Prohibitions on False and Unsubstantiated Claims: Inducing the Acquisition and Revelation of Information through Competition Policy

Kenneth S. Corts*
Rotman School of Management
University of Toronto

June 10, 2011

Abstract

Competition law in many countries prohibits firms from making false claims about product quality or performance and also requires that the truth of specific claims be supported by adequate prior testing. This paper explores the differences between these two policies and asks, among other questions, whether a policy of mandatory prior substantiation is redundant to a ban on false claims. This paper develops a model in which firms have private information about their type–represented by their probability of having a high quality product–and can acquire additional private information about their product quality through costly testing and learning. Penalties for false claims and for unsubstantiated claims create an opportunity for firms to credibly reveal their information and for signaling to emerge in equilibrium. I show that the two kinds of penalties affect the possibility of signaling in different ways, and that the mandatory substantiation requirement in many circumstances improves buyer information and social welfare beyond what is achieved by a ban on false claims alone. It is therefore not redundant to a false claims ban, but is a useful additional policy tool in many markets characterized by asymmetric information.

1 Introduction

Competition law in many countries prohibits firms from making false claims about product quality and performance and also requires that specific performance claims be supported by adequate prior testing. This paper explores the differences between these two policies and asks, among other questions, whether a policy of mandatory prior substantiation is redundant to a false claims ban. This is a timely and important question in Canadian competition policy; in a number of recent cases firms have mounted a defense to charges of unsubstantiated claims in part by arguing that the mandatory prior substantiation requirement should be held to be an unconstitutional infringement of free speech.

*Rotman School of Management, 105 St. George St., Toronto, Ontario M5S 3E6; kenneth.corts@rotman.utoronto.ca.
I thank Abigail Brown, Joe Farrell, Tom Ross, and seminar participants at UC-Berkeley, the University of Rochester, Queen's University, the Federal Trade Commission, the 2008 Canadian Economics Association meetings, and the 2009 International Society for the New Institutional Economics meetings for helpful comments.

As I will describe later, this is grounded in a claim that mandatory prior substantiation is both overly broad and redundant to the false claims ban–in particular, that unsubstantiated claims that are false are prohibited by the false claims ban, and that unsubstantiated claims that are true are not problematic and should not be penalized. Such a defense has succeeded in one recent case and failed in another. This is also a relevant policy question in the US, as a Federal Trade Commission policy statement on the subject suggests ambivalence toward enforcement of the mandatory substantiation requirement, which is also present in US competition law. I show in this paper that mandatory prior substantiation is not in general redundant to a false claims ban, and that in fact both penalties used together can achieve better outcomes than either policy used alone.

I develop a model in which firms have private information about their type, which reflects their probability of having a high quality product, and may acquire additional private information about their actual product quality through costly testing and learning. Penalties for false and/or unsubstantiated claims can support two types of signaling in such a model. Under what I term "strong signaling", firms choose to learn their product qualities for certain and then accurately and completely signal this information to buyers through quality claims that are subject to these penalties. In such a signaling equilibrium, buyers correctly interpret the signal to imply that the products are in fact high quality with certainty. Under what I term "weak signaling", high type firms remain uninformed of their actual product quality but nonetheless choose to signal their types to buyers through "speculative claims" of high quality that are subject to these penalties. In such a signaling equilibrium, buyers correctly interpret the signal to imply that products are of high quality only with some probability. While this provides less information than strong signaling, it nonetheless improves buyer information relative to pooling.

In general, gross social welfare increases with the fineness of information provided to consumers, so that weak separation is always preferred to pooling, and strong separation is preferred to weak separation if testing costs are low enough. I show that for given parameter values penalties can typically be set to induce the socially optimal level of information revelation and that in general this can be accomplished through either type of penalty. However, this is true only for unconstrained penalties and fixed parameter values. If the range of penalties is constrained or if a single penalty will be applied across a range of parameter values, there are three shortcomings of the false claims ban, which I describe below. I show that each of these shortcomings of a false claims ban is mitigated by use of a mandatory substantiation policy, while this substantiation policy has one specific shortcoming of its own. I argue that, in general, a socially optimal policy employs penalties for both false claims and unsubstantiated claims.

1. For a "small but reasonable penalty" equal to the unfairly gained profits due to misrepresentation of quality, there is no strong separation under a false claims ban alone. This penalty makes firms exactly indifferent between making the speculative claim, on one hand, and learning and telling the truth, on the other, but provides no incentive to bear a positive learning cost. Under a mandatory prior substantiation policy with a similar penalty, there is equilibrium strong signaling if learning costs are low. The difference is that optimistic firms don't discount this substantiation penalty; it therefore makes learning, from which they are fairly certain to receive good news, relatively more attractive for these firms.

2. For any given penalty, there is no strong separation for optimistic enough high type firms under a false claims ban alone. Such a firm discounts the false claims penalty, sometimes significantly,

because it believes it there is a good chance its quality claim is true. In contrast, even a firm that is fairly certain it has a high quality product knows it is certainly violating the prior substantiation requirement by making a speculative claim, and thus does not discount that penalty. The mandatory prior substantiation penalty is therefore a more powerful tool for increasing the firm's incentive to invest in learning product quality, especially when firms are optimistic about product quality.

3. Increasing a false claims penalty alone may induce inefficient learning by pessimistic high type firms. Because optimistic firms discount the false claims penalty more than less optimistic firms, a less optimistic firm may be driven to socially wasteful learning by a penalty designed to encourage socially beneficial signaling by a more optimistic firm. Because the mandatory prior substantiation penalty generates a relatively more binding constraint for more optimistic firms, having this second lever available allows the regulator to target penalties in a way that strikes an appropriate balance between the social benefits of learning and its social costs.

4. Increasing either a false claims penalty or a prior substantiation penalty shrinks the range of parameters for which weak signaling is supported; increasing a prior substantiation penalty shrinks the weak separation region more quickly than increasing a false claims penalty. In the model, consumers understand that when firms make speculative claims of high quality (as under weak signaling) these must be discounted. This requires for internal consistency that firms do not learn their type. However, stronger penalties for false or unsubstantiated claims increase the incentive to learn product quality, thereby reducing the willingness of firms to engage in this type of signaling. This problem is more pronounced for the substantiation penalty because it is not discounted by the optimistic firms who are the source of speculative claims under weak signaling.

