Misleading Labeling and Product Quality *

Frances Xu Lee†

February 25, 2022

Abstract

Using a positive term in labeling a product may mislead buyers into expecting the product to have a higher quality than the true quality. However, given the limitation that the labeling language has to be simple, completely eliminating inflation of the buyers' beliefs by a positive term will also eliminate information transmission of the product quality. Information transmission is needed to motivate the seller to invest in the quality in the first place. A "reasonable consumer" standard that holds any labeling that has misled a reasonable consumer liable results in too little use of the positive term and too little effort in providing the quality valuable to the buyers. More market-beneficial labeling behaviors can be encouraged by a legal or regulatory policy that also dismisses cases where buyers are only moderately misled so as to preserve the seller's incentive to label with the positive term even when the quality is not perfect.

Keywords: misleading advertising, reasonable consumer, Bayesian consumer, adverse selection

JEL Classification. K13; D82

*I thank the helpful comments from James Cooper, Nate Atkinson and William Hubbard, the audience at Law and Economics Theory Conference at Northwestern Pritzker School of Law, 19th Annual International Industrial Organization Conference, Antonin Scalia Law School at George Mason University, and Georgetown University Law School.

†Quinlan School of Business, Loyola University Chicago. E-mail: francesxu312@gmail.com.

1. Introduction

Numerous lawsuits have been filed against sellers for labeling their products with a positive term, such as "natural", "locally grown", "low sugar", "heart healthy", "Made in USA" and so on. At least 100 such lawsuits were filed between 2011 to 2013 alone.[1] Silverman (2018) recorded over 425 active food marketing class action lawsuits in the federal court between 2015 to 2016. Many of these cases settle. For example, Pepsico Inc. settled for $9 million for its potato chips. Kashi settled for $3.99 million for its cereals.[2] Some lawsuits were dismissed. Often, the court dismisses such class action lawsuits based on whether a "reasonable consumer" would have been misled. For example, in 2019, a federal court judge tossed out a class action lawsuit alleging that Rachael Ray Pet Food's "natural" labeling is false and misleading advertising. The court argued that no "reasonable consumer" would expect a product with a label stating it is "natural" to be completely free from any amount of glyphosate, a type of herbicide[3]. Similarly, the court dismissed cases targeting terms like "handmade" or "handcrafted" against alcohol beverage makers because a "reasonable consumer" would not expect bourbon to be entirely made without machinery.[4] The key question in these cases are how a "reasonable consumer" would interpret a positive term.[5]

Any policy that uses a party's belief as the benchmark has to consider the endogeneity of such a belief. Consumers' interpretation of a positive term would depend on the sellers' behavior. If all sellers in the market overstate their quality, then the buyers would rationally discount the seller's statements, just as how rampant grade inflation would cause employers to discount the GPA on the transcripts. In this article, I explore how the "reasonable consumer" legal standard would affect the market outcome and whether alternative legal standard can benefit the market outcome.

Outside the court, the various government agencies have relied on the concept of a reasonable consumer as well. The U.S. Food and Drug Administration (FDA) interpreted the term "natural" to mean that "nothing artificial or synthetic (including all color additives regardless of source) has been included in, or has been added to, a food that would not *normally be*

[1]See Esterl (2013).

[2]*Id*.

[3]Markeith Parks v. Ainsworth Pet Nutrition LLC d/b/a Rachael Ray Nutrish, Case No. 1:18-cv-06936-LLS, in the U.S. District Court for the Southern District of New York.

[4]See Salters v. Beam Suntory, 2015 WL 2124939.

[5]These are cases where the puffery defense typically does not apply because the positive terms here are factual speech rather than nonfactual, such as the term "the best". See Hoffman (2005) for how the puffery defense was successful in defending terms such as "the world's best aspirin" and "very best chocolate", etc.

expected to be in that food (emphasis added)," while declining to establish a formal definition of the term "natural"[6]. The Federal Trade Commission (FTC), to which the courts offer deference, has over the time changed the standards on deception. Prior to 1984, FTC insisted on the most literal truthfulness of the advertisements.[7] Then in 1984 it redefines deception, more in favor of the defendants, as "a representation, omission or practice that is likely to mislead the consumer acting *reasonably* in the circumstances, to the consumer's detriment (emphasis added)."[8]

An obvious challenge in defining a term such as "natural" without relying on a "reasonable consumer" is that food contains a wide range of ingredients and each ingredient can be in varying amounts. Even if people agree on a particular ingredient, such as a pesticide, as being not natural, a question remains as to how low the levels of these unnatural ingredients have to be in order to render the final product worthy of the term "natural". Similarly, the production process of a product takes many steps that involve different levels of automation and different ingredients from different locations, so it is hard to decide whether a production process deserves a term like "handmade", "made in USA" or "locally grown".

Very often, the quality desired by the buyers, such as more "natural" or more "handcrafted", is costly to the seller. For example, Pepsico Inc's Frito-Lay snack unit reformulated more than 60 products in order to make its potato chips more natural. It removed about three dozen artificial ingredients including FD&C Red 40, a food coloring, and replaced it with beets, cabbage and carrots. The company experimented with more than 300 versions of its barbecue potato chips alone before finding one that tasted like the original.[9]

In this paper, I provide a theoretical framework to analyze the adoption of a positive labeling term when buyers are Bayesian and rational. The seller effectively sends a binary message to the buyers by either adopting the term or not in the labeling. I am specifically interested in the amount of information transmitted to the buyers and the seller's ex ante incentive to pay effort to improve this dimension of quality. A key assumption is that the labeling of the products on the product packaging has to be simple relative to the nuanced product quality. This is due to the limitation of space, the limited time and attention span of the buyers, or the need to protect trade secrets. Given that the language is coarse relative to the nuanced and varied

[6]"Use of the Term Natural on Food Labeling" available on FDA website.

[7]The court decision in *Charles of the Ritz Distributors Corp. v. FTC* allows the FTC to "insist upon the most literal truthfulness in advertisements . . . and should have the discretion, undisturbed by the courts, to insist, if it chooses, upon a form of advertising clear enough so that, in the words of the prophet Isaiah, 'wayfaring men, though fools, shall not err therein."'

[8]See "FTC Policy Statement on Deception" available at the FTC's website.

[9]*Supra* Footnote 1.

underlying product quality, this paper shows that inflating the buyers' belief from the true product quality is not always a sign of socially undesirable behavior. It is actually necessary in order for information to be maximally communicated through labeling. To see that, one can simply look at the extreme case where only products with 0% of pesticide are labeled "natural", in which case most products, with varying levels of pesticides, are not differentiated at all in the labeling.

When a seller's effort leads to non-deterministic quality, to encourage the maximal amount of information transmission and thus to motivate the seller's effort, any regulation or law should target labeling behavior that involves a certain chance of inflating the buyers' belief, or in other words, should only deter egregious inflation of the buyers' beliefs.

Following the tradition of economics, the belief of a "consumer acting reasonably" or a "reasonable consumer" should be the equilibrium expectation of quality (in the specific dimension the labeling language concerns) of a Bayesian consumer. Then the "reasonable consumer" standard is a binary decision rule for the court: a seller is liable if its positive term on the labeling leads a Bayesian consumer to form an equilibrium expectation of quality (the "reasonable belief") that is higher than the truth, and not liable otherwise. This paper shows that such a decision rule, however, is likely to lead to an undesirable equilibrium, where all sellers do not label their products with the positive term. Such failure is based on the old wisdom of adverse selection from Akerlof (1978). If the court only permits a positive term when the quality is above the "reasonable belief", then the positive term signals an average quality that is above this "reasonable belief" cutoff, which will cause a rational buyer to adjust up the expectation of the quality to a new higher "reasonable belief". This raised bar will push out a segment of lower-quality sellers from adopting the positive term and this process can only stop when only the absolute highest-quality seller uses the positive term without liability, while attaining that quality is almost technologically impossible. That is, this endogenously moving benchmark causes the collapsing of the information transmission.

