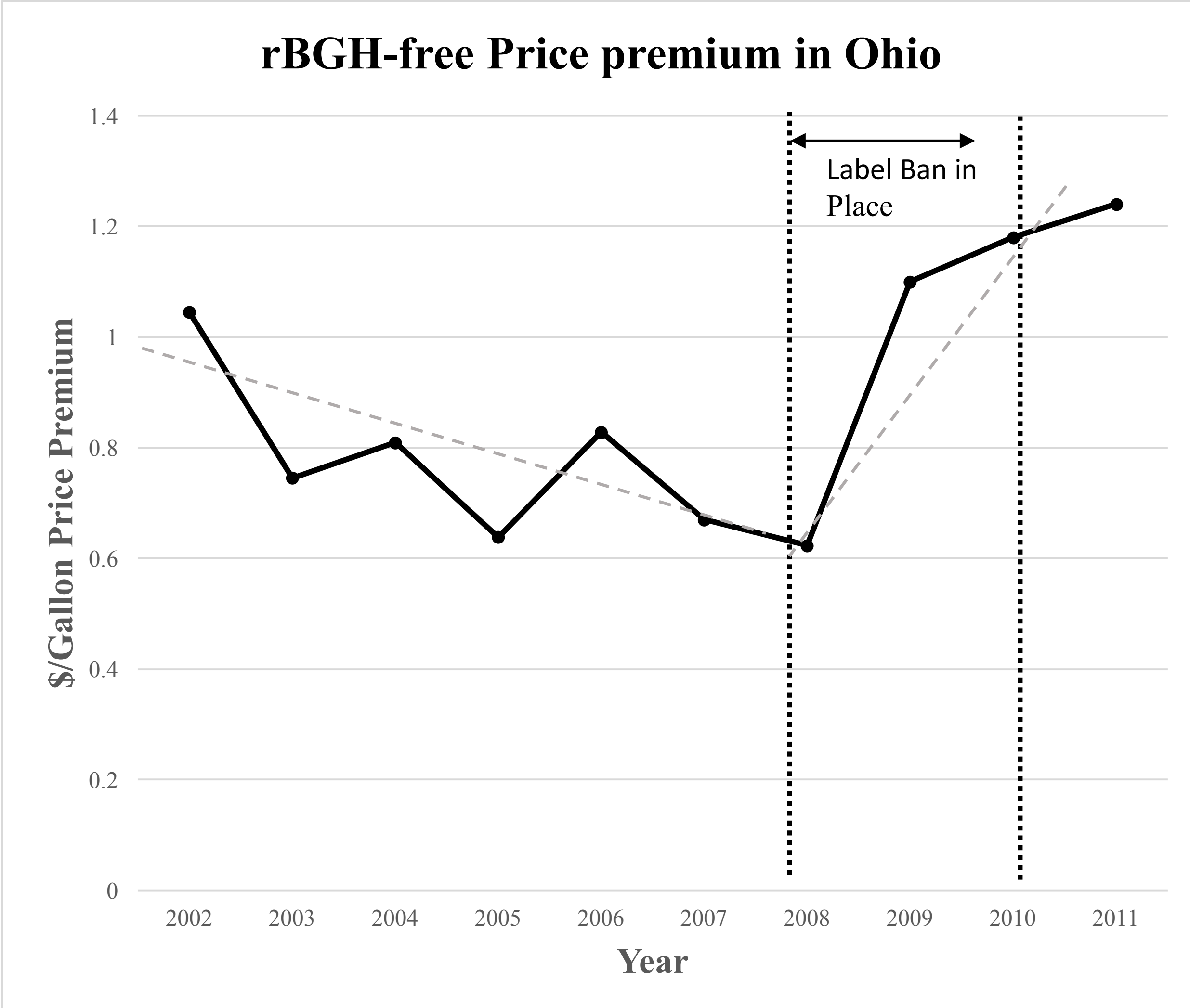
Vandenbergh, M. P., Dietz, T., & Stern, P. C. (2011). Time to try carbon labelling. *Nature Climate Change*, *1*(1), 4-6.

Varian, H. R. (1980). A model of sales. *The American Economic Review*, 651-659.

Yoo, B., Donthu, N., & Lee, S. (2000). An examination of selected marketing mix elements and brand equity. *Journal of the academy of marketing science*, *28*(2), 195-211.

APPENDIX A

FIGURE 4. EXAMPLES OF TYPICAL rBGH LABELS



Note: The premium tracked is defined as the average price per gallon of all milk ex. rBGH-free, subtracted from the average price of all rBGH free milk. The dashed-grey line breaks the overall trend to two phases: 2001-2007 (prior to the ban), and after 2008.

FIGURE 5. INTERACTION TERMS: PRICE PREMIUMS IN OHIO FOR rBGH FREE MILK

APPENDIX B

TABLE 3. MILK SALES REGRESSIONS (BY VOLUME, PRICE AND DISPLAY USE)

	(1) Volume (OLS)	(2) Price (OLS)	(3) Display (Probit)
	gallons/wk	*$/gallon*	*% sales w. display*
rBGH-free	113.575	0.0693*	-0.1229*
Ohio	1729.821*	-0.1595*	0.5221*
Kansas	-2627.868*	-0.2610*	0.6340*
Indiana	1280.033*	-0.2737*	0.5693*
Ohio x 2008 x rBGH-free	-19162.714*	0.6232*	1.0131*
Ohio x 2009 x rBGH-free	476.801	1.1016*	3.3235*
Ohio x 2010 x rBGH-free	99786.075	1.1765*	3.5684*
Ohio x 2011 x rBGH-free	103532.702	1.2429*	6.0775*
Kansas x 2008 x rBGH-free	3751.822*	-0.4161*	0.0767
Kansas x 2009 x rBGH-free	4426.881	-0.5757*	0.4432*
Kansas x 2010 x rBGH-free	203.396	-0.3998*	-0.0045
Kansas x 2011 x rBGH-free	-436.006	-0.4046*	0.4840*
Indiana x 2008 x rBGH-free	11526.490*	0.1304*	0.9107*
Indiana x 2009 x rBGH-free	-1530.393	0.1112*	-1.1778*
Indiana x 2010 x rBGH-free	-2854.841	0.1501*	-2.0674*
Indiana x 2011 x rBGH-free	-2625.608	0.0045*	-2.1428*
Time Controls (2002-2011)	x	x	x

* Significant at a 1% level ** Significant at 5% † Significant at 10%

Notes: 2008-2010 are the years the ban is in place in Ohio, while 2011 is the first year with relabeling. The omitted category for all years (including interactions w. years) is 2002, while the omitted category for states is everything ex. Ohio, Kansas and Indiana. Source: IRi Consumer Data 2001-2011. Running model with errors clustered by individual stores does not significantly influence results.

TABLE 4. OLS PRICE REGRESSION WITH PREMIUM SEGMENTATION

	Price (OLS)
	$/gallon
rBGH-free	-0.958*
Ohio	-0.18*
Kansas	-0.041*
Indiana	-0.275*
Ohio x 2008 x rBGH-free	-0.2905*
Ohio x 2009 x rBGH-free	0.3265
Ohio x 2010 x rBGH-free	-0.1225*
Ohio x 2011 x rBGH-free	0.307*
Kansas x 2008 x rBGH-free	-0.259*
Kansas x 2009 x rBGH-free	-0.381*
Kansas x 2010 x rBGH-free	-0.222*
Kansas x 2011 x rBGH-free	-0.276*
Indiana x 2008 x rBGH-free	-0.238
Indiana x 2009 x rBGH-free	-0.015
Indiana x 2010 x rBGH-free	0.021
Indiana x 2011 x rBGH-free	-0.102*
Premium rBGH-free x Ohio x Ban (2008-2010)	-0.32
Time Controls (2002-2011)	x
Premium Segmentation	x

* Significant at a 1% level ** Significant at 5% † Significant at 10%
Notes: The premium segmentation model looks only at rBGH-free milk sold at a price over $2.8/gallon, i.e. >5% premium over the average cost of regular milk. Source: IRi Consumer Data 2001-2011.