This paper contributes to the literature by making the first explicit theoretical inquiry into the effects of mandatory substantiation policies. There is certainly a great deal of theoretical work on the effects of asymmetric information in markets (the literature on lemons problems and adverse selection), on strategies firms may undertake to credibly reveal their information (the literature on signaling and certification), and on policies governments may employ to improve information provision (the literature on mandatory disclosure and liability rules). However, there appears to be no existing treatment of mandatory substantiation policies, which are quite different from mandatory disclosure and other related and well-studied policies. While mandatory disclosure rules require that firms inform consumers of all information acquired through testing, mandatory substantiation requires that all information provided to customers be supported by testing. The former prohibits firms from hiding bad news, which generally reduces incentives to acquire information; the latter prohibits firms from making speculative claims, which generally increases incentives to acquire information.

Akerlof (1970) is the seminal citation in this literature. That paper lays out the basic model of informed sellers selling to uninformed buyers and establishes the iterative unraveling process by which consumers revise down their assessment of product quality as they consider which firms would be unwilling to sell at a price reflecting the average quality of the remaining firms. Viscusi (1978) and Grossman (1981) apply this same logic to a setting in which informed sellers can credibly reveal product quality; in this case the uninformed or pooled portion of the market (at least partially) unravels because high-quality firms reveal their information, lowering consumer expectations of average quality among pooled firms and inducing further information revelation. This leads to complete revelation of information if certification (i.e., a mechanism for credible information revelation) is costless, as in Grossman's model, or partial revelation if it is costly, as in Viscusi's model. Farrell (1986) and

Matthews and Postlewaite (1985) pursue a similar analysis but focus on firms that actively undertake learning to become informed. Both papers show that mandatory disclosure regulations dampen the incentive to acquire information and may obstruct the unraveling process and prevent complete information revelation. A large literature on mandatory disclosure has since followed (including an offshoot literature in financial economics focusing specifically on financial disclosure regulations), including papers by Polinsky and Shavell (2006), who investigate how the liability regime affects the impact of mandatory disclosure on information revelation when firms must invest to learn product quality, and Daughety and Reinganum (2008), who analyze how the liability regime affects an exogenously informed firm's decision of whether to signal its information through certification or through prices.

Section 2 provides context for this analysis by discussing the prevailing legal regime in Canada and the US. Section 3 lays out the basic model, while section 4 provides both graphical intuition and analytical results demonstrating the differences between the two types of policies. Section 5 concludes.

2 Background

In Canada, the existence of both a false claims prohibition and mandatory substantiation is quite clear. Section 74.01(1)(a) of the *Competition Act* prohibits any "representation to the public that is false or misleading in a material respect," while section 74.01(1)(b) prohibits any "representation to the public ... of the performance, efficacy, or length of life of the product that is not based on an adequate and proper test thereof..."[1] If the Competition Tribunal finds that a firm has violated these provisions, it may order that the firm cease making the suspect claims and/or engage in corrective advertising and it may levy administrative fines of up to $100,000 for the first violation and $200,000 per violation thereafter.

In the US, the situation is less clear, but the *FTC Policy Statement Regarding Advertising Substantiation* (which was issued in the 1980s but is prominently featured alongside other and more recent guidance documents) summarizes the FTC's position.[2] This document states, seemingly quite clearly, that the FTC "intends to continue vigorous enforcement of this existing legal requirement that advertisers substantiate express and implied claims..." and that "... a firm's failure to possess and rely upon a reasonable basis for objective claims constitutes an unfair and deceptive act or practice in violation of Section 5 of the FTC Act." However, the document goes on to elaborate on the admissibility of post-claim evidence–that is, substantiating tests performed after the representations of quality or performance were made to the public–and states that "the truth or falsity of a claim is always relevant to the Commission's deliberations. Therefore, it is important that the agency retain the discretion and flexibility to consider [post-claim] evidence..." The document proceeds to elaborate on the circumstances under which true but initially unsubstantiated claims will not be pursued as violations of the FTC Act.

In Canada, the mandatory substantiation policy clearly exists as a separate requirement, but it is also under attack. In two recent cases brought by the Commissioner of Competition under section 74.01(1)(b), the firms have defended themselves in part through a constitutional challenge to the

[1] http://laws.justice.gc.ca/en/ShowFullDoc/cs/C-34///en
[2] http://www.ftc.gov/bcp/guides/ad3subst.htm

prior substantiation requirement itself. Such challenges are based on the free speech protections of the Charter of Rights and Freedoms. In broad terms, commercial speech is protected speech, and the provisions of section 74.01(1) are deemed to be infringements of that speech. However, the Charter permits the government to infringe on such rights in some circumstances if such infringement is necessary to advance certain valid policy goals. To successfully justify the infringement, the government bears the burden of proof for certain tests pertaining to the value of the policy goals pursued and the nature of the infringement. Most relevant in these cases, the government must show that the infringement meets the test of "minimal impairment"–that is, that the same policy goal cannot readily be met with alternative restrictions that constitute a less significant infringement on the protected rights. The argument made by the firms in these cases is essentially that section 74.01(1)(a) already prohibits false claims and that 74.01(1)(b) is therefore an overly broad prohibition–catching as it does true but unsubstantiated claims–that fails to meet the minimal impairment test.

A defense along these lines has been made in two recent cases heard by the Competition Tribunal. In *Commissioner of Competition v. Gestion Lebski* (2006 Comp. Trib. 32), the Tribunal found that the government had presented no evidence to meet its burden of proof in justifying the infringement of free speech, and therefore held 74.01(1)(b) invalid for purposes of that case. However, as the government had charged violations of both 74.01(1)(a) and 74.01(1)(b) in this case, Gestion Lebski was ordered to pay a fine for violation of section 74.01(1)(a). In *Commissioner of Competition v. Imperial Brush Co. and Kel Kem (c.o.b. as Imperial Manufacturing Group)* (2008 Comp. Trib. 02), the government did present evidence in justification of its infringement of free speech, and the Tribunal held that 74.01(1)(b) was valid and applicable.[3] Imperial Manufacturing Group was ordered to undertake corrective advertsing and pay a fine for violation of 74.01(1)(b).