We are already seeing and experiencing that trend in the market place, which has led the Wall Street Journal to comment in 2013 that "(f)ood products labeled as 'natural' are starting to disappear." It cites Datamonitor in reporting that "Only 22.1% of food products and 34% of beverage products launched in the U.S. during the first half of 2013 claimed to be 'natural,' down from 30.4% and 45.5%, respectively, in 2009." This trend is likely to result in a decline of efforts to make products more natural. In this paper, I recommend a different approach where the court finds cases that only moderately mislead a reasonable consumer not liable. This approach can be given a new name, such as "egregiously misleading" standard.

Literature

There is a literature in law that looks closely at how courts have made decisions in these labeling class action lawsuits, but not at how courts should make decisions. Negowetti (2013) and Silverman (2018) gave a lot of case by case details on the various legal decisions.

In a different context, the legal literature has recognized the dependence of people's expectation of privacy on the search intensity of the government and the court's decisions. Such dependence creates a degree of circularity in the Katz v United States (389 US 347 (1967))'s "reasonable expectations of privacy" test and can cause the privacy protection of 4th amendment to unravel. Richard Posner (1979) observed that "it is circular to say that there is no invasion of privacy unless the individual whose privacy is invaded had a reasonable expectation of privacy; whether he will or will not have such an expectation will depend on what the legal rule is." See Kugler and Strahilevitz (2017) for a discussion of this literature and the empirical work testing whether and how the expectation of privacy reacts to court decisions.

The "reasonable" person test is also the traditional test for compliance with the duty of care in torts. Negligence is defined as doing what a reasonable person would not do or failing to do what a reasonable person would do. In the context of negligence, there are two major definitions of the reasonableness. It is either a normative definition, what a reasonable person should do, or a positive definition, what an average person empirically would do. See Miller and Perry (2012) for a detailed examination of these concepts in the context of negligence. Here, I look at reasonableness in the context of belief formation. In some sense, I co-mingle both the normative and positive aspects in my definition of the reasonable consumer belief: the rational Bayesian updating can be thought of as what a consumer "should" do, but I also allow the consumers to be gullible, who do not interpret labels in a strictly scientific way, in my extension.

This paper advocates a standard that is less stringent than the "reasonable consumer" standard defined as using the reasonable expectation as the benchmark to determine whether *there is liability*. A recent and very insightful paper, Cooper and Kobayashi (2021), examines the various benchmarks used to determine *the amount of damage* when liability is already established. They also show how the current practice of the court in determining damage is too stringent on the seller from an economic perspective because the court assumes that all sales in the market are coerced by the false claim while only part of the sales are. They also observe that "efficient conduct will include some actions that are potentially deceptive." This stems from the heterogeneity among the consumers. However, our conclusion that the efficient conduct will include some actions that are surely deceptive is based on homogeneous and Bayesian consumers.

The information transmission in this paper is a form of costly lying in the economics literature, pioneered in Kartik (2009). In the literature of costly lying, there is no hard evidence that the seller can reveal so there is no full voluntary disclosure of information on the product quality. On the other hand, there is some cost of inflating the language from the truth, which is in the form of legal liability in this paper. In Kartik (2009), the language of the reports can be as rich as the set of information sender's types, which allows for a range of full separation: where the private information is precisely communicated through the language. However, the main feature here is that the labeling language is not rich and can provide at best only a partition of the set of possible qualities. Labels here have to be concise because of either the attention capacity of the buyers, or the space constraint on the packaging, or the need for a seller to keep its business secrets private.

Rhode and Wilson (2018) also reached the conclusion that a certain amount of equilibrium false advertising is good for the social welfare. However, the reasoning is different from this paper. In Rhode and Wilson (2018), the language of the advertising is rich (as rich as the number of different quality levels), but the seller may not sell to all buyers. False advertising induces the buyers to over-estimate the quality of a low-quality product and the seller will expand the quantity sold as a result, which counters the inefficiently low output level of a monopoly. In this paper, this source of benefit of false advertising is purposefully avoided by letting the seller always cover the full market. Certain amount of equilibrium false advertising is also socially optimal in Corts (2014) because it is costly for sellers to find out their own true quality. The high-quality seller can then signal the high quality using "speculative" claims that will prove to be false with some probability. In this paper, the seller is fully informed of its own quality before labeling the product. Piccolo et. al. (2015) shows that false advertising can also benefit the buyers by making the products more homogeneous between competing sellers. More broadly, this paper is related to works that study advertising as communication to Bayesian buyers, whose meaning is endogenously determined. See, for example, Chakraborty and Harbaugh (2014) and Anderson and Renault (2006).

The rest of the paper is organized as follows. Section 2 builds a model to investigate the market outcome given a court's decision rule. It shows that, in particular, if the court's decision rule satisfies the reasonable consumer standard, the equilibrium is undesirable from an information transmission and effort motivation perspective. Section 3 investigates what the market-optimal labeling behavior is and how a court policy can implement that. I illustrate it with a numeric example. Section 4 allows some of the buyers to be gullible and redefines the reasonable belief as an average of the sophisticated and the gullible consumers' beliefs. Section 5 links the analysis to the certification of labeling terms. Section 6 discusses limitations and extension of the model. Section 7 concludes.

2. The model

A monopoly seller can choose whether to exert effort in improving a certain quality dimension. Normalize the cost of no effort to be zero and the cost of effort to be $C > 0$. This cost is the private information of the seller, drawn from a continuous distribution G over $[0, 1]$. Given no effort, the quality θ is drawn from a distribution F_L over $[0, 1]$ and given effort, the quality θ is drawn from a distribution F_H over $[0, 1]$. Assume F_H first-order stochastic dominates F_L. That is, $F_H(\theta) < F_L(\theta)$ for all $\theta \in (0, 1)$. Assume both distributions have positive density everywhere on the support. Let $E_H[\theta]$ denote the expected quality under distribution F_H and $E_L[\theta]$ denote the expected quality under distribution F_L.

As an example, one can think of this quality θ being the percentage of the ingredients in a cereal bar that is free of synthetic additives. Effort by the seller in monitoring each supplier and trying out different ingredients in laboratories reduces the chance of synthetic additives, but this effort is costly to the seller.

Higher θ is more desirable to the buyers. Assume there is a unit measure of identical buyers. They all have unit demand each and their value of consuming a product with quality θ is directly $v + \theta$ with $v > 0$. The effort cost C, the realized quality θ, and whether the seller has exerted effort are all private information of the seller and are unobservable to the buyers.

The seller can choose to label the product, at no cost, with either a positive term or not. That is, the seller can send a binary message m to the buyer, where $m \in \{\Gamma, \phi\}$, with Γ denote the positive term and ϕ denote the lack of such a term. To continue the example of cereal bar, the Γ message to the buyers is printing "natural" on the label and the ϕ message is not printing it.[10] A key assumption is that the quality level is continuous, but the message is binary.

The seller also sets a price, p, for the product. The buyers form a belief about the quality given the message m and the price p. The buyers make purchase their decisions based on their shared belief of the quality.

If the seller does not use the positive term Γ, the seller will not be liable in court. If a seller has used the term Γ and is sued, the court learns about θ. The court sets a quality cutoff $\hat{\theta}$ and

[10]This paper assumes that the labeling language is exogenously coarse. There is a literature that endogenizes the use of coarse grade when finer grades are available. See Harbaugh and Rasmusen (2018) Section V for a review of this literature.

holds any $\theta < \hat{\theta}$ liable and holds any $\theta \geq \hat{\theta}$ not liable. Assume that there is also a positive but arbitrarily small probability ε that the court does not follow its policy and instead finds any seller with a positive term liable and there is an equal probability ε that the court does the opposite: it acquits any seller with a positive term. This ε captures the uncertainty in the litigation.[11] The theoretical role of this uncertainty is to ensure that the seller still benefits from a higher price even when knowing that the court intends to hold the seller liable for using the positive term.

If the seller is deemed liable, the court applies restitution. That is, the seller is liable for the "price premium", defined to be the actual price paid by the buyers minus the "but for" price, i.e., the price the buyers would have paid if the positive term was not used. When liable, the seller incurs a reputation and litigation costs $K > 0$.