Chapter 3: Regulated Label Premiums and Labeling "Leakage": USDA Organic Labels and Natural Products in Packaged Foods

Abstract

The primary motivation for a producer to label their goods "organic" or "natural" is to increase consumer demand (through elevated willingness-to-pay). The organic label became centrally regulated by the USDA in 2001, while natural labels continued to have much less stringent requirements. This divergence created the possibility for both an increased premium for the now regulated, and thus more uniform (and perhaps credible) organic label, as well as an occurrence of significant "leakage" into the less regulated natural labels (and thus a reduced premium for the regulated label), with potentially equal appeal to consumers. To capture any such effects on the market accurately there must be controls for the shifting product mix in the market, as well as the endogenous relationship between price and sales volume. This paper utilizes a like-for-like model holding the product set fixed over time, as well as a total-market specification including new entrants to the market. This allows us to identify both the changes in price and volume, and decompose any labeling premium or leakage accurately. The findings support the need for such multi-faceted analysis, as several product categories examined yield mixed results depending on the model and metric chosen.

INTRODUCTION

Organic labels in the United States were unified under the National Organic Program (NOP) controlled by the United States Department of Agriculture (USDA) in 2001. In the "green" consumer goods market, a related, much more lightly regulated, category of "natural" products are also prevalent. Typical studies on the effects of such labels tend to focus on the immediate, short-term effects on the market of existing products, ignoring the important role played by new entrants into the marketplace, such as Wal-Mart's low-price organic range[1], a notable departure from the traditional "premium"-priced organic product offerings. A better understanding of the role of product changes is necessary to accurately estimate the potential problem of labeling leakage.

[1] See, e.g., Martin, Andrew (Nov 06, 2014) "Wal-Mart Promises Organic Food for Everyone" BLOOMGERG BUSINESS *available at* http://www.bloomberg.com/bw/articles/2014-11-06/wal-mart-promises-organic-food-for-everyone

This paper takes a two-pronged approach, addressing the policy concern of labeling "leakage"[2], while also raising the methodological issue of a changing product mix not addressed in typical studies on labeling effects. This paper highlights the importance of shifting product portfolio in organic and natural goods over time and identifies the demand-shift seen for pre-existing products (in a like-for-like model, tracking a fixed set through time). This is separated from the effects of new product entries (reflected in a total market model, including all products sold in a given year).

A regulator without any paternalistic motive to either increase (e.g. retirement savings) or reduce (e.g. tobacco) the consumption of a particular goods is merely interested in maximizing the efficiency of the markets. This might require regulating a label, allowing for a separating equilibrium to emerge where previously asymmetric information and/or a credibility problem caused a partial or complete unraveling of the premium market (Akerlof 1970), or even outright fraud being perpetrated on consumers.

In contrast, producers want to label their goods in order to increase their demand (through increased consumer willingness to pay). A common concern in the labeling literature is the potential for labeling "leakage", whereupon any public regulation of a particular label producers simply shift from the newly regulated categories to similar, less stringent labels.

If we assume public labeling regulations are weakly binding (zero or positive costs to producers), it follows that affected producers would opt out of the regulated category if a similar product positioning was possible with a net loss of profits smaller than the compliance and productions cost of the regulated label.[3] For example, if natural labels allow producers to capture 90% of the increase in a consumers' willingness-to-pay for organic products, while maintaining significantly cheaper production and compliance costs, several producers could simply shift from

[2] See Section II for detailed discussion
[3] Not all regulated labels can be avoided in this way: for example calorie labeling requirements impose a universal requirement on producers.

providing organic products into the “natural” goods market. For large producers with vast portfolios of products, it is exceedingly likely that at least *some* product lines might be optimally moved into the less regulated category, resulting in reduced societal benefits of any regulatory action.

Not all leakage should be considered negative. It is entirely possible that there are valuable product attributes that happen to be less regulated (e.g. carbon labeling), and producers moving from currently regulated labels to these categories could even result in a net gain for both producers and consumers. Concerns of leakage are typically linked to the idea of producers being able to “escape” a regulation and extract more of the value generated in a market, at an overall lower benefit to society. In environmental labeling this often takes the form of “greenwashing”[4], a practice where producers capitalize on a real consumer demand for environmentally friendly products by modifying their products, e.g. using misleading labeling and packaging without actually providing the desired environmental attributes.

A related phenomenon is that of affective marketing, referring to the practice of changing consumer preferences through emotional, i.e. uninformative or less than fully rational, appeals. An example of these phenomena playing out would be if a producer, instead of producing organic peanut butter simply called its existing peanut butter “natural”, while repackaging the product in green colors and a picture of squirrels frolicking in a lush forest. Both products are attempting to tap into a consumer’s demand for environmentally friendly goods, potentially resulting in roughly equivalent willingness-to-pay, but offering very different actual attributes.

Natural labels are often a mixture of actual informative content (to the extent that there are any actual restrictions on the use of the term, e.g. no use of laboratory-produced additives) and a greenwashing/affective marketing component in that consumers may mistakenly believe natural is broadly speaking equivalent to organic and pay a premium solely due to this mistaken belief. This

[4] See e.g., Ramus&Montiel (2005).

chapter will assume, on the basis of existing greenwashing literature[5], that "natural" labels are primarily made up of the latter, non-informative component.

Observing several product categories through the introduction of the USDA regulation in 2001 to 2010, this paper attempts to measure the emergence of any regulated label premium on the one hand, and the prevalence of leakage into the potentially greenwashing "natural" label on the other hand. Measuring market changes in both price and volume, as well as running separate models for a static like-for-like product comparison and a dynamic total market model, this paper is able to decompose the observed market effect into its constituent parts.

GREEN LABELS

Environmental information in consumer goods is one of the more significant expansion proposals for mandatory consumer disclosure, with potentially tremendous economic and environmental impacts. The overall market-size for "green" products alone[6] in the United States could currently be as high as $280 billion.[7] Initially, green consumer products appeared to be a temporary fad created by the surplus economics of ever-growing spending power, or by the energy saving attempts of consumers facing abnormally high oil and electricity prices.[8] However, even after the

[5] See e.g., Coppolechia (2010), and TerraChoice Report (2010), *available at* http://sinsofgreenwashing.com/findings/greenwashing-report-2010/

[6] In fact, environmental disclosure would affect virtually the entire consumer market, not just the segment already environmentally friendly. However, the market for green products is a useful approximation of how much consumers currently value the attributes that are disclosed.

[7] Dossey (2010) (Note that his estimate includes all "lifestyles of health and sustainability" products which is a broader category than just environmentally friendly).

[8] James R. Healey & Barbara Hagenbaugh, *Record fuel prices blow budgets,* USA TODAY, March 11, 2008, at B 1-2 (discussing the fact, mainstream Green advertising first began in the 1970s when a recession led to sky-high oil prices and a focus on environmental issues. The second wave has resurfaced now due to similar problems, with record-level fuel prices).

financial crisis of 2008, and the subsequent recession, as well as with oil prices as far apart as $150 to $50 per barrel, sales of green products have mostly increased.[9]

There are several dimensions of green available for consumers today at the supermarket, covering products that are among others organic, low-carbon, recyclable/made from recycled materials, and/or supporting conservation of wildlife habitat. A product's environmental benefits can manifest in many ways such as reduced costs (e.g. lower electricity bills), health and safety from absence of toxins, or convenience (e.g. fluorescent light-bulbs do not have to be replaced as often).[10] However, for environmentally responsible behavior to be profitable for companies they must be able to convey these characteristics to the consumer. Similarly, for environmentally harmful producers to survive they must keep this information from the consumers, or offer a price discount.

The proliferation of these pro-environmental products has created the potential for consumer confusion, and it has driven subsequent private and public efforts to enhance information disclosure through standardization. Such disclosure can take place through two broad categories: general information campaigns and educational outreach, or point-of-sale interventions through labeling. The consumer goods context makes labeling a more attractive option to both marketers and regulators. With a labeling, point-of-sale approach there is no need to pay to reach the consumers, nor do the marketers or regulators have to worry about the consumers' attentiveness as much. There is also no need—or attempt—to ensure information retention and salience up to the time of purchase. It is not surprising then that most of the marketing and regulatory effort has taken place in product packaging and labeling.

[9] Dossey, *supra* note , at 17.
[10] Ottman, et al.,(2006)(presents different positioning strategies taken in marketing by green producers).

The primary means for producers or regulators of products to convey a product's environmental impact is through labeling programs. Initiatives such as EnergyStar[11], Energy Guide[12], organic labeling[13], carbon footprints[14] etc. all rely on a point-of-sale exhibition of the product's attributes. The labels can be formatted as either "report card" type information disclosures, or a simpler seal of approval indicating the product has met some threshold requirements.[15] Much like nutritional labels, the notion is that consumers can, without having engaged in any prior research, evaluate the product in the store and adjust their shopping habits accordingly.