To fix ideas, it is useful to consider in more detail the latter of these cases recently heard by the Canadian Competition Tribunal.[4] Imperial Manufacturing Group manufactured and marketed, among other products, the "Supersweep Chimney Cleaning Log", an artificial log of the sort that one burns in a fireplace. The product packaging contained performance claims such as "Helps Prevent Chimney Fires" and "Helps Eliminate Dangerous Creosote In Your Chimney". The Commissioner of Competition requested evidence of tests supporting these performance claims. Imperial Manufacturing responded with results of tests they had conducted and with field-based evidence from the use of the log and its chemical additives in other forms over the course of years of experience. The Commissioner of Competition considered this evidence unsatisfactory due to a lack of replication and other issues and ultimately brought allegations of violation of 74.01(1)(b)—the substantiation provision of the Competition Act—before the Competition Tribunal. Imperial Manufacturing argued both that the substantiation provision itself was unconstitutional and also that they had relied on reasonable tests and field experience to support their performance claims. The Commissioner argued both that the substantiation requirement was constitutional and that these tests and field experience did not constitute an adequate and proper prior test. The Competition Tribunal ultimately found in favor of the Commissioner on both issues and ordered Imperial Manufacturing to undertake corrective advertising and pay a fine.

Note that the Commissioner did not allege that the performance claims were false in this case,

[3] I provided expert testimony for the Commissioner of Competition in the constitutional challenge aspect of this case.

[4] The following summary is gleaned from the filings available at http://www.ct-tc.gc.ca/CasesAffaires/CasesDetails-eng.asp?CaseID=278.

only that they were not backed by adequate prior testing. This case therefore highlights the difference between the false claims ban and the prior substantiation requirement.

It seems clear that the Commissioner's ability to request evidence of testing to support the performance claims, and then to take legal action solely on the failure of the firm to present such evidence to the Commissioner's satisfaction, greatly simplified the Commissioner's task and reduced the costs of investigating and prosecuting the case, compared to a situation in which the Commissioner would have had to contest the actual truth or falsity of the performance claims themselves. This difference between the substantiation requirement and a false claims ban could alone create enormous differences in the costliness of enforcement, the speed with which action can be taken, the number of cases that can be pursued with given resources, and so on. However, this paper abstracts from all of these considerations to consider other differences between false claims penalties and substantiation requirements that do not have to do with the cost of enforcement or any effect through an endogenous intensity of enforcement. Throughout the analysis, I assume simply that an exogenous enforcement regime leads to certain expected penalties for violation of either of these policies.

What the model emphasizes is not so much the harm caused by any particular unsubstantiated claim—in equilibrium consumers properly discount all claims to account for the true expected quality of the product—but the harm caused by the consumers' collective lack of confidence in quality claims when penalties are insufficient to dissuade firms from making speculative claims. The problem is not that consumers are duped by a particular claim (i.e., it is not that they take the Supersweep claims, which may turn out to be false, at face value), but rather that their skepticism prevents the firm from being fully rewarded for its actual high quality (i.e., knowing that the firm's incentives are not sufficient to ensure testing leads consumers to undervalue Supersweep logs in the case when the performance claims are in fact true). Strong penalties that ensure learning by the firm (including, as the analysis will show, penalties for violating a substantiation requirement) restore consumers' faith in performance claims and allow truly high quality products to be recognized as such. This improvement in consumer information in turn yields social benefits.

3 Model

This section lays out a simple one-shot game in which a firm with private information about its type–which reflects its probability of having a high quality product–sells to consumers in the presence of a regulator that may enforce penalties for false or unsubstantiated claims about product quality.

3.1 A game of asymmetric information

A single price-setting firm offers to potential buyers at a single price a product that may be of either high (H) or low (L) quality. A firm may be one of two types, where the type represents the firm's probability of having a high quality product. The firm knows its type, which may be either 0, meaning that it is certain that it is has a low-quality product, or 1, meaning that it has a $\gamma \in (0, 1)$ probability of having a high-quality product and a $1 - \gamma$ probability of having a low-quality product. The proportion of type 1 firms is λ.

The firm may learn with certainty its product quality at cost t; this represents an investment in testing. Both the act of testing/learning and the resulting knowledge of product quality are unverifiable to consumers, but verifiable to the regulator. The firm's type is not verifiable to either

consumers or the regulator, except inasmuch as it can be inferred with certainty by the regulator in the case of a product demonstrated through testing to be of high quality.

Demand for the product and therefore the firms' maximum profit depend on only the perceived quality of the product. Let π_i be the firm's maximum profits for a given consumer perception, where $i \in \{H, L, 1, 0, P\}$ denotes the consumer perception. Perceived quality of H or L corresponds to buyer belief that the product is certainly of that particular quality. Perceived quality of 1 or 0 corresponds to buyer belief that the firm is certainly of that type, and therefore that the product is of quality H with probability γ or 0, respectively. Perceived quality of P corresponds to buyer belief that the types are pooled, and therefore that the product is of quality H with probability $\lambda\gamma$. I assume that maximized profits are increasing in perceived quality: $\pi_H > \pi_1 > \pi_P > \pi_L = \pi_0 > 0$.

In an attempt to influence consumer perceptions, firms may make claims about product quality. Formally, I allow the firm to state that its product is low or high quality (L or H) or remain silent on quality. Restricting the set of potential claims in this way is without loss of generality, because this space is as rich as the information that is verifiable to the regulator. Since the regulator can verify and therefore condition penalties on only the actual product quality and not firm type, only claims about product quality will generate the differential costliness (through differential exposure to penalties) that can ensure credibility of the signal. The natural alternative to consider is that the firm signals its actual ex ante private information–that is, its type. Imagine that a firm does so by truthfully claiming that "I am a type 1 firm; my product will be of high quality a proportion γ of the time." One of two interpretations of the false claims ban and substantiation requirement must apply. One possibility is that this deemed to be an unverifiable claim and so not subject to these penalties. In this case, there is no differential cost to the signal that would permit its credibility. The other possibility is that this could be deemed to be a claim about actual product quality that is subject to penalties when false or unsubstantiated. In this case, this statement of type is equivalent to a statement of quality.

The model incorporates a competition policy authority that may enforce penalties for false or unsubstantiated claims. A "false claims ban" (FCB) prohibits statements of product quality that are in fact false. The exogenous and unspecified investigative and legal regime is such that firms that claim their product is high quality when in fact it is low quality face an expected penalty Δ_1. A "mandatory prior substantiation" (MPS) policy prohibits claims that are not based on prior testing, regardless of their actual truth or falsity. Firms that claim high product quality but have not paid the cost t to learn their product's quality face an expected penalty Δ_2. A firm that paid to learn its product quality and learned its product was low quality, but which nonetheless claimed high quality, is subject to both penalties.[5]

3.2 Timing

The game proceeds in the following stages.