We assume the court's policy, $\hat{\theta}$, is pre-committed and stable over time, so the public learns about it by closely observing the court's past decisions and the reasoning written in the opinions on summary judgement and motion to dismiss.[12]

For the purpose of the analysis, we assume a non-strategic plaintiff attorney that sues any seller that uses a positive term. This can also be endogenized by assuming that the experienced plaintiff attorney's cost of filing is zero and they are tempted by the uncertainty in the litigation to sue. Silverman (2018) observes that "(t)he small group of attorneys who bring these cases will continue shopping for lawsuits and generating cut-and-paste complaints so long as there is more than a nominal chance of a settlement and no risk to asserting even the most far-fetched claims."

In equilibrium, the buyers form their belief using Bayes' Rule taking into account the seller's equilibrium effort and labeling strategies. Denote the buyers' updated belief distribution conditional on message m by F_m, and denote the expectation of quality given F_m by θ_m ($m \in \{\Gamma, \phi\}$). The timing of the game is the following:

1. Court's past ruling reveals $\hat{\theta}$.
2. Nature determines the seller's effort cost C, privately revealed to the seller.
3. The seller chooses whether to exert effort or not.
4. The quality θ is realized and privately revealed to the seller.

[11] In the discussion (Section 6.1), I consider more smooth uncertainty in the court's ruling. This uncertainty could also be built into the assumptions about the plaintiffs' behaviour rather than the court's.

[12] Language such as "this case is dismissed because a reasonable consumer is only moderately misled" can help communicate the court's policy to the public. Other possible and language that does not rely on "reasonable consumer" belief can be, for example, "this case is dismissed because the percentage of herbicide found in the product is lower than the average of the industry".

5. The seller chooses a message $m \in \{\Gamma, \phi\}$.
6. A unit measure of buyers form expectation θ_Γ if the message is Γ and forms expectation θ_ϕ if the message is ϕ.
7. The seller posts a price p.
8. Buyers choose whether or not to buy after observing the price.
9. If message Γ was sent, the plaintiff attorney sues.
10. If the seller was sued, the court observes θ and punishes according to the policy, with a chance of error.

The payoff of the seller is the profit minus the punishment and the reputation/litigation cost if any. The payoff of the buyer is the value of the product minus the price paid. The buyers get outside option normalized to 0 for not buying. We assume each buyer gets at most a negligible amount of compensation from the lawsuit even if the seller is found liable, so the prospect of a lawsuit does not enter into the buyer's purchase decision.[13] We assume that the buyers buy when indifferent, the seller does not exert effort when indifferent, and the seller uses the positive term when indifferent. All players are risk neutral. Given any policy of the court, the equilibrium concept is Perfect Bayesian Equilibrium. Since both the effort cost and product quality are private information of the seller, we use "cost-types" and "quality-types" to refer to these two dimensions of private information.

Recall that the expectation of quality following the positive term is θ_Γ. Define a "reasonable consumer" policy as one that holds any inflation of the buyers' expectation above the truth to be liable when the positive term is used.

Definition 1. A policy $\hat{\theta}$ satisfies the "reasonable consumer" standard if the buyers' equilibrium expectation of the quality given this policy satisfies $\theta_\Gamma = \hat{\theta}$.

We will investigate what $\hat{\theta}$ gives a "reasonable consumer" policy, the impact of such a policy and whether an alternative policy is better.

The seller's pricing decision

We will do backward induction. First, we consider the subgame after the seller has sent the message/labeled the product and the buyers have form the expectations θ_Γ and θ_ϕ.

If the seller does not use the positive term, then there is no litigation risk, and the buyers will purchase as long as the price is less or equal to $v + \theta_\phi$. Therefore, a seller with the positive term charges price $v + \theta_\phi$ and gets payoff $v + \theta_\phi$.

[13]The compensation to each buyer is typically tiny because the class size is big and the attorneys take a big cut form the settlement amount. For example, when Red Bull settled for $13 million in 2013, many class members received only $4.23 cash from the settlement. See Silverman (2018).

If all buyers purchase and the seller is not found liable in court, the subgame payoff of the seller is p. If all buyers purchase, but the seller has used Γ and is found liable in court, the subgame payoff is $p - [p - (v + \theta_\phi)] - K = v + \theta_\phi - K$. That is, the seller has to pay back the price premium calculated based on the but-for price, $v + \theta_\phi$. Notice that the payoff of being liable is not related to the actual price charged to the buyers, while the payoff of not being liable is strictly increasing in the price. Moreover, after the effort cost is sunk, a higher quality product does not incur a higher marginal cost, therefore, price signaling is not possible here. The proof for the following Lemma is in the appendix.

Lemma 1. *The unique subgame equilibrium after labeling is a pooling equilibrium, where all quality-types who have used the positive term all charge the same price* $p = v + \theta_\Gamma$*; and all types who have not used the positive term all charge the same price* $p = v + \theta_\phi$*.*

The seller's labeling decision

We go backward to consider the subgame after the quality has realized. Let $\hat{F}$ denote the cumulative distribution of quality given the equilibrium effort strategy of all cost-types of the seller. The following lemma states that the court policy $\hat{\theta}$ results in a labeling strategy of the seller that also has $\hat{\theta}$ as the cutoff. High-quality types use the positive term and low-quality ones do not.

Lemma 2. *For any court policy* $\hat{\theta}$ *and a distribution* $\hat{F}$ *of quality-types, there is a unique subgame equilibrium of a cutoff feature, with cutoff* $\hat{\theta}$*, such that any* $\theta < \hat{\theta}$ *chooses not to use the positive term and charges price* $v + E_{\hat{F}}[\theta|\theta < \hat{\theta}]$ *and any* $\theta \geq \hat{\theta}$ *chooses to use the positive term and charges price* $v + E_{\hat{F}}[\theta|\theta > \hat{\theta}]$*, where the expectations are taken according to* $\hat{F}$*.*

Proof. Let $L(\theta)$ denote the probability of type θ being found liable after using the positive term Γ. A seller with quality θ's payoff of using the positive term is $(1 - L(\theta))(v + \theta_\Gamma) + L(\theta)(v + \theta_\phi - K)$ and the payoff of not using the positive term is $v + \theta_\phi$. Therefore, the difference is $(1 - L(\theta))(\theta_\Gamma - \theta_\phi) - L(\theta)K$. Then, for $\theta < \hat{\theta}$, the payoff difference is $\varepsilon(\theta_\Gamma - \theta_\phi) - (1 - \varepsilon)K$, so for ε sufficiently small the difference is negative. For $\theta \geq \hat{\theta}$, the payoff difference is $(1 - \varepsilon)(\theta_\Gamma - \theta_\phi) - \varepsilon K$, which is positive for sufficiently small ε.

In equilibrium, the belief of the buyers given the message has to be correct, so $\theta_\phi = E_{\hat{F}}[\theta|\theta < \hat{\theta}]$ and $\theta_\Gamma = E_{\hat{F}}[\theta|\theta > \hat{\theta}]$. □

That is, with a vanishingly small error, the court's policy $\hat{\theta}$ induces a labeling behavior characterized by $\hat{\theta}$. Note that, if the court's policy has $\hat{\theta} = 1$ (the highest standard possible), then only the seller with the perfect quality will use the positive term, while if the court has $\hat{\theta} = 0$ (the lowest standard possible), then all quality-types will use the positive term.