Environmental labeling brings up the interesting issue of competing disclosures. If the primary purpose of labels is information disclosure, not a desire to influence consumer behavior beyond that, there should be little harm in the proliferation of labels. Consumers can simply read what each label is signifying and adjust their behavior accordingly, accepting the disclosure proponents' arguments at face value. The proliferation should simply provide them with additional information, and therefore more empowerment than a narrower, uniform labeling system. Debates surrounding the criteria for certification within such narrow environmental labels (such as carbon footprints) illustrate that there are in fact several legitimate factors that different consumers may wish to consider, which in turn would justify the existence of competing labels.

Unfortunately, little work has been done on the empirical effects of label proliferation, and research is undoubtedly stifled by the lack of readily available data. The lack of empirical evidence has not dissuaded several authors from advocating for the need of a single, or at least more regulated,

---

[11] An EPA program started as a response to the Energy Policy Act of 1992, Pub. L. No. 102-486, 106 Stat. 2776 (1992).
[12] Labeling for Covered Products, 16 C.F.R. § 305.11 (2013).
[13] General Requirements for Accreditation, 7 C.F.R. § 205.501 (2013).
[14] *See* Weidema, et al. (2008) (discussing the carbon footprint program in the United Kingdom. Note that unlike the US energy labeling programs mentioned, carbon labeling is currently purely private and voluntary).
[15] Banerjee&Solomon (2003) (overview and taxonomy of US eco-labeling, primarily in the field of energy efficiency).

labeling scheme for environmental claims.[16] The conflict between decentralized, voluntary disclosure as opposed to centralized mandatory disclosure highlights the uncertain effect of government intervention on empowerment. Instead of merely adding information to the mix, government regulation could be forcing out privately provided information. Under a completely unregulated system, at least in theory, producers could counteract one distortionary disclosure with another. However, if the government intervenes and regulates the disclosure "market" for one dimension, it can effectively bar producers from offsetting these distortions.

The positive case for such regulations is often framed as reducing confusion, increasing the credibility behind a label and thus allowing for new markets to emerge and thrive where previously impossible. One way of representing this dynamic is through Akerlof's "market for lemons" (Akerlof 1970). Due to asymmetric information, the market for high-quality used cars collapses without a regulatory intervention. Consumers do not know if a particular car is good or bad quality, but do know there is a mixture of good and bad cars that is being offered. Therefore, they are only willing to pay a fraction of the price for a good used car (reflecting the probability the car they are bidding on happens to be good, i.e. the overall share of good cars in the market). This offer means sellers of quality used-cars are unwilling to sell their cars (at a fraction of their true value), resulting in a further skew of the mix of available cars to poor quality. This process iteratively leads to the collapse of the market for good used cars, bar a regulatory intervention such as a lemon law.[17]

Similarly, in environmental labeling markets, uncertainty about the meaning or true attributes of a labeled product would lead a consumer to only be willing to pay some fraction of their "true" environmentally-friendly product price, causing a similar unraveling of the market. A publicly regulated label acts as an explicit warranty: all goods carrying the label satisfy the specified

[16] *See*, *e.g*. Fliegelman (2010), Coppolecchia (2010).
[17] Whereby used car sales come with an implicit warranty (bad cars can be returned by buyer within a preset period after the sale).

environmental characteristics. Therefore, the intervention of a regulatory body can increase the willingness-to-pay for environmental goods, and help generate a separate market for organic products.

It should be noted that most of the market for environmental information disclosure is already heavily regulated. Not only are some types of disclosure required, such as the use of additives, but the private provision of other information is regulated as well. The Federal Trade Commission (FTC) has regulated the types of environmental claims producers can make since 1992.[18] Recently in October 2012 the FTC issued newly updated guidance specifying federal standards for advertising claims like "biodegradable".[19] Several states, such as California and Indiana, have already incorporated portions of these FTC Green Guides into their own laws.[20] Private pseudo-regulators have also stepped in, such as the National Advertising Division of the Council of Better Business Bureau ("NAD"),[21] which interprets and adjudicates matters related to the Green Guides, as well as state environmental-marketing laws. However, one product type that is both relatively commonly observed, and relatively lightly regulated is "natural" goods.

Regulated v. Unregulated Labeling: Greenwashing

This paper assumes that the benefits of an unregulated label accrue mostly to unscrupulous producers who are able to use the label to "greenwash" their product. In other words, the unregulated label product is in fact somehow inferior to the labeled product, while still presenting itself as possessing the broad qualities of the regulated label. This is a reasonable assumption in the narrow

[18] Guides for the Use of Environmental Marketing Claims, 57 Fed. Reg. 36,363 (Aug. 13, 1992).
[19] 16 C.F.R § 260: Guides For the Use of Environmental Marketing Claims: Adoption of Revised Guides, FTC File No. P954501.
[20] Coppolecchia (2010).
[21] NATIONAL ADVERTISING DIVISION, http://www.bbb.org/us/national-advertising-division/ (last accessed Dec. 18, 2015).

sense of "inferior" as any product fulfilling the criteria of the regulated product could be labeled under that scheme instead (assuming compliance costs are not the driving factor). However, several legitimate but lightly regulated labels exist (e.g. carbon footprint), thus not all leakage represents greenwashing.[22]

In our case, products labeled "natural" are presumed to have traits that bar them from using the organic label (e.g. use of pesticides during the production process), while still wishing to convey to consumers an image of similarly pure/holistic qualities. For many other types of product and/or labels these assumptions will not hold, e.g. comparing organic products (regulated) to carbon footprint labels (largely unregulated) does not satisfy the "similar underlying qualities" part of our definition for leakage. Some pairs will be more difficult to demarcate as leakage or competing disclosures due to the amorphous nature of the underlying "quality" to which consumers are responding.[23]

One result of unregulated labels when utilized in "greenwashing" might be a rebound effect (Greening et al. 2000) where consumers exhaust their environmental "budget" on the misleadingly marketed products and decrease other expenditures relating to environmental protection. Suppose a consumer is ready to spend fifty dollars more a month to buy environmentally friendly products, but spends thirty dollars of this on greenwashed products. The result is only twenty dollars spent on actual conservation efforts out of the potential fifty. Such an exhaustion of the consumers' willingness to pay for the environment seems plausible, as there are often significant budget constraints limiting the consumer's ability to increase expenditures even when more salient appeals and concerns come to

[22] Even in the case of natural products, the label does impose substantive requirements such as the absence of synthetic additives.

[23] For example, is "organic" primarily about not containing pesticides, or as is a widespread misbelief among consumers, superior nutrition, or even about an idyllic small farm? Depending on this the less regulated labels "locally grown", "no added sugar" or "triple-washed" might all or none be relevant to the quality sought.

their attention. Assuming a social preference for environmentally friendly goods (i.e. a paternalistic regulator), such a rebound undermines the usefulness of creating the regulated label in the first place.

Alternatively, the greenwashing could result in a spillover effect[24], where the salience of environmental issues to consumers is heightened by both the regulated and unregulated marketing efforts. This could lead to increased expenditure on "true" green products. If environmental cues become a part of consumers' everyday shopping experience, this could conceivably lead to more awareness of environmental issues in other parts of their lives as well. Here, rather than operating within a pre-set budget, the consumer responds to increased environmental appeals by simply spending more than they initially planned. This mechanism seems plausible in the context of broadening the base of environmental consumers, i.e. engaging previously passive consumers, who perhaps were not operating on the upper limit of their environmental-spending budget yet. Another context where a spillover could occur is boosting green expenditures by consumers currently only making "rational" environmental purchases, assuming they also have the capacity to relax their budgetary restrictions further.[25] In the case of competing products within the same product category such spillovers seem unlikely, and it is conceivable that "natural" products heighten the salience of environmental attributes enough to drive more consumers to organic goods. However, any such effect would likely be outweighed by the cannibalization of true organic sales by the pseudo-environmental natural products.

Thus the overall effect of labeling and greenwashing on "true green" spending will depend on the relative effects on each of these sub-populations of consumers, as well as their relative sizes. Conceivably in product markets with little pre-existing green market share even disingenuous green marketers could be paving the way for socially responsible companies by "activating" a previously

[24] Kals et al. (1999).