1. Nature chooses the firm's type (0 or 1), which the firm observes.

2. The firm can a pay a cost t to learn its product quality (H or L).

[5] These penalties can be thought of expansively to include legal costs and reputational costs incurred upon the regulator's finding of improper claims. These "penalties" could even include reputational or other costs incurred independent of the regulator's activities if, for example, there were other routes of consumer discovery of quality. For simplicity, I simply assume that the regulator controls this expected penalty.

3. The firm can make a claim about its product quality (H or L) and sets price.

4. Consumers make inferences about product quality; demand and profits are realized.

5. The regulator assesses penalties for false and/or unsubstantiated claims.

3.3 Information revelation regimes

The analysis of this paper will focus on three different information revelation regimes. Much of the analysis will focus on which penalties support which regime, but for now it is useful just to define these regimes without regard to how they are achieved or supported. The pooling regime (denoted P) is the case in which no information is credibly revealed to buyers, and buyers therefore believe that the product is quality H with probability $\lambda\gamma$. Strong separation (denoted S) is the case in which product quality is perfectly known or inferred by buyers. Weak separation (denoted W) is the case in which firm type, but not product quality, is perfectly known or inferred by buyers. Buyers therefore believe that the product is high quality with probability γ if and only if the firm is known or inferred to be type 1.

While demand and profits of a particular firm depend on only consumer perceptions and are indexed by $i \in \{H, L, 1, 0, P\}$, ex ante expected total welfare (the metric by which one evaluates the social desirability of particular policies) depends on the regime that prevails (which implies a weighting over particular realizations of types and qualities and the consequent signalling behavior and consumer inferences). I will denote gross total welfare T_j, where $j \in \{P, S, W\}$ indexes the regime. This is defined gross of the learning/testing cost, which is incurred only by type 1 firms and only under strong signaling. Thus, the net total welfare comparison is among T_P, T_W, and $T_S - \lambda t$.

The social value of information is complex and beyond the scope of this paper. In a related paper (Corts, 2011) I demonstrate that in many familiar models it is true that gross total welfare is increasing in the fineness of the information that is conveyed to consumers–that is, $T_P < T_W < T_S$–and I will maintain that assumption throughout the paper. Note that when weak signaling can be achieved at no social cost–as will be the case in this paper–it is always preferred to pooling. Because the cost of testing and learning affects the net total welfare associated with strong signaling, strong signaling will generally be preferable when testing costs are low, with weak signaling preferable when testing costs are high. Clearly, as the learning cost t approaches 0, strong separation is always preferable.

The following two subsections derive the conditions on the two penalties under which weak and strong separation are feasible. Both weak separation and strong separation are standard signaling equilibria in which I require incentive compatibility for firms and consistency of inferences by consumers. In principle, three kinds of incentive compatibility constraints apply: participation constraints, constraints that ensure the appropriate amount of firm learning for that equilibrium, and incentive constraints that ensure truthful signaling conditional on the extent of learning. In fact, the participation constraints are irrelevant to this analysis since by assumption the firm makes nonnegative profit even under the most pessimistic consumer inferences.

3.4 Weak Separation

This subsection lays out the incentive constraints that are required to support the weak separation equilibrium. Recall that in this equilibrium, consumers infer that a claim of high quality implies the

firm is type 1 and that the absence of such a claim implies the firm is type 0, and that firms make high quality statements if and only if they are type 1. Thus, the firm in this case is making what I will call a "speculative claim"–that is, they are claiming high product quality when in fact they know only that this is a possibility, not a certainty.

The incentive compatibility constraints that ensure truth-telling conditional on knowing only one's type are:

$$IC_W(1) : \pi_1 - \pi_L \geq (1-\gamma)\Delta_1 + \Delta_2$$

$$IC_W(0) : \pi_1 - \pi_L \leq \Delta_1 + \Delta_2.$$

The left-hand side, $\pi_1 - \pi_L$, is the gross benefit to signaling. The right-hand sides of these inequalities represent the cost of signaling. In both cases the firm is exposed to the expected MPS penalty Δ_2 since it has not learned. The FCB penalty Δ_1 is added to the cost of signaling if the firm is type 0, but for the type 1 firm this is discounted by the probability $1-\gamma$ that the claim is false. This pair of incentive constraints simply ensures that signaling pays for the high type firms but not for the low type firms; this can hold because, though the benefit to signaling is the same for the two types, the exposure to penalties for false claims differs between the two types of firms.

For these to be the right incentive constraints and consumer inferences, it must be that the firm chooses not to learn its type. This leads to a "learning constraint" that ensures the firm does not have an incentive to learn, which would allow it to avoid making speculative claims, thereby eliminating its exposure to the FCB and MPS penalties:

$$LC_W(1) : \pi_1 - (1-\gamma)\Delta_1 - \Delta_2 \geq \gamma\pi_1 + (1-\gamma)\pi_L - t.$$

The left-hand side represents the net profit of making the speculative claim. The firm is identified as a type 1 but is subject to both types of penalties. The right-hand side reflects the alternative profit the firm would earn by learning. It is clear that a firm that learned it had a low-quality product would not signal, since this is precisely the condition given in $IC_W(0)$. Thus, the profit that follows from learning is the weighted average of the profits that accrue to a firm who has revealed itself as a type 1 or a type 0 firm, less the cost of learning. Such a firm has no exposure to either of the penalties.

3.5 Strong Separation

A similar set of inequalities determines the set of parameters for which strong separation can be sustained. Recall that in this equilibrium, consumers infer that a claim of high quality implies the firm has a high quality product and that the absence of such a claim implies the firm has a low quality product, and that the firm makes the high quality claim if and only if its product is in fact high quality. For this to be true, of course, the firm must pay to learn its product quality. This gives rise to a similar set of incentive and learning constraints.

The incentive constraints that ensure truth-telling conditional on knowing product quality are:

$$IC_S(H) : \pi_H - \pi_L \geq 0$$

$$IC_S(L) : \pi_H - \pi_L \leq \Delta_1 + \Delta_2.$$

As in the case of the weak separation incentive constraints, the left-hand side represents the gross benefit to signaling, while the right-hand side represents the cost of expected penalties. Note that a firm that has learned it has a high-quality product faces no exposure to either kind of penalty, while the firm who has learned that it has a low-quality product faces both expected penalties.[6]

For these to be the right incentive constraints, it must be that a type 1 firm pays to learn its true product quality. This generates two distinct learning constraints, one that requires learning to be more profitable than a speculative claim, and a second that requires learning to be more profitable than simply not signaling at all:

$$LC_S^1(1) : \gamma\pi_H + (1-\gamma)\pi_L - t \geq \pi_H - (1-\gamma)\Delta_1 - \Delta_2$$

$$LC_S^2(1) : \gamma\pi_H + (1-\gamma)\pi_L - t \geq \pi_L.$$

In both cases, the left-hand side represents the profit earned by learning and then truthfully signaling. In LC_S^1, the right-hand side represents the profits from making a speculative claim, including the exposure to penalties. In LC_S^2, the right-hand side reflects the profits the firm earns by simply not signaling and being viewed as a seller of a low-quality product. I will refer to LC_S^2 being satisfied as the condition that "learning is not cost-prohibitive".