The seller's effort decision

Going further backward, the seller's effort strategy has to follow a cutoff pattern too: a higher-effort-cost seller chooses not to exert the effort, and a lower-effort-cost seller chooses to exert the effort. This is because the benefit of the effort is the same for all cost-types, but the costs of effort vary. Let $\hat{C}$ denote the cost cutoff. Expecting this cutoff $\hat{C}$, prior to observing the label from the seller, the buyers would expect the quality to be distributed according to

$$\hat{F}(\hat{C};\theta) = G(\hat{C})F_H(\theta) + (1 - G(\hat{C}))F_L(\theta)$$

If $\hat{C} = 0$, which means no cost-type exerts effort, then $\hat{F}$ coincides with F_L. On the other hand, if $\hat{C} = 1$, which means all cost-types exert effort, then $\hat{F}$ coincides with F_H. For conciseness of the notations, we sometimes suppress the arguments for the function $\hat{F}$. The seller's effort strategy, summarized by the cutoff $\hat{C}$, influences the buyers' expectations of the quality θ_Γ and θ_ϕ through the distribution $\hat{F}$, which in turn determines how much the seller can charge the buyers. The seller's revenue of exerting the effort is:

$$\overline{R} = v + [1 - F_H(\hat{\theta})]((1-\varepsilon)E_{\hat{F}}[\theta|\theta > \hat{\theta}] + \varepsilon(E_{\hat{F}}[\theta|\theta < \hat{\theta}] - K)) + F_H(\hat{\theta})E_{\hat{F}}[\theta|\theta < \hat{\theta}]$$

This reflects that, with effort, the quality will be distributed according to F_H. When using the positive term (with probability $1 - F_H(\hat{\theta})$), there is a small chance ε of the seller being held liable despite the high quality. Similarly, the seller's revenue of not exerting the effort is:

$$\underline{R} = v + [1 - F_L(\hat{\theta})]((1-\varepsilon)E_{\hat{F}}[\theta|\theta > \hat{\theta}] + \varepsilon(E_{\hat{F}}[\theta|\theta < \hat{\theta}] - K)) + F_L(\hat{\theta})E_{\hat{F}}[\theta|\theta < \hat{\theta}]$$

This reflects that, with effort, the quality will be distributed according to F_L. Therefore, the benefit of the effort is the difference $\overline{R} - \underline{R}$:

$$\overline{R} - \underline{R} = [F_L(\hat{\theta}) - F_H(\hat{\theta})]((1-\varepsilon)(E_{\hat{F}}[\theta|\theta > \hat{\theta}] - E_{\hat{F}}[\theta|\theta < \hat{\theta}]) - \varepsilon K)$$

Because ε is arbitrarily small, the benefit $\overline{R} - \underline{R} \in (0, 1)$. The effort cost C is distributed over $[0, 1]$, so there is at least one cost-type who is indifferent between exerting effort or not. That is, the cutoff cost-type $\hat{C}$ exists and must satisfy the indifference condition that the cost is equal to the benefit:

$$\hat{C} = [F_L(\hat{\theta}) - F_H(\hat{\theta})]((1-\varepsilon)(E_{\hat{F}(\hat{C})}[\theta|\theta > \hat{\theta}] - E_{\hat{F}(\hat{C})}[\theta|\theta < \hat{\theta}]) - \varepsilon K)$$

To the extent that this equation admits multiple solutions for $\hat{C}$ because $\hat{F}$ depends on $\hat{C}$, pick the largest $\hat{C}$ which allows for the highest chance of effort because the chance of effort is $G(\hat{C})$. Therefore, we can view this largest $\hat{C}$ as a function of $\hat{\theta}$. That is, the court's cutoff for dismissal $\hat{\theta}$ determines the equilibrium effort strategy of the seller, which means the distribution of quality $\hat{F}$ depends on $\hat{\theta}$. The following proposition then follows from the analysis so far, so a proof is omitted.

Proposition 1. *Given a court's policy $\hat{\theta}$, the equilibrium has the following feature: a seller with cost below $\hat{C}$ exerts effort and a seller with cost above $\hat{C}$ does not exert effort. After quality is realized, a seller with quality below $\hat{\theta}$ does not use the positive term and a seller with quality above $\hat{\theta}$ uses the positive term, where the cost cutoff $\hat{C}$ is a solution to the following equation:*

$$\hat{C} = [F_L(\hat{\theta}) - F_H(\hat{\theta})]((1-\varepsilon)(E_{\hat{F}(\hat{C})}[\theta|\theta > \hat{\theta}] - E_{\hat{F}(\hat{C})}[\theta|\theta < \hat{\theta}]) - \varepsilon K) \tag{1}$$

where $\hat{F}(\hat{C};\theta) = G(\hat{C})F_H(\theta) + (1 - G(\hat{C}))F_L(\theta)$.

Recall that the policy fits the "reasonable consumer" standard if and only if $\hat{\theta} = \theta_\Gamma$. This can only happen when the policy is extremely stringent, as shown by the following corollary.

Corollary 1. *The only policy that satisfies the "reasonable consumer" standard is one with $\hat{\theta} = 1$. Under this policy, only a seller with $\theta = 1$ uses the positive term and the seller exerts no effort in equilibrium to improve the quality.*

Proof. Because $\hat{\theta} = \theta_\Gamma = E_{\hat{F}}[\theta|\theta > \hat{\theta}]$, it must be $\hat{\theta} = 1$, which in turn implies $F_H(\hat{\theta}) = F_L(\hat{\theta}) = 1$, so $\overline{R} - \underline{R} = 0$. Therefore, no cost-type exerts effort. □

To understand how the reasonable consumer standards leads to an extreme equilibrium standard, one can also borrow the arguments behind how "adverse selection" (Akerlof 1978) leads to a market failure: if only cases with quality above the reasonable belief is dismissed, cases with quality below the reasonable belief will stop using the positive term in the label, which in turn pushes up the reasonable belief associated with the positive term in the label. This process keeps pushing lower qualities away from using the positive term, until this positive term "selects" only the top quality. But this is an undesirable outcome because then the positive term does not help the consumers to discern most of the qualities in the market.

Under $\hat{\theta} = 1$, sellers with virtually all quality levels will not use the positive term, are not distinguishable in the mind of the buyers, and will charge the same price $v + E_{\hat{F}}[\theta]$. There is no incentive to influence the quality through effort.

The next section explores what policy is better.

3. The market-optimal policy

Here, the buyers all purchase. The purchase price is a transfer between the buyer and the seller within the society. The judgement or settlement in a litigation is also a transfer within the society, so the social welfare (the sum of all parties' payoffs) here is simply the gains-of-trade in the market place minus the litigation cost if any. The litigation cost is incurred only due to the error of the court. When ε is infinitely small, the social welfare converges to the gains-of-trade, so in this paper we will use the gains-of-trade as the welfare objective, and focus on the "market-optimal" outcome, defined to be the equilibrium outcome that maximizes the gains-of-trade in the market place subject to the constraint that it is an equilibrium outcome achievable under some court's policy. Formally, the gains-of-trade as a function of the cost cutoff is:

$$W(\hat{C}) \equiv \int_0^{\hat{C}} \left[v + \int_0^1 \theta \, \mathrm{d}F_H(\theta) - C \right] \mathrm{d}G(C) + \int_{\hat{C}}^1 \left[v + \int_0^1 \theta \, \mathrm{d}F_L(\theta) \right] \mathrm{d}G(C)$$

This reflects that when $C < \hat{C}$, effort is exerted and effort cost is incurred and the quality is distributed according to F_H, whereas for $C > \hat{C}$, effort cost is not incurred and the quality is distributed according to F_L.

Because the policy cutoff is directly the labeling cutoff, this is simply a problem of finding the market-optimal labeling cutoff. The market-optimal θ^* is the solution to the following problem:

$$\max_{\hat{\theta}} \quad W(\hat{C}(\hat{\theta}))$$
$$s.t. \quad \hat{C}(\hat{\theta}) \text{ is the largest solution to equation (1).}$$

The seller's effort creates an expected consumption benefit of:

$$\int_0^1 \theta \, \mathrm{d}F_H(\theta) - \int_0^1 \theta \, \mathrm{d}F_L(\theta) = E_H[\theta] - E_L[\theta] > 0$$

The effort also creates a cost C to the seller. An effort strategy that makes the seller pay effort whenever the consumption benefit exceeds the cost of effort creates the first-best level of gains-of-trade. That is, a cutoff $C^* = E_H[\theta] - E_L[\theta]$ solves an unconstrained problem of maximizing the gains-of-trade by choosing $\hat{C}$ directly:

$$C^* = E_H[\theta] - E_L[\theta] = \arg\max_{\hat{C}} W(\hat{C})$$

Following standard terminology, we call this C^* the first-best cost cutoff. It results in an upper bound of the gains-of-trade possible in this market place.