[25] Note that the model used in this paper does not rely on adjustable budgets.

apathetic segment of consumers. However, it seems likely that in aggregate the crowding-out or rebound effect, due to limited overall environmental budgets, will counteract any gains made by such spillovers.

Greenwashing can be especially appealing to companies since the demand primer — environmental concern — is supplied by governments and NGOs. Unlike in a typical product-marketing scenario, the company does not have to spend significant sums to generate the "want", but has a ready-made market. Effectively taxpayers and donors to charities might be subsidizing the for-profit companies that engage in greenwashing, resulting in a larger amount of faux-green merchandise than is socially optimal. Notably, any efforts to increase environmental concern would also have the unintended consequence of driving even more consumer expenditure to unscrupulous marketers, resulting in an ongoing leakage to inefficient products.

In the traditional context, the "free rider" problem relates to lack of contributions towards public goods.[26] For example, everybody arguably benefits from national defense or police expenditure, but simultaneously have an incentive to shirk from contributions and rely on others paying for it. With greenwashing, the free riders go a step further, and might actually be "stealing" some of the total contribution of others. Relying on governments and NGOs to retain the salience and urgent need for action on environmental issues, affective marketers can simply siphon some of the potential pro-environmental spending created by these public awareness campaigns into expenditure on their products. Note that there is another possibility, where the affective marketers and greenwashers are actually also promoting genuine green spending. Whether greenwashing constitutes detrimental free riding, or ends up indirectly helping the social goals it is exploiting depends on how consumers budget their total environmental expenditures and is beyond the scope of this paper.

[26] Clarke (1971).

Suffice it to say, leakage into unregulated label alternatives is a serious concern for the overall health of both the regulatory framework, and environmental issues in general.

Since the FTC regulates the "market" for information in advertising, companies relying on information in their advertising have their hands tied and become potentially liable for their marketing strategies. Companies using such an informational strategy might have to go through onerous certification processes, perhaps even several times as the requirements might change over time. For example, the FTC Green Guides are not rules or regulations, and while under the 2012 standards a product may be called "biodegradable" if it breaks down within a year, the FTC could easily limit this further to six months in a subsequent guide with little forewarning. Relying on regulated information thus leaves companies both with more regulation and more uncertainty over the long-term viability of their strategy.

On the other hand, companies that choose to green-wash through less regulated labels face no increased scrutiny from either the FTC or consumers (if they are unaware of the manipulation), and would thus minimize the risk for legal liability or consumer backlash. Further, bar an unprecedented regulation of labeling and packaging more generally, producers need not worry about a moving regulatory standard for their claims, as "natural" can be mimicked through e.g. "wholesome" "natures' goodness" and similar labels, or even non-verbally through the use of green packaging, recycled materials etc. The market experience so far suggests that consumers are not becoming less sensitive to such manipulations either, as affective marketing strategies have been utilized successfully for several decades, and greenwashing has gone largely unpunished by consumers in the marketplace (or the FTC through regulatory action) for at least three years.[27]

[27] The time between the first TerraChoice report in 2007 and the latest in 2010. See TerraChoice, The Sins of Greenwashing Home And Family Edition 2010: A Report on Environmental Claims Made in the North American Consumer Market (2010)

DATA

This paper uses the iRI Academic Data set, consisting of individual product sales observations from a selection of grocery and drug stores, grouped weekly from 2001 to 2010 (this time period is especially conducive to our analysis due to USDA's creation of their organic label in 2000). For each sales event the exact good sold (SKU)[28], volume of goods sold, the price, and whether the product was being promoted at that store during the time of the sale (e.g. aisle-end displays). The dataset contains some 30+ product categories, however several of these are not edible (razor blades, household cleaners etc.). Of the 15 edible product categories, three contain reliable details on the products' characterization as natural and/or organic: peanut butter, spaghetti sauce and hot dogs. For this chapter, all variations of "natural" labels are included, for example "all-natural" and "100% natural". The prevalence of these features in the total product offering (i.e. SKU-share) is detailed below:

[28] Stock-keeping unit (SKU), the unique identifier for a product, e.g. two different flavors of the same brand peanut butter, or two different sizes of jars would each have their own unique identifier.

TABLE 1. ORGANIC/NATURAL SKU-SHARE WITHIN PRODUCT CATEGORIES

	Organic SKU-share		Natural SKU-share	
	2001	2011	2001	2011
Peanut Butter (n=513)	9%	12%	20%	18%
Pasta Sauce (n=5215)	3.2%	0.3%	14.7%	12%
Hot Dogs (n=1865)	0.2%	0.3%	4%	5%

Notes: Natural category includes variations of the label such as "100% Natural" "All Natural" etc.

Source: irI Academic Data Set, product stub files.

SKU-share allows us to measure the product mix in the market. The 20% share for natural peanut butter in 2001 (see Table 1 above) indicates that one out of five available choices for peanut butter were labeled natural.[29] The product categories included included a range of dynamics from very prevalent "green" markets in peanut butter (organic and natural products jointly account for approximately 30% of products), to a marginal but growing market in hot dogs. The shifting product portfolio is apparent in, e.g., the three percentage point increase in SKU-share of organic peanut butter, and a similar decrease in pasta sauces. Note that while the joint market share of both green categories does not shift much (increase of <1% in peanut butter and hot dogs, and decrease of

[29] Note that SKU-share does not measure the *volume* in each product category, as a large proportion of sales can be driven by only a few popular products. In 2001 natural peanut butter only had an approximate 8.5% market share of total sales, less than half its SKU-share.

approximately 5% in pasta sauce), there are significant shifts in the mix between organic and natural goods.

The data yields roughly six million annual unique observations in each category, resulting in over 30,000 data points in even the most thinly populated category of interest (organic hotdogs and natural pasta sauce).

MODEL & METHODOLOGY

Two models are applied to each product category: a like-for-like static SKU (only looking at products sold already in 2001, not updating for newer products) and a dynamic "total market" model including all products in the category regardless of their year of introduction.

Like-for-Like (LfL)

The rationale for using this model is to control for the possibility of the product mix offered under natural and organic labels changing over time. For example, if cheaper store-brands enter the organic market during the time observed here, the *average* price premium for organic products will decrease. However, the original "premium" organic brands might keep their price premium or even enjoy an increase that would be masked by the newcomers in the total market analysis. By restricting the model to the products that are initially available we are able to separate this product-mix effect and more accurately assess the change in popularity and pricing power due to potential leakage.

Total Market (TM)

Labeling leakage could primarily occur during the timeframe studied here, with the number of "natural" products being offered increasing each year (as is the case in the hot dog market). To capture this leakage, we must adjust the products in both the organic and natural category each year to reflect such entries/exits. This approach is not able to identify the effect on existing producers accurately, but represents the overall market much more accurately.

Suppose we observe a decrease in the price premium for natural products in the like-for-like model, while observing an increase in the total market model. This would indicate a strengthened market position for natural products generally, combined with the entry of several lower-cost alternatives into the natural market. See the table below for a comprehensive guide to interpreting results:

TABLE 2. INTERPRETATION OF RESULTS: LIKE FOR LIKE VS. TOTAL MARKET

	Relative Magnitude of Effect	Absolute Magnitude of Effect	Interpretation
I	Like-for-like >Total Market	LfL<0, TM<0	Loss of pricing power/market share, exacerbated by entry of lower price products in the mix.
II	Like-for-like > Total Market	LfL>0, TM>0	Increase in pricing power/market share, partially muted by entry of lower price products/exit or products in mix.
III	Like-for-like < Total Market	LfL>0, TM>0	Some increase in pricing power/market share, further amplified by entry of higher price products/new products in mix.
IV	Like-for-like < Total Market	LfL<0, TM<0	Loss of pricing power/market share, partially mitigated by entry of higher price products/new products in mix.
V	Like for-like < Total Market	LfL<0, TM>0	Loss of pricing power/market share

			completely masked by entry of higher price products/new products in mix.
VI	Like-for-like> Total Market	LfL>0, TM<0	Increase in pricing power/market share, completely masked by entry of lower price products/exit of products in mix.

The potential difference in results between the like-for-like and total market models also acts as a useful proxy for the extrapolative value of typical studies on labeling, as these tend to focus on a short time period holding the product mix constant (i.e. forcing a like-for-like type analysis). Should we see total market results that are vastly different from the like-for-like model, this suggests such studies are relatively inaccurate in predicting the *true* market effects of a labeling change, once producers have had a chance to respond to the new environment.