4 Analysis

This section explores how the feasible information revelation regimes depend on the particular penalties chosen by the regulator. It also highlights the strengths and weaknesses of each type of penalty under different assumptions on the penalties and sheds light on how a net total welfare-maximizing regulator would chose optimal penalties.

4.1 Separating and pooling equilibrium existence

The inequalities laid out in section 3, representing the incentive and learning constraints, determine the set of parameters for which each type of signaling equilibrium exists. That is not to say that such an equilibrium is the only equilibrium for a given set of parameters that satisfies those constraints. Rather, given the fact that the participation constraints are trivially satisfied, it is straightforward that for any set of parameter values there exists a pooling equilibrium in which no firm signals; in such an equilibrium, consumers infer nothing from statements about product quality but believe that the firm's product is high quality with probability $\lambda\gamma$. That said, it is true that there is at most one type of signaling equilibrium for any set of parameter values. This follows immediately from the fact that LC_S^1 and LC_W cannot be simultaneously satisfied. Rearranging and combining these two constraints yields $\pi_H - \pi_L \leq \Delta_1 + \frac{\Delta_2 - t}{1-\gamma} \leq \pi_1 - \pi_L$, which contradicts the assumption that $\pi_H > \pi_1$. This proves the following proposition.

[6] While such a firm has undertaken testing and borne the learning cost, it nonetheless fails to have prior testing *that supports the high-quality claim* and is therefore subject to the substantiation penalty.

Proposition 1 *For any set of parameter values, there exists a pooling equilibrium and at most one type of separating equilibrium.*

One other general observation that can be made about the existence of a separating equilibrium is that no signaling equilibrium of either type will exist in the absence of regulatory penalties. It is only these penalties that create the differential cost of signaling that permits the signal to be credible and prevents imitation by low type firms or firms with low-quality products. This follows immediately from the fact that neither $IC_W(0)$ nor $IC_S(L)$ can hold at $\Delta_1 = \Delta_2 = 0$, given that the assumption that $\pi_H > \pi_L$ implies that $\pi_H - \pi_L > \pi_1 - \pi_L > 0$.

Proposition 2 *For any set of parameter values, no separating equilibrium of either type exists in the absence of regulatory intervention.*

4.2 Supporting socially optimal information revelation

For any given parameters it is straightforward that the socially optimal level of information revelation is strong separation if $T_S - \lambda t \geq T_W$ and weak separation if $T_W > T_S - \lambda t$. Pooling is never socially optimal. In addition, given the analysis above of the conditions under which strong and weak separation prevail, it is easy to see that penalties can always be chosen to implement weak signaling when it is optimal and they can often be chosen to implement strong separation when it is optimal.

Proposition 3 *For any set of parameter values, there exist penalties that implement the socially optimal information revelation regime if (1) weak separation is socially optimal, or (2) strong separation is socially optimal and learning is not cost-prohibitive (i.e., $LC_S^2(1)$, which does not depend on the penalties, is satisfied).*

For any particular parameter values, it is straightforward to find a penalty that will support weak separation. It is easy to check that, for example, $\Delta_1 = \pi_1 - \pi_L$ and $\Delta_2 = 0$ satisfy all the weak signaling incentive and learning constraints for any parameter values. Similarly, an arbitrarily high Δ_1 (or Δ_2) satisfies the $LC_S^1(1)$ constraint for all parameter values, supporting strong signaling for any parameters such that $LC_S^2(1)$, which does not depend on the penalties, is satisfied.

What this demonstrates is both that the determination of socially optimal penalties is fairly straightforward and that the availability of a substantiation penalty is not especially useful if the penalties are unconstrained and are being designed to apply to a single, specific set of parameter values. The situation is more complex if penalties are constrained or if a single penalty regime must be chosen to apply across multiple possible parameters either because the parameters are not yet known when penalties are fixed or because one set of penalties must apply across many industries or time periods over which parameters may vary. The analysis of these more general considerations is the subject of the remainder of this section.

4.3 A graphical representation

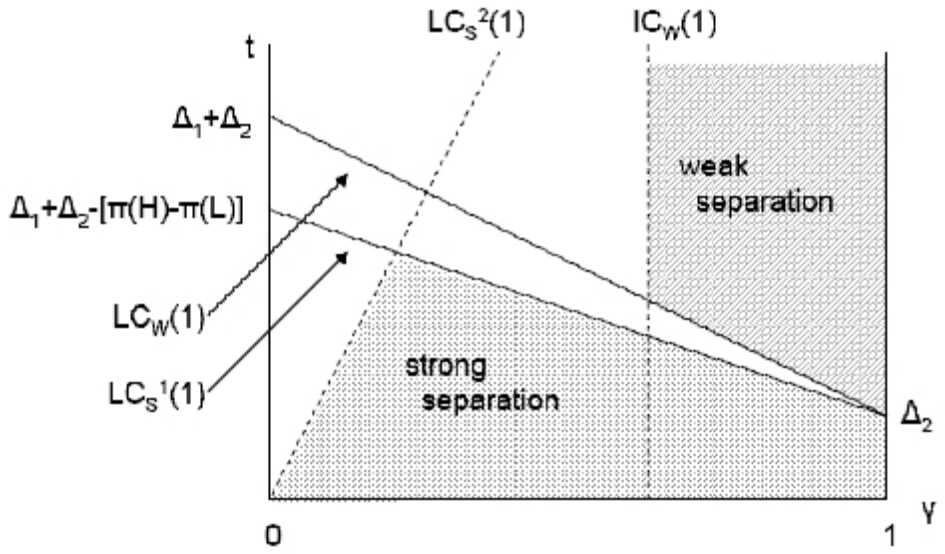
Figure 1 portrays in γ, t space all the relevant constraints that determine the boundaries of the regions in which strong and weak separation are feasible. Visualizing how these constraints shift with increases in one or both penalties is the clearest way to develop an intuitive understanding of the

differences between the two penalties and the tradeoffs that figure into the determination of optimal penalties. This figure is drawn for $\Delta_1 + \Delta_2 \geq \pi_H - \pi_L$, which ensures that $IC_S(L)$ and $IC_W(0)$ hold. In addition, $IC_S(H)$ holds throughout the paper by the assumptions on the profit functions. The remaining constraints are portrayed in the figure as follows.