Proposition 2. *Under any court's policy, the equilibrium cost cutoff is weakly lower than the first-best cost cutoff. That is, in general too little effort is exerted in equilibrium compared to the first best.*

Proof. By first order stochastic dominance, for any $\hat{\theta} \in [0, 1]$,

$$E_H[\theta|\theta > \hat{\theta}] \geq E_{\hat{F}}[\theta|\theta > \hat{\theta}] \geq E_L[\theta|\theta > \hat{\theta}],$$

$$E_H[\theta|\theta < \hat{\theta}] \geq E_{\hat{F}}[\theta|\theta < \hat{\theta}] \geq E_L[\theta|\theta < \hat{\theta}].$$

Therefore,

$$\begin{aligned}
&\overline{R} + (1 - F_H(\hat{\theta}))\varepsilon(E_{\hat{F}}[\theta|\theta > \hat{\theta}] - E_{\hat{F}}[\theta|\theta < \hat{\theta}] + K) - v \\
&= (1 - F_H(\hat{\theta}))E_{\hat{F}}[\theta|\theta > \hat{\theta}] + F_H(\hat{\theta})E_{\hat{F}}[\theta|\theta < \hat{\theta}] \\
&\leq (1 - F_H(\hat{\theta}))E_H[\theta|\theta > \hat{\theta}] + F_H(\hat{\theta})E_H[\theta|\theta < \hat{\theta}] = E_H[\theta]
\end{aligned}$$

$$\begin{aligned}
&\underline{R} + (1 - F_L(\hat{\theta}))\varepsilon(E_{\hat{F}}[\theta|\theta > \hat{\theta}] - E_{\hat{F}}[\theta|\theta < \hat{\theta}] + K) - v \\
&= (1 - F_L(\hat{\theta}))E_{\hat{F}}[\theta|\theta > \hat{\theta}] + F_L(\hat{\theta})E_{\hat{F}}[\theta|\theta < \hat{\theta}] \\
&\geq (1 - F_L(\hat{\theta}))E_L[\theta|\theta > \hat{\theta}] + F_L(\hat{\theta})E_L[\theta|\theta < \hat{\theta}] = E_L[\theta]
\end{aligned}$$

Therefore, $\hat{C} = \overline{R} - \underline{R} \leq E_H[\theta] - E_L[\theta] - (F_L(\hat{\theta}) - F_H(\hat{\theta}))\varepsilon(E_{\hat{F}}[\theta|\theta > \hat{\theta}] - E_{\hat{F}}[\theta|\theta < \hat{\theta}] + K) < E_H[\theta] - E_L[\theta] = C^*$. □

The intuition is that the coarse labeling provides some information to the buyers, but less than the full information, so the seller does not get the full benefit of a higher quality. Moreover, any error of the court causes the seller to incur some reputation or litigation cost K for communicating through the coarse labeling. Next, we will investigate what can be achieved under a court policy. A market-optimal labeling cutoff is one that gives rise to the highest possible incentive for effort (the highest $\hat{C}$). It is clear that such a labeling cutoff has to be interior:

Corollary 2. *The problem* $\max_{\hat{\theta}} W(\hat{C}(\hat{\theta}))$ *is equivalent to* $\max_{\hat{\theta}} \hat{C}(\hat{\theta})$*. Moreover, the solution, denoted by* θ^**, must satisfy* $\theta^* \in (0,1)$*.*

Proof. The second order derivative of the gains-of-trade $W(\hat{C})$ with respect to $\hat{C}$ is $-1 < 0$. Then Proposition 2 implies that the gains-of-trade is strictly increasing in $\hat{C}$ for $\hat{C} < C^*$. When $\hat{\theta} = 1$ or $\hat{\theta} = 0$, because $F_H(\hat{\theta}) = F_L(\hat{\theta})$, $\hat{C} = \overline{R} - \underline{R} < 0$. When $\hat{\theta} \in (0,1)$, $F_H(\hat{\theta}) < F_L(\hat{\theta})$, for sufficiently small ε, $\hat{C} = \overline{R} - \underline{R} > 0$. □

This in particular implies that the "reasonable consumer" standard is undesirable because it results in $\hat{\theta} = 1$ and no effect because $\hat{C}(1) = 0$. In practice, the reasonable consumer standard is so attractive because the reasonable consumer's expectation of quality is already reflected in the observable purchase price. The problem is not using the endogenous consumers' expectation as a benchmark, the problem is that the "reasonable consumer" standard tries to punish any positive gap between the consumers' expectation and the truth. The following corollary shows two ways of achieving the market-optimal labeling cutoff of θ^*, with the latter still using the consumer expectation as a benchmark:

Corollary 3. *The following policies maximizes the gains-of-trade.*

1. *Quality-based: a court policy that finds the seller liable if and only if* $\theta < \theta^*$*.*
2. *Belief-based: a court policy that finds the seller liable if and only if the amount of belief inflation is greater than* $E_{F^*}[\theta|\theta > \theta^*] - \theta^*$*, where* $F^* = \hat{F}(\hat{C}(\theta^*), \theta^*)$*.*

The former policy uses a cutoff quality as a benchmark, while the latter uses the consumer expectation as a benchmark while allowing an appropriate amount of deviation.[14] Next, a numeric example illustrates the market-optimal policies.

Numeric example:

Let G be uniform over $[0,1]$. Let $v = 2$. Let $K = 0.1$. Let $\varepsilon \to 0$. Let F_H be uniform over $[1/2, 1]$ and F_L be uniform over $[0, 1/2]$.[15] Recall that the policy element $\hat{\theta}$ directly determines

[14]Following a belief-based policy that allows a specific gap, it is not guaranteed that the equilibrium is unique, but there is at least one equilibrium where the labeling cutoff is θ^*. The lack of uniqueness is because if the labeling cutoff is lower than θ^*, the effort cutoff drops, so the distribution of quality is worse which reduces the expectation of quality given a positive term, which in turns allows a seller at a lower labeling cutoff to induce the exact same gap between the truth and the expectation.

[15]The distributions are not everywhere positive on $[0,1]$, but they provide an easy example and a full support is not crucial.

the labeling cutoff $\hat{\theta}$. First we investigate how the equilibrium cost cutoff is related to the labeling cutoff $\hat{\theta}$.

For any cost cutoff $\hat{C}$, the distribution $\hat{F}$ has two segments: one segment is uniform on $[0, 0.5]$ with density $1 - \hat{C}$, and the other segment is uniform on $[0.5, 1]$ with density $\hat{C}$.

(1) For $\hat{\theta} \in (0, 0.5]$, we have $F_L(\hat{\theta}) = 2\hat{\theta}$ and $F_H(\hat{\theta}) = 0$, so $F_L(\hat{\theta}) - F_H(\hat{\theta}) = 2\hat{\theta}$.

$$E_{\hat{F}}[\theta|\theta > \hat{\theta}] = \frac{\frac{1}{4} + \frac{\hat{C}}{2} - (1-\hat{C})\hat{\theta}^2}{1 - 2\hat{\theta} + 2\hat{\theta}\hat{C}}, \qquad E_{\hat{F}}[\theta|\theta < \hat{\theta}] = \frac{\hat{\theta}}{2}.$$

$$\hat{C} = \hat{\theta}\frac{\frac{1}{2} + \hat{C} - \hat{\theta}}{1 - 2\hat{\theta} + 2\hat{\theta}\hat{C}} \Rightarrow \hat{C}(\hat{\theta}) = \frac{3\hat{\theta} - 1 + \sqrt{1 - 6\hat{\theta} + 13\hat{\theta}^2 - 8\hat{\theta}^3}}{4\hat{\theta}}.$$

(2) For $\hat{\theta} \in [0.5, 1)$, we have $F_L(\hat{\theta}) = 1$ and $F_H(\hat{\theta}) = 2\hat{\theta} - 1$, so $F_L(\hat{\theta}) - F_H(\hat{\theta}) = 2(1 - \hat{\theta})$.