Further, in interpreting our results we need to consider the endogenous relationship between price and sales. We observe a single point on the price/volume slope, where by lowering prices a producer could sell more of the product at a lower profit per unit, or by raising prices a producer would sell fewer units but at a higher profit for each unit. The exact shape of this slope, and the producer's chosen trade-off between price and volume depend both on the consumers' demand price elasticity (for each x% change in price, what is the change y% in volume demanded), and the producers' cost functions (at what level is the marginal cost of a product sold equal to the marginal revenue). Note that in a perfectly competitive market the long run supply function is a flat line set at the producers' cost (i.e. there is no repositioning on the line). However, the existence of a higher priced (compared to conventional goods) natural goods market, and an even higher priced organic market suggests there are no perfect substitutes, and the market has some monopolistic characteristics.

As we are unable to observe the unrealized combinations of price and volume, we are limited to a cruder method of measuring the underlying market position of the natural and organic goods. They can be seen as either absolutely deteriorating (price *and* volume decreasing), absolutely improving (price *and* volume increasing), or a mixed market indicative of producer repositioning (price and volume changes have opposite directions). See Table 3 below for a full representation of the possible combinations observed:

TABLE 3. INTERPRETATION OF RESULTS: SALES V. PRICE CHANGES

Change in Volume	Change in Price	Interpretation
Greater than zero	Greater than zero	Increased demand
Greater than zero	Less than zero	Mixed-repositioning through discount
Less than zero	Greater than zero	Mixed-repositioning through niche strategy
Less than zero	Less than zero	Decreased demand

The discussion above highlights the complexity of analyzing changes in the marketplace due to a labeling change. Studies merely looking at the market share or price premiums in isolation can confound and misinterpret the underlying change in the markets if the period observed is long enough to allow for producer responses through pricing or product portfolio decisions. The likelihood of unrepresentative results is heightened due to the common practice of focusing on a single product category, and extrapolating these effects to the overall market as even in overall relatively stable product markets, any particular category is likely to experience a shift in the product mix available. This paper explicitly breaks down the total market effect observed into its constituent parts, allowing us to capture the potential effects of labeling regulation and leakage more accurately.

Results

The results show strong evidence of labeling leakage taking place, especially in the spaghetti sauce category, and a noticeable difference between the like-for-like and total market models used. Further, across the three product categories there is evidence of a non-existent or smaller price premium for natural products, and a significant premium for organic products (see Table 7 in Appendix A). Due to the extensive sample (millions of observations for each year) being used, almost all results are statistically significant.[30]

Peanut Butter

In the peanut butter market, we observe a very high starting level of both organic and natural products, with a shift in the product mix towards more organic goods by the end of the sample in 2010. The results (see Figures 1 and 2) show the transformation of the organic peanut butter market from predominantly discounted "value" products in 2001-2002 (up to a 50 cent discount, or roughly 15% off the average total price of peanut butter) to a more premium niche by 2007 (average total market premiums up to 55 cents, or over 17% of the total average price, see Table 4). Year 2008 and later years see a sharp reversal of this pricing power, for both the like-for-like comparison and total market, perhaps capturing the effects of the financial crisis in 2008 and the subsequent recession. Interestingly after 2008 the like-for-like price model produces higher premiums, which could indicate the introduction of more "value" oriented organic brands into the mix again in a sharp reversal of the changes between 2001 and 2007.

[30] The notable exception is organic hotdogs, where the miniscule marketshare of those products (<0.1%) results in smaller sample sizes, and very large standard errors.

In the peanut butter market, a static like-for-like study would vastly underestimate the positive market growth for organic goods, as the like-for-like product mix lags the total market both in terms of price premia and to a lesser degree in sales volume. This suggests producers had a dynamic response to the new labeling scheme and consumer demand, and significantly adjusted the product mix offered to the market, as evidenced by the growing SKU-share for organic products (see Table 1 above).

Natural peanut butter seems to have no consistent price premium during the whole time studied here, and even has a slight (<5% of total price) overall discount to the overall market. However, the price metrics are missing an immense increase in the sales for natural products at a steady price-point, indicating a strong increase in demand for the product. Unlike organic peanut butter, here there is no evidence of significant changes in the product mix offered: SKU-share actually shrunk from 2001 to 2010 and comparison of the like-for-like and total market coefficients reveals no obvious re-pricing or introduction of popular new products.

Overall, there is some evidence of labeling leakage in the peanut butter markets, with natural sales volumes easily outpacing organic goods. However, while for most of the period under study (2003-2007) organic peanut butter enjoyed a significant price premium (driven by entrants into the market after the 2001 USDA label regulation), no such premium is captured by the "greenwashed" natural products.

Spaghetti Sauce

Spaghetti sauce sales follow a similar (see Figures 3 and 4), albeit slightly different pattern: after an initial dip in prices for both natural and organic products to below the national average pricing, organic products start to capture a higher price premium on both a like-for-like and total market basis. On the volume side, however, we see the growth dominated by new natural products (with cumulative growth three to four times the size of organic products) even in the face of sharply falling sales for the portfolio of natural products that existed already in 2001. Accounting for the drag on the total natural sales caused by this initial portfolio, the growth coming from new natural products is nearly ten times as much as the total growth in organic spaghetti sauce.

The growth in volume of natural markets is especially striking when considering that natural products enjoyed on average a higher premium (or lower discount) than organic products. The overall average across years for natural products is a discount of 10 cents (approximately 4% of the total price), while organic products sold at an average discount of 20 cents (8% discount)(see Table 5). Even in the later years with increased premia for both types of products, we can see the total market for organic products achieving a meaningful premium over natural products in 2009, and to a lesser extent in 2007.

Partially this is due to the introduction of lower-price organic products into the mix, evidenced by the significantly higher like-for-like figures. This is especially true after 2007, when the average like-for-like premium was 15 cents higher. This difference translates to roughly 6% of the average total price, and in some years more than three to four times the total market price premium). Natural products also saw an influx of lower priced products into the mix, which seem to have been the primary driver of the increased sales (as discussed above). Similar to organic products, the average like-for-like premia are approximately 15 cents higher than the total market, but due the overall higher premiums in the natural category, this represents a smaller percentage difference between the two.

These results suggest that for the peanut butter category, both organic and natural labels saw a strengthened market position from 2001 to 2011. Both categories enjoyed growing price premia, while also increasing sales in this time period. A short-term like-for-like model would have significantly underestimated the leakage into natural goods, while overstating the growing sales and prices for the organic market. As was the case in peanut butter, here too we observe significant evidence for leakage, with natural sales growth outpacing the organic market by a comfortable margin. The case for leakage is further bolstered by the higher price premiums. Whereas in the peanut butter markets only organic goods enjoyed positive premiums, here after 2005 natural products have a 6-8% premium over the total average price, which is comparable or even higher than the average organic premiums.[31]

Hot Dogs

Both sales prices and volumes (see Figures 5 and 6) seem to have been significantly boosted by new entrants into the market (total market effects are higher than like-for-like, especially after 2006). Out of all categories observed the price effects in hot dogs are by far the most dramatic, with the final average price premium for organic goods reaching $1.4 by 2010, or roughly 50% over the total average market price (see Table 6). However, due to the very small market share of organic hot dogs (approximately 0.3% in 2010, less than 6,000 observations) none of the organic price effects are statistically significant.

For natural hot dogs, a smaller 20 cents (6% premium over total average price) price premium emerges by 2010, and is statistically significant. This effect seems to be driven entirely by new entrants, as the like-for-like 2001 portfolio of natural hot dog products consistently sold for a discount to the average price after 2006.

Again, we find strong evidence of labeling leakage, as the overall sales growth for natural products is three to four times higher than for organic products. While organic products potentially enjoy higher price premiums, this effect was not statistically significant. This leads us to

[31] Note that unlike with peanut butter, the "green" natural and organic markets in spaghetti sauce do not seem to have been as dramatically affected by the macroeconomic shock and recession after 2008.

conclude that the overall market position for natural goods improved more than that for organic alternatives after the 2001 USDA labeling regulation. Further, assuming a static product mix (like-for-like model) significantly underestimates the extent of this leakage, as well as in this case potentially the boost in demand for organic products.

CONCLUSION

The effect of a newly regulated label, and any ensuing label "leakage" in a dynamic market, is nuanced and not easily captured by simple models looking at any one variable (price, sales etc.). Producers might respond by re-pricing their offerings in the short-run, and introducing new products into the marketplace in the medium-to-long run, while withdrawing less successful offerings.