Being below $LC_S^2(1)$ ensures that under strong separation inferences a type 1 firm would rather learn than make no high-quality claim. This is the condition referred to previously as learning not being prohibitively expensive. Being below $LC_S^1(1)$ ensures that under strong separation inferences a type 1 firm would rather learn (and then signal if and only if its product is high quality) than make a speculative high-quality claim.

Being to the right of $IC_W(1)$ ensures that under weak signaling inferences a type 1 firm would rather make an untested speculative high-quality claim than not signal. Being above $LC_W(1)$ ensures that under weak separation beliefs a type 1 firm will not learn in order to avoid exposure to the penalties (which is required for consistency with the weak signaling inferences). Note that while I have depicted the two learning constraints $LC_S^1(1)$ and $LC_W(1)$ as downsloping, they are not in fact necessarily downsloping in general. This will be clear when their intercepts are discussed below.

Figure 1



$LC_S^2(1)$ does not move with the penalties, as the comparison of learning and then truthfully signaling with simply not signaling does not depend on the penalties. This means that the range supporting strong signaling varies with the penalties only through movements in $LC_S^1(1)$. Note that the left intercept of $LC_S^1(1)$ $(\Delta_1 + \Delta_2 - (\pi_H - \pi_L))$ is determined by the sum of the penalties, while the right intercept (Δ_2) is determined by only the prior substantiation penalty Δ_2. The intuition for this is simple. Toward the left side of the figure–that is, with respect to firms that know they are almost certain to have a low quality product–the two penalties are nearly perfect substitutes. Making a speculative claim is very likely (in the limit, certain) to lead to paying the FCB penalty

Δ_1, while it is certain to lead to exposure to the MPS penalty Δ_2. In contrast, at the right side of the figure–that is, for firms that know they are almost certain to have a high quality product–the FCB penalty Δ_1 is of little concern. These high-γ firms discount the false claims penalty considerably (in the limit, completely). As a result, the mandatory substantiation penalty is generally a stronger incentive for learning, with this difference being more pronounced when γ is higher.

Now consider the constraints determining the region of weak signaling. $IC_W(1)$ shifts to the right with increases in either penalty, and moves faster for increases in the prior substantiation penalty Δ_2 than for increases in the false claims penalty Δ_1. $LC_W(1)$ moves up with increases in either penalty, and moves faster for increases in the prior substantiation penalty Δ_2 than for increases in the false claims penalty Δ_1. In both cases, the more responsive reaction to increases in the substantiation penalty Δ_2 follows from the fact that the false claims penalty Δ_1 is discounted by the chance of having a high quality product, so that firms, especially those with a high γ, react more strongly to increases in the substantiation penalty Δ_2.

Putting these observations together, one can see that increases in either penalty shift more of the parameter space into regions in which strong separation is supported, at the expense of shrinking the region in which weak signaling is supported. Recall that I emphasize three shortcomings of the false claim ban; it is useful to outline them in the context of this figure before proving them analytically in the following subsections.

First, the false claims ban alone supports no strong signaling for "small but reasonable penalties" totalling the ill-gotten gains from a false speculative claim (i.e., $\Delta_1 + \Delta_2 = \pi_H - \pi_L$). This is evident from the position of $LC_S^1(1)$. $LC_S^1(1)$ intersects the $\gamma = 0$ axis at $\Delta_1 + \Delta_2 - (\pi_H - \pi_L)$, which equals 0 at these "small but reasonable" penalties and intersects the $\gamma = 1$ axis at Δ_2, which is 0 under a false claims ban alone. Thus, no strong signaling is possible for a positive learning cost. In contrast, at small but reasonable penalties totalling $\pi_H - \pi_L$, but including some positive prior substantiation penalty, there is in fact strong separation for some portion of the parameter space as $LC_S^1(1)$ slopes up to a positive intercept at $\gamma = 1$.

Second, for any particular false claims penalty alone, there is no strong signaling for optimistic enough firms, even if the learning cost is very low (i.e., there is no strong signaling for some region of low-t, high-γ parameters). This is evident from the fact that $LC_S^1(1)$ intersects the right axis at 0 under a false claims penalty alone. However, under a prior substantiation penalty this constraint has a positive intercept and strong signaling prevails for even the most optimistic firms as long as the learning cost is low enough.

Third, a false claims ban alone that ensures strong signaling for some particular γ, t will also always support strong signaling for some region including higher learning costs. That is, ratcheting up the false claims penalty to ensure strong signaling for some particular parameters (e.g., where it is desirable) may also induce strong signaling in regions with higher learning costs (e.g., where it is not desirable). This is clear from the fact that $LC_S^1(1)$ slopes down to 0 at $\gamma = 0$. In fact, this is true also of the prior substantiation penalty used in isolation, since in that case $LC_S^1(1)$ slopes up to Δ_2 at $\gamma = 0$. However, when the two penalties are used together, in particular with $\Delta_1 = \pi_H - \pi_L$, the $LC_S^1(1)$ constraint can be made flat (with its position determined by the level of Δ_2), ruling out the inducement of learning for regions with higher learning costs.

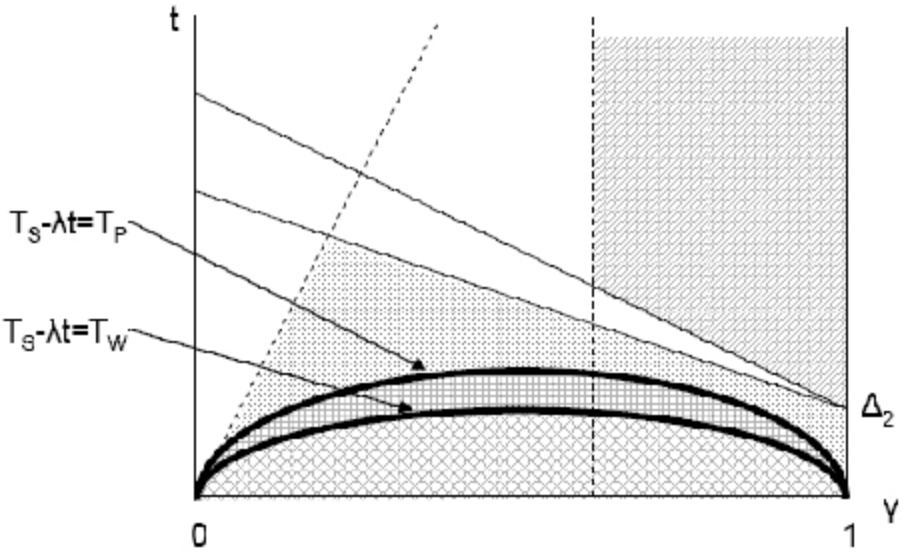
And, finally, I also emphasize one significant shortcoming of the substantiation penalty: it undermines weak signaling through its direct penalty on all speculative claims. As described above,