$$E_{\hat{F}}[\theta|\theta > \hat{\theta}] = \frac{1 + \hat{\theta}}{2}, \qquad E_{\hat{F}}[\theta|\theta < \hat{\theta}] = \frac{\frac{1}{4} - \frac{\hat{C}}{2} + \hat{C}\hat{\theta}^2}{1 - 2\hat{C} + 2\hat{\theta}\hat{C}}.$$

$$\hat{C} = (1-\hat{\theta})\frac{\frac{1}{2} + \hat{\theta} - \hat{C}}{1 - 2\hat{C} + 2\hat{\theta}\hat{C}} \Rightarrow \hat{C}(\hat{\theta}) = \frac{2 - \hat{\theta} - \sqrt{-4\hat{\theta} + 13\hat{\theta}^2 - 8\hat{\theta}^3}}{4(1 - \hat{\theta})}.$$

That is, there is a unique equilibrium cost cutoff $\hat{C}(\hat{\theta})$ for any labeling cutoff $\hat{\theta}$. Moreover, $\hat{C}(\hat{\theta})$ is increasing in $\hat{\theta}$ for $\hat{\theta} \in (0, 0.5)$ and is decreasing in $\hat{\theta}$ for $\hat{\theta} \in (0.5, 1)$. Therefore, to maximize $\hat{C}(\hat{\theta})$, $\theta^* = 0.5$, which results in $\hat{C}(\theta^*) = 0.5$.

With $\hat{C} = 0.5$, the quality distribution $\hat{F}$ is uniform over $[0, 1]$. Therefore, the expectation of quality given the positive term is 0.75. To implement the market-optimal labeling cutoff θ^*, the quality-based policy sets $\hat{\theta} = 0.5$. That is, the court finds liable any case where the true quality is below 0.5. Note that the buyer's expected quality given the positive term is however 0.75, so the court will have to dismiss cases where buyers are misled by an amount less than $0.75 - 0.5 = 0.25$. The market-optimal policy for this numeric example can therefore be phrased in two ways:

1. Quality-based: a court policy that finds the seller liable if and only if $\theta < 0.5$.
2. Belief-based: a court policy that finds the seller liable if and only if the amount of belief inflation is greater than 0.25.[16]

[16] Under this policy, however, there can be another non-market-optimal equilibrium where the labeling cutoff is approximately 0.28. At this labeling cutoff, the effort is lower than at the market-optimal labeling cutoff of 0.5, so the expectation of quality given a positive term is also lower, while the gap between the expectation and the labeling cutoff happens to be also 0.25.

In practice, the appropriate amounts of deviation in belief are hard to articulate and would likely have to be expressed in terms of price differentials. For example, in the numeric example assuming values are measured in dollars, the market-optimal belief-based policy could say that if the buyers over-paid by less than 25 cents (or a quarter of the total premium a buyer is willing to pay for perfection), then the buyers are deemed to have been only moderately misled and the seller is therefore not liable.

4. Some buyers are gullible

Our base model assumes that the buyers are all rational and Bayesian. They have a correct understanding of the seller's equilibrium strategy and the legal environment, and makes inferences of the labels accordingly. It has been argued that buyers do not exert the mental energy to interpret the advertising language carefully.[17] It remains a question what these lazy buyers actually think when they purchase. When they do not exert effort to interpret the label, do they assume any labeling a complete lie (extremely skeptical) or do they take the literal meaning of the label (extremely gullible)? The latter is more plausible, as skeptical buyers tend to think more. A disagreement over whether buyers are gullible or sophisticated can lead to different rulings over the same case based on the same "reasonable consumer" standard. For example, 7th court district judge dismissed the claim that "100% grated Parmesan cheese" labeling is deceptive because no reasonable consumer would expect shelf-stable cheese to be completely free of preservatives. However, the U.S. Court's Appeal's judge thought it was reasonable for consumers to believe Parmesan cheese can be shelf-stable and reversed the lower court's ruling.[18]

In this section, I allow the court to define "reasonable consumer" as potentially being gullible as well. Let $\delta \in (0, 1)$ denote the proportion of the buyers that are gullible: they believe that only sellers with products 100% natural will use the label "all natural" regardless of the court's policy. That is, the expectation of quality is $\theta = 1$ if the positive term is used and $E_{\hat{F}}[\theta]$ if the positive term is not used. [19] Notice that such an interpretation is more favorable to the seller than the one made by the rational buyers under both a positive term and the lack thereof.

[17]See Klass (2020) p11.

[18]In re: 100% Grated Parmesan Cheese Marketing and Sales Practices Litigation, U.S. District Court, Northern District of Illinois, No. 16-05802.

[19]The gullible buyers' belief does not have to this extreme. It just needs to be more optimistic about the quality underlying the positive term than the actual distribution of qualities. They can, for example, believe that seller with $\theta > \tilde{\theta}$ uses the positive term with $\tilde{\theta}$ strictly greater than the actual labeling cutoff $\hat{\theta}$.

When the proportion of the gullible buyers is limited, the seller will still price according to the rational buyers' willingness-to-pay. This ensures that the market is still fully covered by the seller and the channel through which misleading the buyers can increase social welfare in Robert and Wilson (2018) is excluded. A sufficient but not necessary condition for full market coverage is

$$\delta < \min\left\{\frac{v + E_L[\theta]}{v+1}, \frac{v}{v + E_H[\theta]}\right\}$$

The condition $\delta < (v + E_L[\theta])/(v+1)$ implies that for any $\hat{F}$, a seller who has used the positive label is better off charging a price equal to the rational buyers' willingness-to-pay and selling to all buyers than charging a higher price and selling only to the gullible buyers. The condition $\delta < v/(v + E_H[\theta])$ implies that for any $\hat{F}$, a seller who has not used the positive term is better off charging a price equal to the rational buyers' willingness-to-pay and selling to all buyers than charging a higher price and selling only to the gullible buyers. In our earlier numeric example, this condition is satisfied if the proportion of gullible buyers is below $8/11$ which is approximately 73%. Under this assumption on the fraction of the gullible buyers, the equilibrium characterization under policy $\hat{\theta}$ (Proposition 1) extends verbatim.

The presence of the gullible buyers means that the belief of the "reasonable consumer" is a weighted average of those of the rational buyers and the gullible buyers. That is, in equilibrium, the reasonable consumer belief following a positive labeling term is $(1-\delta)\theta_\Gamma + \delta$, reflecting that a gullible buyer's belief is extreme at 1. Accordingly, Definition 1 should be updated:

Definition 2. A policy $\hat{\theta}$ follows the "reasonable consumer" standard if the buyers' belief in the equilibrium given this policy satisfies $(1-\delta)\theta_\Gamma + \delta = \hat{\theta}$.

As in the main setup, Corollary 1 holds: The only policy that satisfies the "reasonable consumer" standard is where $\hat{\theta} = 1$. A "reasonable consumer" standard that partially depends on the inflated belief of the gullible ones is higher than the standard that only depends on the rational belief. This standard also drives sellers with any quality less than perfection away from using the positive labeling term.

The quality-based market-optimal policy stays the same as in the main model, but because the expectation of a "reasonable consumer" is now higher as the court accepts a gullible buyer as "reasonable" as well, the belief-based market-optimal policy should effectively permit a larger extent of misleading the buyers. In the earlier numeric example, instead of permitting a belief inflation of 0.25 when all buyers are rational, the court should permit a belief distortion of $(1-\delta)0.75 + \delta - 0.5 = 0.5 + 0.25\delta$ when the δ proportion is gullible.

5. Certification as a quality-based policy

A quality-based policy in our setup is equivalent to relying on certification by a third party to build credibility of the labeling terms. For example, in U.S., certification organizations such as the U.S. Department of Agriculture (USDA) certifies the term "organic" and the non-profit organization, Non-GMO Project, for the term "Non-GMO". Using a positive term without the approval of these organization or defrauding the certification organizations subjects the sellers to fines and litigation. For example, Randy Constant, a Missouri farmer, was sentenced to over 10 years in prison for selling non-organic grains as organic (New York Times, November 8 2021) and a South Dakota man was sentenced to 51 months in prison for selling non-organic seeds as organic (AP news, February 23 2021). Instead of relying on buyer's belief as a benchmark, certification agencies set a detailed standard and grant the certification based on evidences from testing and inspection.