Apparent effects of a label picked up in any particular dimension are not necessarily reflective of the overall market shift, for example in the hot dog market a like-for-like methodology (implicit in typical studies) would significantly underestimate the demand increase for both organic and natural goods. In the peanut butter market, the like-for-like methodology would overestimate the overall premium paid for organic products, due to the entry of low-cost organic alternatives. Finally, perhaps the starkest example of the limitations of assuming a static product mix comes from the spaghetti sauce market, where the like-for-like model misidentifies a significant boom in the natural goods market as a *weakening* due to the driving role of new product entries.

This chapter set out to model the overall market response in an attempt to capture any potential regulatory leakage effect fully. The product categories studied are not a representative sample of the overall consumer goods market affected by the USDA regulations (for example, excluding any fresh produce, which potentially has different dynamics due to different levels of packaging/labeling space available to producers).

Further, "natural" products are conceivably more attractive in packaged goods with the possibility of preservatives and other additives[32], in fresh goods the leakage might well be more driven by other labels like "family owned", "locally grown" etc. The product mix chosen represents a variety of markets: those with large "green" segments as peanut butter, those dominated by less regulated natural labels (pasta sauce), and largely conventional markets witnessing an increase in demand for natural products like that for hot dogs.

The results clearly demonstrate the need for more nuanced modeling, rather than relying on headline price and market share figures. Further research is needed to map out the contours of any particular label change/leakage, or a specific product type (foods v. cosmetics, pantry items v. fresh produce). However, all three categories examined here produce consistently strong evidence for significant "leakage" into the natural category. Further, all three categories show organic goods on average retaining a "premium" niche category, with natural goods largely positioned at lower price premiums, and even at a discount to conventional goods. Analysis of such results is needed to better understand the potential benefits of publicly regulated consumer labels, and the threat posed by "leakage" into less regulated alternatives.

[32] For example, hot dogs are a product traditionally associated with the use of less-than desirable parts of animals, but also sodium nitrates, MSG and other "unnatural" additives.

REFERENCES

Akerlof, G. A. (1970). The market for" lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 488-500.

Banerjee, A., & Solomon, B. D. (2003). Eco-labeling for energy efficiency and sustainability: a meta-evaluation of US programs. *Energy Policy*, *31*(2), 109-123.

Clarke, E. H. (1971). Multipart pricing of public goods. *Public choice*, *11*(1), 17-33.

Coppolecchia E.K. (2010), *The Greenwashing Deluge: Who Will Rise Above the Waters of Deceptive Advertising?*, 64 U. MIAMI L. REV. 1353.

Dossey, L. (2010). LOHAS market size: US consumer sales soar close to $300bn. *LOHAS Journal 2010*, 16.

Fliegelman J.E. (2010), *The Next Generation of Greenwash: Diminishing Consumer Confusion Through a National Eco-labeling Program*, 37 FORDHAM URB. L.J. 1001.

Greening, L. A., Greene, D. L., & Difiglio, C. (2000). Energy efficiency and consumption—the rebound effect—a survey. *Energy policy*, *28*(6), 389-401.

Hartman, C. L., Ottoman, J. A., & Stafford, E. R. (2006). Avoiding green marketing myopia: Ways to improve consumer appeal for environmentally preferable products. *Environment*, *48*(5), 22-36.

Kals, E., Schumacher, D., & Montada, L. (1999). Emotional affinity toward nature as a motivational basis to protect nature. *Environment and behavior*, *31*(2), 178-202.

Ramus, C. A., & Montiel, I. (2005). When are corporate environmental policies a form of greenwashing?. *Business & Society*, *44*(4), 377-414.

Weidema, B. P., Thrane, M., Christensen, P., Schmidt, J., & Løkke, S. (2008). Carbon footprint. *Journal of Industrial Ecology*, *12*(1), 3-6.

APPENDIX A

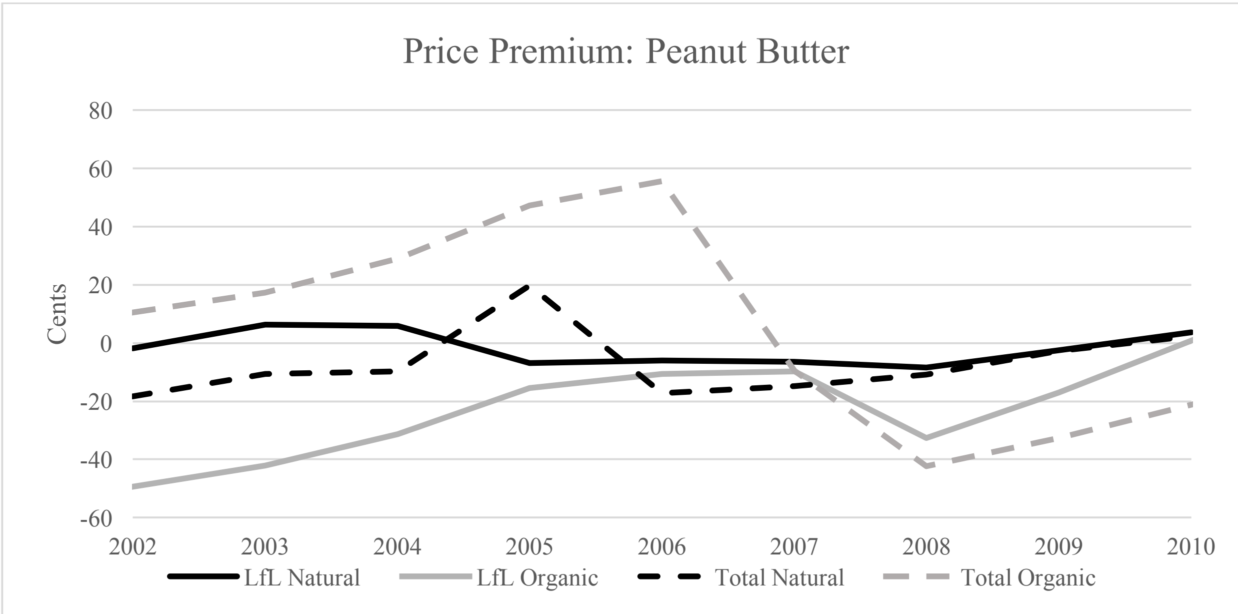


FIGURE 1. PEANUT BUTTER PRICE PREMIUM 2002-2010
Source: iRI Academic Data Set 2001-2010