the region of weak signaling defined by $IC_W(1)$ and $LC_W(1)$ shrinks as penalties are increased, with both constraints moving more quickly to reduce the weak signaling region when the substantiation penalty is increased than when the false claims penalty is increased.

To see the relationship of this intuition and these features of the two policies to the implementation of socially optimal policies, it is necessary to further develop this graphical representation to include a representation of the socially optimal information revelation. This in turn requires putting more structure on the model to derive socially optimal information revelation regimes in this parameter space.

Corts (2011) develops a model in which perceived quality improvements increase the intercept of a linear demand curve. Specifically, the demand curve is $P = a_i - bQ$, where a_i is determined by the perceived quality: $a_L = a$; $a_H = a + \delta$; $a_1 = a + \gamma\delta$; and so on. In this specific model I can calculate the total social surplus under the various regimes–strong separation, weak separation, and pooling. This yields an expression for the boundary between the regions in which strong signaling is socially optimal and the regions in which weak signaling in socially optimal. Specifically, strong separation is preferred to weak separation when $t < \frac{3}{8}\delta^2\gamma(1-\gamma)$, and strong separation is preferred to pooling if $t < \frac{3}{8}\delta^2\gamma(1-\lambda\gamma)$. The shapes of these regions can be portrayed as in Figure 2; this is drawn for arbitrary penalties, so these boundaries may or may not lie wholly below $LC_S^1(1)$ as pictured. The area under the lower curved line is the region in which strong separation is socially optimal; above it, weak separation is socially optimal. Note that, quite intuitively, strong separation is optimal when learning costs are low and when there is relatively more to be gained through learning (γ is not too close to 0 or 1, so that knowledge of firm type alone is not itself extremely informative). In the area beween these curves, weak signalling is optimal, but strong signalling would be preferred to pooling if weak signalling were not feasible.

Figure 2



Combining this figure with the graphical intuition developed earlier gives a sense of the strengths and weaknesses of the two kinds of policies, which the subsequent sections make more concrete in analytical propositions. Throughout this discussion it is important to keep in mind that the differences in these policies or the advantages of having both policies do not arise when one is unconstrained in choosing penalties that are designed to apply to a single set of parameter values. All of the interest in the combination of the two types of policies is driven by consideration of possible constraints in penalties or the need to have a single set of penalties apply over a range of the parameter space. The following subsections show that in such circumstances use of both false claims and prior substantiation penalties is valuable in achieving socially optimal information revelation. As in the construction of the figures above, I assume throughout the remainder of this section that $\Delta_1 + \Delta_2 \geq \pi_H - \pi_L$, the necessary conditions for strong and weak signaling.

4.4 Supporting separation with small but reasonable penalties

One difference between the false claims ban and the prior substantiation penalty arises when one considers a regulator constrained to "small but reasonable" penalties. In particular, I consider penalties equal to exactly the "unfair profit" gained by a firm that falsely claims to be high quality under the consumer inferences associated with strong separation (i.e., $\pi_H - \pi_L$). By "reasonable" I do not mean to imply that this is a penalty that constitutes a good policy by some measure. I mean only that it is not a pathological case of arbitrarily high or low penalties, but is a penalty that is related in a reasonable way to the model's parameters. The penalty is also reasonable in the sense that it does not entail large punitive damages but essentially taxes away the profit gains achieved through the false quality claims. To the extent that the firm's profit increase represents a pure transfer of consumer surplus, this penalty is also a good proxy for the harm suffered by consumers, which could be relevant under liability rules that might constrain the size of the penalty. The penalty is also small in the sense that it is the minimal penalty that satisfies $IC_S(L)$, the informed, low-quality firm's incentive constraint. It therefore permits strong separation with respect to that constraint. The content of the following proposition lies in showing that a penalty that just satisfies this incentive constraint strictly fails to satisfy the strong separation learning constraint LC_S^1 for all positive learning costs. The subsequent proposition demonstrates that this is not the case for the prior substantiation penalty.

Proposition 4 *Assume the learning cost t is strictly positive. Then for "small but reasonable" penalties ($\Delta_1 = \pi_H - \pi_L$) a false claims ban alone achieves no strong separation regardless of the other parameter values.*

The learning constraint LC_S^1 can be rearranged in the FCB-only case ($\Delta_2 = 0$) to yield $\pi_H - (1 - \gamma)(\pi_H - \pi_L) - t \geq \pi_H - (1 - \gamma)\Delta_1$. The left side of this inequality is the profits from learning, while the right side is the profits from making a speculative and potentially false claim. This can be read as follows: by claiming to have high quality, the firm will earn profits π_H in some states regardless of which path it follows; however, in each case it suffers a different downside risk. If the firm learns, the downside is that the firm spends t and $1 - \gamma$ of the time the firm gets only π_L. The downside of making the speculative claim instead is that $1 - \gamma$ of the time the firm is exposed to the FCB penalty Δ_1. As this inequality makes clear, at these small but reasonable penalties ($\Delta_1 = \pi_H - \pi_L$) the firm is indifferent between the risk of discovering it is a low type and the risk of paying the FCB penalty.

However, the firm strictly prefers not to learn since learning is costly. With the failure of the LC_S^1 constraint, there can be no strong separation, which proves the proposition.

Proposition 5 *For "small but reasonable" penalties ($\Delta_2 = \pi_H - \pi_L$), mandatory prior substantiation alone achieves strong separation if learning is not cost-prohibitive (i.e., if $t \leq \gamma(\pi_H - \pi_L)$ so that $LC_S^2(1)$ holds), regardless of the other parameters.*

Rearranging LC_S^1 in this case yields $\pi_H - (1-\gamma)(\pi_H - \pi_L) - t \geq \pi_H - \Delta_2$. Here, the left side is unchanged, with the firm accounting for the learning cost it will bear and the $1-\gamma$ chance that it will discover it does not have a high-quality product. However, the right side is different, with the firm realizing that it will pay the undiscounted MPS penalty Δ_2 if it does not learn but signals anyway, regardless of the truth of its claim. When $\Delta_2 = \pi_H - \pi_L$, this inequality is satisfied for small t. The undiscounted penalty for unsubstantiated claims looms large in comparison to the learning cost and the chance that the firm will learn it has a low quality product.