The existence of a single certification organization responsible for a positive term largely takes away the need for the court to decide whether a labeling term has misled the consumers. Instead, the court only needs to decide whether the product was actually certified or not and whether the seller has defrauded or corrupted the certification agencies. Therefore, there is barely any misleading labeling lawsuits targeting the term "organic". Consistent with the policy suggestion from our base model, the U.S. certification of "organic" is based on a standard that is not at the extreme of the quality spectrum. For example, there is a list of synthetic ingredients that are allowed in an "organic crop".[20] Therefore, certification theoretically allows targeting a specific desirable labeling cutoff: θ^*. Relying on the certification organization also has the advantage of not using the court as a tool for social engineering.

When there are multiple certification organizations that offer certification of similar terms, but have different standards, the court may still be called in to decide whether the labeling is misleading and using the reasonable consumer standard can again cause a positive term to be under-used. This happened with the case Latiff v. Nestle USA Inc. Nestle labelled some products with "No GMO Ingredients" which was certified by a third party named SGS. However the standard behind this term is less stringent than that behind the term "Non GMO project verified". For example, SGS allows diary or meat from animals that are fed non-GMO feed, but the latter does not. US district Judge Otis D. Wright rejected the motion to dismiss and said in a 2019 order that the plaintiff had "adequately alleged that Defendant's No GMO Ingredients Label could deceive a reasonable consumer". This suggests that, there is an advantage in allowing only one certification organization. However, the downside is the

[20]See §205.601 "synthetic substances allowed for use in organic crop production".

potential ability of a monopolistic certification organization to extract too much certification fee from the sellers, which dampens the incentive to provide quality.

6. Discussion

6.1. Uncertainty in the court's standard

In the base model analysis, the court uses a definitive cutoff quality type $\hat{\theta}$. The analysis can extend to accommodate any uncertainty in the court's policy or how a policy translates to legal outcomes.

Suppose, instead of choosing one cutoff, the court chooses a distribution H of quality cutoff over $[0,1]$. Let $\varepsilon = 0$. Then, the probability of quality-type θ to be liable for using the positive term is equal to the probability that this cutoff falls above θ. That is, type θ's probability of being liable when using the positive term is $1 - H(\theta)$. Therefore, the net benefit of using the positive term is:

$$B_H(\theta, \hat{\theta}) \equiv H(\theta)(E_{\hat{F}}[\theta|\theta > \hat{\theta}] - E_{\hat{F}}[\theta|\theta < \hat{\theta}]) - (1 - H(\theta))K$$

This benefit is strictly increasing in θ. Therefore, the equilibrium has a labeling cutoff, $\hat{\theta}$, that is determined by H through the condition $B_H(\hat{\theta}, \hat{\theta}) = 0$. Then the equilibrium cost cutoff is determined by:

$$\hat{C} = \int_{\hat{\theta}}^{1} B_H(\theta, \hat{\theta}) dF_H(\theta) - \int_{\hat{\theta}}^{1} B_H(\theta, \hat{\theta}) dF_L(\theta)$$

where $\hat{F}(\hat{C}; \theta) = G(\hat{C})F_H(\theta) + (1 - G(\hat{C}))F_L(\theta)$.

In the base model, the policy influences the effort cutoff only through influencing the labeling cutoff. Here, the policy H directly influences the effort cutoff. In the earlier numeric example, to implement a labeling cutoff $\hat{\theta} = 0.5$, a sufficient and necessary condition is that $H(0.5) = 1/6$. To see that, a seller with quality 0.5 is indifferent between using and not using the positive term if and only if:

$$0 = H(0.5)(0.75 - 0.25) - (1 - H(0.5))0.1 \Rightarrow H(0.5) = \frac{1}{6}.$$

A uniform distribution over $[0.4, 1]$ preserves 0.5 as the market-optimal labeling cutoff because $\hat{C}(\hat{\theta})$ is still maximized at $\hat{\theta} = 0.5$, as shown in the Appendix. That is, the uncertain liability cutoff should have $1/6$ probability of being below quality level 0.5, which is still a violation of the reasonable consumer standard given that the rational consumers' expectation is at 0.75

and the gullible consumers' expectation is at 1.

6.2. Clearer language

Certain labeling terms are less ambiguous in meanings. For example, "no artificial sweeteners", "no preservatives", "no artificial flavor" and etc. Typically, a potential plaintiff can find hard evidence through testing the ingredients of the product and the legal question is a simpler matter of finding the facts about the ingredients.[21] However, these terms suffer from a similar problem of not being able to partition the quality space somewhere in the middle, for example, the lack of the term "no preservatives" does not distinguish between a little amount of preservatives and a large amount of preservatives. An argument can be made that these dimensions of qualities are more certain and controllable by the seller. With sufficient effort, the seller can ensure that the product has zero amount of preservatives. When the reasonable consumer standard renders the term "natural" unable to communicate meaningful information, the seller would choose to use the limited space on the packaging for terms like "no preservatives", which motivates the seller to improve the sub-dimension of quality regarding the preservatives.

7. Conclusion

This paper gives the "reasonable consumer" standard widely used in the false advertising litigation a rigorous meaning that is consistent with the Bayes Rule. When interpreting the advertising languages they see, consumers are sophisticated enough to take into account the potential litigation's impact on the sellers and their resulting behavior. This paper argues for the court to be more lenient in the form of not punishing cases that only moderately mislead the consumers so as to maintain the seller's incentive to use the positive term so that the labeling will transmit valuable information about the product quality to the buyers, which will in turn sustain the sellers' motivation in improving the quality.

The Bayes' Rule has been incorporated into the analysis of how jury can update their beliefs in the face of evidences.[22] However, for litigation where liability depends on whether information receivers were misled, Bayes' Rule is also extremely relevant. This goes beyond the labeling or advertising litigation, to also, for example, allegations of libel, accounting fraud, and lack of disclosure, where Bayes' Rule should be very relevant to the questions of whether

[21] Occasionally, those terms carry some ambiguity as well. In Branca v. Bai Brands, LLC, the plaintiff and defendant argued over whether a synthetic form of malic acid is a "flavor" or not in light of the phrase "no artificial flavor" on the packaging. The case settled prior to class certification.

[22] See Ayres and Nalebuff (2015).

the plaintiffs were harmed and if so by how much. This paper hopefully leads to more interest in incorporating Bayes' Rule into these analysis.

References

[1] Akerlof, G.A., 1978. The market for “lemons”: Quality uncertainty and the market mechanism. Uncertainty in Economics, pp. 235-251. Academic Press.

[2] Anderson, S.P. and Renault, R., 2006. Advertising content. American Economic Review, 96(1), pp.93-113.

[3] Ayres, Ian, and Barry Nalebuff. ”The rule of probabilities: a practical approach for applying Bayes’ rule to the analysis of DNA evidence.” Stan. L. Rev. 67, pp.1447.

[4] Chakraborty, A. and Harbaugh, R., 2014. Persuasive puffery. Marketing Science, 33(3), pp.382-400.

[5] Cho, I.K. and Kreps, D.M., 1987. Signaling games and stable equilibria. The Quarterly Journal of Economics, 102(2), pp.179-221.

[6] Corts, K.S., 2013. Prohibitions on false and unsubstantiated claims: Inducing the acquisition and revelation of information through competition policy. The Journal of Law and Economics, 56(2), pp.453-486.

[7] Corts K. 2014. Finite Optimal Penalties for False Advertising. Journal of Industrial Economics, 62, pp.661-681.

[8] Cooper J. and Kobayashi B., 2021. Equitable Monetary Relief under the FTC Act: An Opportunity for a Marginal Improvement. Antitrust Law Journal, 83(3), pp.645-695.

[9] Esterl M. 2013. Some Food Companies Ditch “Natral” Label. Wall Street Journal, November 6, 2013.