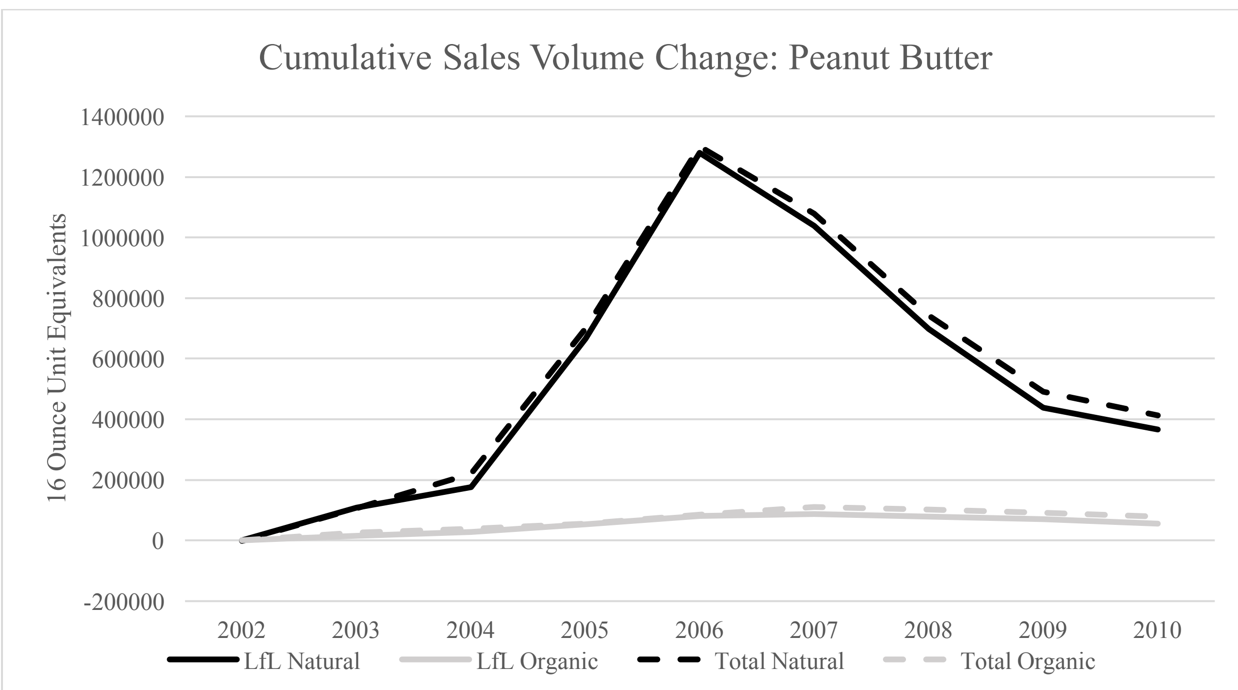


FIGURE 2. PEANUT BUTTER SALES 2002-2010
Source: iRI Academic Data Set 2001-2010

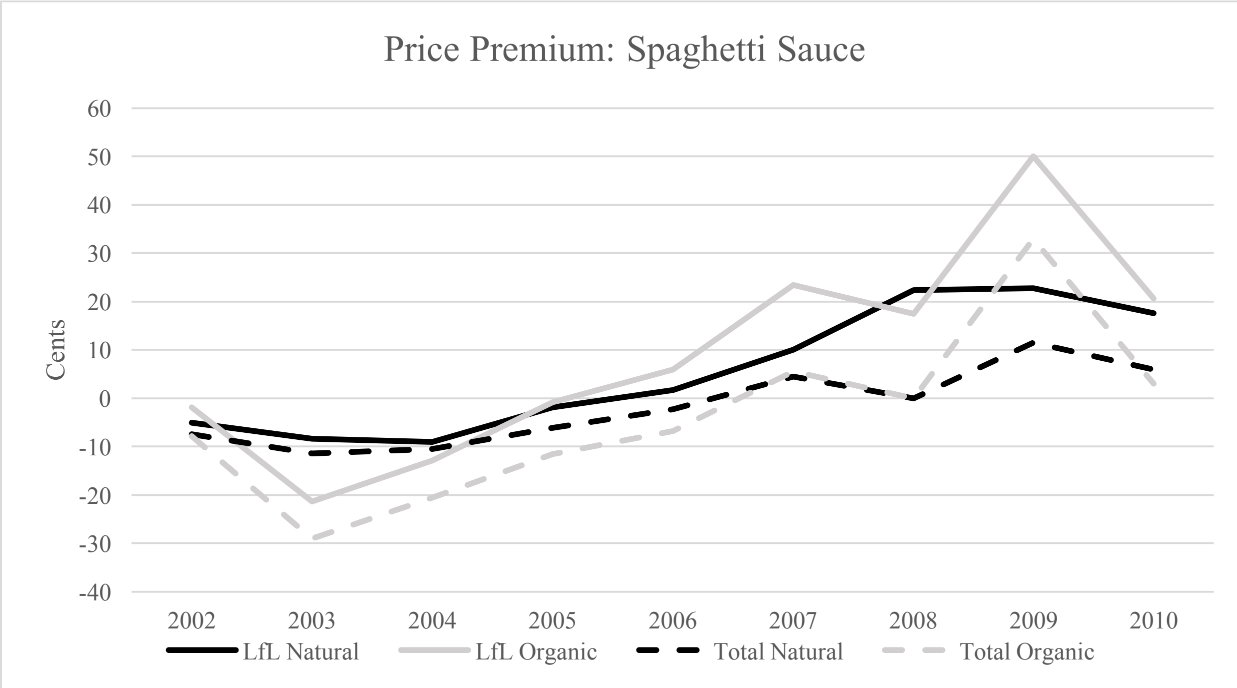


FIGURE 3. SPAGHETTI SAUCE PRICE PREMIUM 2002-2010
Source: iRI Academic Data Set 2001-2010

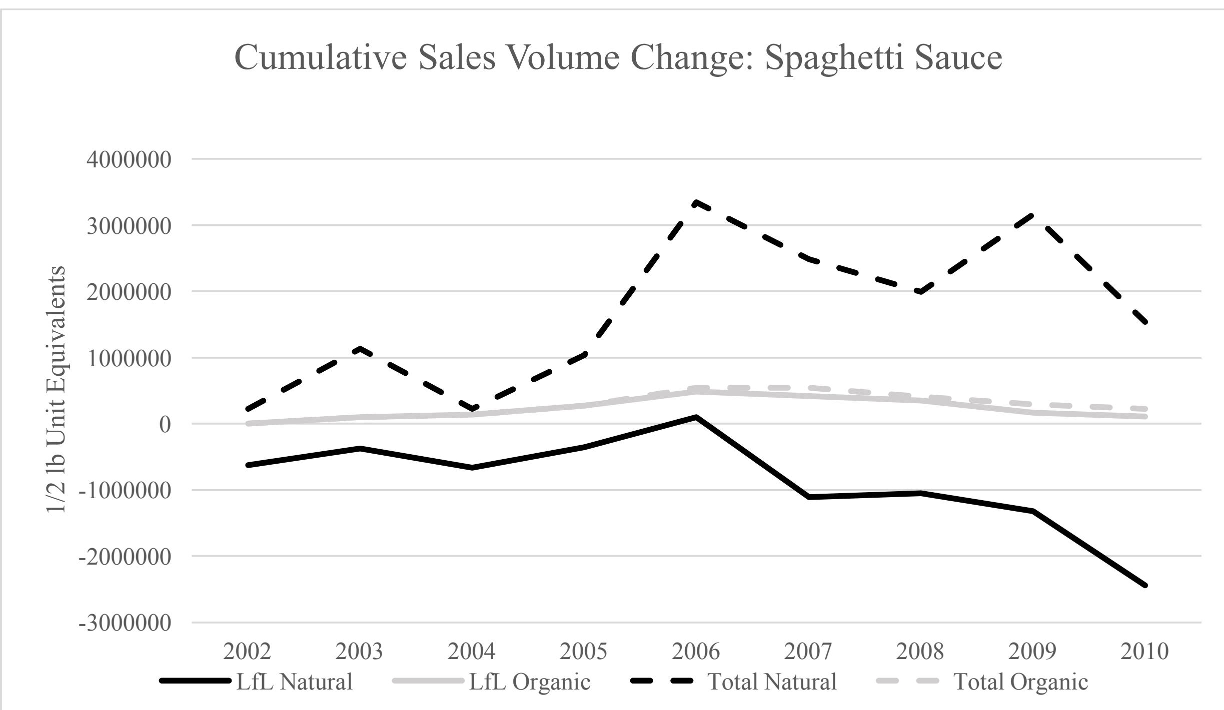


FIGURE 4. SPAGHETTI SAUCE SALES 2002-2010
Source: iRI Academic Data Set 2001-2010