[10] Harbaugh, R. and Rasmusen, E., 2018. Coarse grades: Informing the public by withholding information. American Economic Journal: Microeconomics, 10(1), pp.210-35.

[11] Hoffman, D.A., 2005. The best puffery article ever. Iowa L. Rev., 91, pp.1395.

[12] Kartik, N., 2009. Strategic communication with lying costs. The Review of Economic Studies, 76(4), pp.1359-1395.

[13] Klass, G., 2020. False Advertising Law and New Private Law. False Advertising Law, in Oxford Handbook of New Private Law (Andrew Gold et al. eds., Oxford Univ. Pr. forthcoming).

[14] Kugler, M. B., and L. J. Strahilevitz. 2017. The Myth of Fourth Amendment Scrutiny. U. Chi. L. Rev. 84, pp.1747.

[15] Miller, Alan D., and Ronen Perry. 2021. The reasonable person. NYUL Rev. 87, pp.323.

[16] Negowetti, N.E., 2013. Defining natural foods: The search for a natural law. Regent UL Rev., 26, pp.329.

[17] Posner, R.A., 1979. The Uncertain Protection of Privacy by the Supreme Court, S Ct Rev 173, pp.188.

[18] Rhodes, A. and Wilson, C. M. 2018. False advertising. The RAND Journal of Economics, 49(2), pp.348-369.

[19] Silverman, C., 2018. In Search of the Reasonable Consumer: When Courts Find Food Class Action Litigation Goes Too Far. U. Cin. L. Rev., 86, p.1.

8. Appendix

Proof of Lemma 1.

Proof. We first prove existence of such a no-price-signal subgame equilibrium. Since prices do not vary according to the quality, buyers' expectation given Γ remains at θ_Γ and buyers' expectation given ϕ remains at θ_ϕ after observing the price. Therefore, buyers will reject prices above $v + \theta_\Gamma$ and $v + \theta_\phi$ respectively.

Define the following off-equilibrium belief. If any seller with message Γ charges a price other than $v + \theta_\Gamma$, the buyers keep the same belief that the expected quality is θ_Γ. If any seller with ϕ charges a price other than $v + \theta_\phi$, the buyers keep the same belief that the expected quality is θ_ϕ.

Let $L(\theta)$ denote the probability of type θ being found liable after using the positive term Γ. Because of uncertainty ε, for any θ, $L(\theta) \in (0, 1)$. Then, for quality-types who have used Γ, their payoff as a function of price p is $(1 - L(\theta))p + L(\theta)(v + \theta_\phi - K)$ if the buyers buy, which is strictly increasing in p, so their optimal price is $v + \theta_\Gamma$ as higher prices will be rejected by the buyers. For quality-types who have used ϕ, they are not sued so their payoff is p if the buyers buy, which is also increasing in p, so their optimal price is $v + \theta_\phi$. This concludes the proof of the existence.

Next, we show that there is no equilibrium where price signals. Suppose not. First suppose there are at least two distinct prices that are charged $p_1 \neq p_2$ and they result in different expectation of qualities $\beta(p_1) < \beta(p_2)$, after observing message ϕ. Because there is no cost difference for different quality-types and all quality-types sell to all buyers, profit is strictly increasing in the belief of the buyers. Then, a type that charges p_1 wants to deviate to p_2. A contradiction.

Second, suppose there are at least two distinct prices that are charged and accepted by buyers: $p_1 < p_2$ after observing message Γ. Any quality-type that charges a price that is not accepted by the buyers would deviate to charging v which would be accepted even if the buyers hold the worst possible belief about the quality. Pick a type that charges p_1 and label it θ_1. The proposed equilibrium payoff for θ_1 is $(1 - L(\theta_1)p_1 + L(\theta_1)(v + \theta_\phi - K) < (1 - L(\theta_1))p_2 + L(\theta_1)(v + \theta_\phi - K)$. That is, θ_1 wants to deviate to p_2 because the legal liability does not depend on the price but only on the true quality. $\square$

Numeric example with uncertain liability cutoff

Here, I show that under any H distribution of the liability cutoff, the effort is maximized when the labeling cutoff is 0.5 in the numeric example.

Let $D(\hat{\theta}, \hat{C}) \equiv E_{\hat{F}}[\theta|\theta > \hat{\theta}] - E_{\hat{F}}[\theta|\theta < \hat{\theta}]$ denote the expectation differential. It is a function of the labeling cutoff $\hat{\theta}$ and the effort cutoff $\hat{C}$, and

$$\hat{C} = (D(\hat{\theta}, \hat{C}) + K)\left(\int_{\hat{\theta}}^{1} H(\theta)dF_H(\theta) - \int_{\hat{\theta}}^{1} H(\theta)dF_L(\theta)\right) - K(F_L(\hat{\theta}) - F_H(\hat{\theta})).$$

For the numeric example,

$$D(\hat{\theta}, \hat{C}) = \begin{cases} \frac{\frac{1}{2}+\hat{C}-\hat{\theta}}{2(1-2\hat{\theta}+2\hat{\theta}\hat{C})} \text{ for } \hat{\theta} \in (0, 0.5] \\ \frac{\frac{1}{2}+\hat{\theta}-\hat{C}}{2(1-2\hat{C}+2\hat{\theta}\hat{C})} \text{ for } \hat{\theta} \in [0.5, 1) \end{cases}$$

For $\hat{\theta} \in (0, 0.5]$, $\hat{C}$ satisfies:

$$\hat{C} = (D(\hat{\theta}, \hat{C}) + K)\left(\int_{0.5}^{1} H(\theta)dF_H(\theta) - \int_{\hat{\theta}}^{0.5} H(\theta)dF_L(\theta)\right) - 2(0.5 - \hat{\theta})K$$

$$\frac{\partial}{\partial\hat{\theta}} D(\hat{\theta}, \hat{C}) = \frac{\hat{C}(1 - 2\hat{C})}{2(1 - 2\hat{\theta} + 2\hat{\theta}\hat{C})^2} > 0$$

$$\frac{\partial}{\partial \hat{C}} D(\hat{\theta}, \hat{C}) = \frac{1 - \hat{\theta} - 2\hat{\theta}^2}{2(1 - 2\hat{\theta} + 2\hat{\theta}\hat{C})^2} < 1$$

Because $\hat{C} \leq C^* < 0.5$ and because the term $\int_{0.5}^{1} H(\theta) dF_H(\theta) - \int_{\hat{\theta}}^{0.5} H(\theta) dF_L(\theta)$ is strictly increasing in $\hat{\theta}$, implicit function theorem implies that $\hat{C}(\hat{\theta})$ is weakly increasing in $\hat{\theta}$.

For $\hat{\theta} \in [0.5, 1)$, $\hat{C}$ satisfies:

$$\hat{C} = (D(\hat{\theta}, \hat{C}) + K) \int_{\hat{\theta}}^{1} H(\theta) dF_H(\theta) - 2(1 - \hat{\theta})K$$

$$\frac{\partial}{\partial \hat{\theta}} D(\hat{\theta}, \hat{C}) = \frac{(1 - \hat{C})(1 - 2\hat{C})}{2(1 - 2\hat{C} + 2\hat{\theta}\hat{C})^2} > 0$$

$$\frac{\partial}{\partial \hat{C}} D(\hat{\theta}, \hat{C}) = \frac{\hat{\theta}(1 - 2\hat{\theta})}{2(1 - 2\hat{C} + 2\hat{\theta}\hat{C})^2} < 0$$

Implicit function theorem implies that $\hat{C}(\hat{\theta})$ is strictly decreasing in $\hat{\theta}$. Therefore, $\hat{C}(\hat{\theta})$ is maximized at $\hat{\theta} = 0.5$.

At $\hat{\theta} = 0.5$, the belief differential does not depend on $\hat{C}$: $D(0.5, \hat{C}) = 0.5$, Then when H is a uniform distribution over $[0.4, 1]$,

$$\hat{C} = (0.5 + 0.1)\frac{7}{12} - 0.1 = 0.25$$

This effort cutoff is lower than the one in the numeric example for the base model with no uncertainty because the significant uncertainty here in H makes a seller liable for using the positive term with a positive probability.