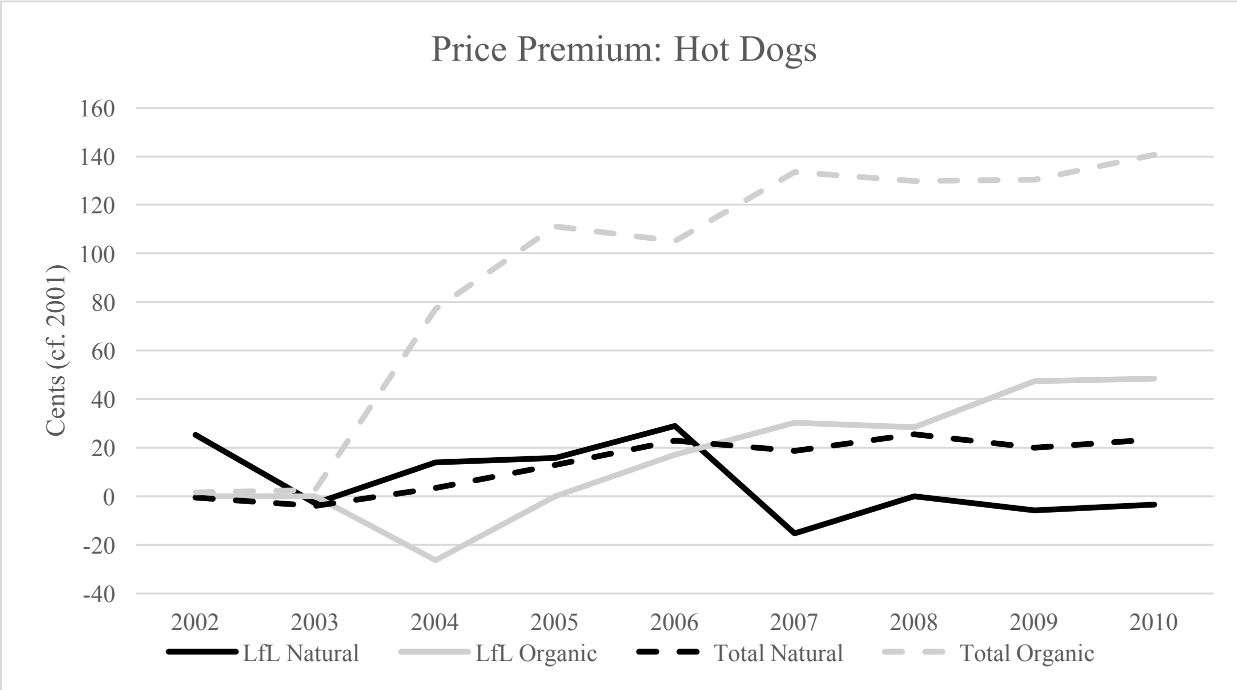


Figure 5. Hot Dog Price Premium 2002-2010
Source: iRI Academic Data Set 2001-2010

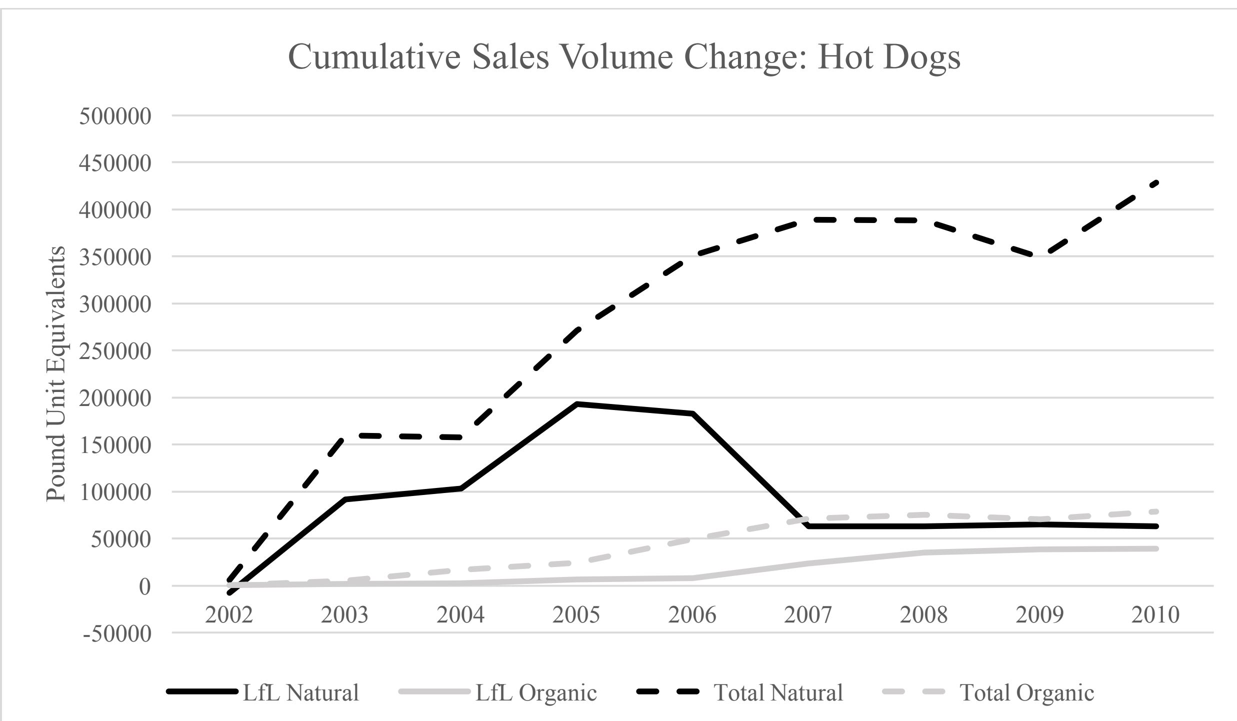


Figure 6. Hot Dog Sales 2002-2010
Source: iRI Academic Data Set 2001-2010

TABLE 4. OLS REGRESSION: PEANUT BUTTER PRICE PREMIUM (DOLLARS)

	Peanut Butter	
	Like-for-Like	Total Market
Natural	-0.076	0.284
Organic	1.762	1.349
Natural x 2002	-0.018	-0.183
Natural x 2003	0.062	-0.106
Natural x 2004	0.059	-0.098
Natural x 2005	-0.070	0.197
Natural x 2006	-0.060	-0.172
Natural x 2007	-0.064	-0.147
Natural x 2008	-0.084	-0.109
Natural x 2009	-0.026	-0.028
Natural x 2010	0.036	0.024
Organic x 2002	-0.493	0.105
Organic x 2003	-0.422	0.172
Organic x 2004	-0.312	0.290
Organic x 2005	-0.154	0.472
Organic x 2006	-0.106	0.555
Organic x 2007	-0.099	-0.095
Organic x 2008	-0.327	-0.423
Organic x 2009	-0.170	-0.326
Organic x 2010	0.008†	-0.211
Time Controls (Year)	x	x

Note: All values are statistically significant at the 1% level, except those marked by † which all have p>0.25

Source: iRI Academic Data Set 2001-2010

TABLE 5. OLS REGRESSION: SPAGHETTI SAUCE PRICE PREMIUM (DOLLARS)

	Spaghetti Sauce	
	Like-for-Like	Total Market
Natural	-0.160	0.022
Organic	1.177	1.378
Natural x 2002	-0.050	-0.074
Natural x 2003	-0.084	-0.114
Natural x 2004	-0.090	-0.105
Natural x 2005	-0.019	-0.061
Natural x 2006	0.016	-0.023
Natural x 2007	0.101	0.045
Natural x 2008	0.224	0.067
Natural x 2009	0.227	0.115
Natural x 2010	0.176	0.059
Organic x 2002	-0.019	-0.079
Organic x 2003	-0.213	-0.291
Organic x 2004	-0.128	-0.205
Organic x 2005	-0.008†	-0.115
Organic x 2006	0.059	-0.068
Organic x 2007	0.234	0.055
Organic x 2008	0.175	0.148
Organic x 2009	0.500	0.329
Organic x 2010	0.206	0.030
Time Controls (Year)	x	x

Note: All values are statistically significant at the 1% level, except those marked by † which all have p>0.25
Source: iRI Academic Data Set 2001-2010

TABLE 6. OLS REGRESSION: HOT DOGS PRICE PREMIUM (DOLLARS)

	Hot Dogs	
	Like-for-Like	Total Market
Natural	1.137	0.185
Organic	2.119	0.922†
Natural x 2002	0.254	-0.006
Natural x 2003	-0.030	-0.041
Natural x 2004	0.141	0.034
Natural x 2005	0.160	0.130
Natural x 2006	0.290	0.230
Natural x 2007	-0.154	0.186
Natural x 2008	-0.017	0.256
Natural x 2009	-0.058	0.201
Natural x 2010	-0.033	0.235
Organic x 2002	-0.046†	0.016
Organic x 2003	-0.173†	0.027
Organic x 2004	-0.263†	0.772†
Organic x 2005	-0.035†	1.111†
Organic x 2006	0.174†	1.052†
Organic x 2007	0.302†	1.335†
Organic x 2008	0.285†	1.300†
Organic x 2009	0.476†	1.304†
Organic x 2010	0.485†	1.407†
Time Controls (Year)	x	x

Note: All values are statistically significant at the 1% level, except those marked by † which all have $p>0.25$

Source: iRI Academic Data Set 2001-2010

TABLE 7. OLS REGRESSION: PRICE PREMIUMS ($) OF PRODUCT CATEGORIES

	Peanut Butter	Spaghetti Sauce	Hot Dogs
Natural	1.324	0.011	0.346
Organic	2.911	1.346	2.451

Note: All values are statistically significant at the 1% level.
Source: iRI Academic Data Set 2001-2010