The inability of the false claims ban alone to achieve strong separation with these constrained penalties is a significant weakness of this type of penalty; the ability of the prior substantiation penalty to achieve strong separation with the same constraints demonstrates a general point about its ability to induce more learning and signaling with any given penalty. In the case of these particular constrained penalties, the contrast is very stark and the prior substantiation penalty is clearly preferred when low learning costs imply that strong signaling is socially optimal for the relevant region of the parameter space.

4.5 Supporting separation among very optimistic firms

In this subsection I continue to pursue the contrast between what the two types of penalties can achieve with a constrained penalty, although here the penalty is some arbitrary fixed penalty rather than the specific small but reasonable penalty described above. The following pair of propositions shows that the prior substantiation penalty is more effective in achieving strong separation for low-t, high-γ portions of the parameter space, which the graphical analysis above showed is likely to involve socially optimal strong separation.

Proposition 6 *No matter how high the penalty or how low the strictly positive learning cost, a false claims ban alone supports no strong separation if the firm is very confident that it has a high quality product. (For any Δ_1 and any c, there exists some $\widehat{\gamma} < 1$ such that no strong separation equilibrium exists for $\gamma > \widehat{\gamma}$.)*

The learning constraint LC_S^1 can be rearranged to yield $t \leq (1-\gamma)[\Delta_1 - (\pi_H - \pi_L)]$, both sides of which are continuous in γ. In the limit as $\gamma \to 1$, this becomes $t \leq 0$, which is clearly strictly false. It is therefore strictly false for γ near enough 1. (Alternatively, the learning constraint is strictly false for $\gamma > \widehat{\gamma} = \max\{0, 1 - t/[\Delta_1 - (\pi_H - \pi_L)]\}$.) Without satisfaction of the learning constraint, there is no strong separation, which proves the proposition.

The intuition for this result is very simple. The incentive to learn in this model derives from the opportunity to avoid making a false quality claim that exposes the firm to FCB penalties. However, a firm that is very confident about its product quality prior to learning (that is, has a very high γ) discounts the possibility of being subject to the penalty (in the limit, it ignores the penalty). This

leaves the firm no incentive to learn, which invalidates the consumers' inference under strong signaling that high quality claims are made only by firms who have learned they have a high quality product with certainty. The firm does not similarly discount the prior substantiation penalty, rendering it much more powerful in inducing learning among optimistic firms.

Proposition 7 *There exists a range of learning costs $t < \Delta_2$ for which a given mandatory substantiation penalty alone supports strong separation even for a firm with arbitrarily high confidence in its product's high quality (specifically, for any $\gamma \in (\widehat{\gamma}, 1)$, where $\widehat{\gamma} = \max\{\frac{c}{\pi_H - \pi_L}, 1 - \frac{\Delta_2 - c}{\pi_H - \pi_L}\}$).*

Intuitively, the FCB penalty fails to support strong separation because the penalty is discounted by the small probability that the firm making a speculative claim will be exposed to that penalty. In the case of the unsubstantiated claim penalty, there is no such discounting. By making a speculative claim, the firm is exposing itself to the MPS penalty regardless of its level of confidence in its product quality. The firm compares learning, followed by truthfully signaling its high quality (incurring only the learning costs), with signaling without learning (exposing itself to the MPS penalty). If the learning cost is less than the MPS penalty, then learning is preferable even if the firm is certain to have a high quality product. The threshold γ given in the proposition is determined by the binding learning constraint.

Since it is clear from the analysis of the socially optimal information revelation regime that some low-t, high-γ portions of parameter space will have socially optimal strong separation, the difficulty of the false claims ban alone in achieving strong signaling in this range is a significant shortcoming. Under these kinds of constrained penalties the availability of the prior substantiation penalty clearly improves the regulator's ability to implement the socially optimal information revelation regime.

4.6 Supporting strong signaling evenly across different firms

Either a false claims ban alone or a prior substantiation requirement alone suffers an additional difficulty when the regulator must establish a single policy that applies across a broad range of parameters. In particular, the penalty that induces socially beneficial learning at one level of γ may induce socially wasteful learning at other levels of γ. This can be seen by noting that strong signaling may prevail at higher t as γ varies, whereas higher t always makes strong signaling less socially beneficial.

Proposition 8 *Assume that $LC_S^2(1)$ strictly holds. Then a false claims penalty alone or a prior substantiation penalty alone that is just sufficient to induce strong separation for a particular γ, t (i.e., such that $LC_S^1(1)$ binds) also induces strong separation for some higher t.*

Note that when $LC_S^2(1)$ holds it is $LC_S^1(1)$ that is key to determining whether strong signaling can occur. In the case of a false claims penalty only this reduces to $t \leq (1-\gamma)[\Delta_1 - (\pi_H - \pi_L)]$. Beginning from a γ, t where this holds with equality, a reduction in γ relaxes the constraint and allows it to be satisfied for a larger t. In the case of a prior substantiation penalty only this reduces to $t \leq \Delta_2 - (1-\gamma)[\pi_H - \pi_L]$. Here, increases in γ relax the constraint and allow it to be satisfied at higher t.

This result is of interest because higher learning costs t make it less likely that strong separation will be socially optimal. Thus, the false claims ban that induces socially optimal learning at one level

of γ risks inducing socially wasteful learning at other levels of γ. This problem can be eliminated by the combined use of the policies.

Proposition 9 *Assume that $LC_S^2(1)$ strictly holds. Then a false claims penalty of $\Delta_1 = \pi_H - \pi_L$ combined with some positive prior substantiation penalty $\Delta_2 > 0$ that just induces strong separation for a particular γ, t does not induce strong separation for higher t regardless of γ.*

With this false claims penalty the relevant constraint $LC_S^1(1)$ becomes simply $t \leq \Delta_2$, which does not depend on γ and is obviously never satisfied for a higher t as other parameters change.

4.7 Supporting weak signaling

While MPS does allow strong signaling over increased ranges of parameters, it has a weakness of its own. Because it penalizes failures to learn, it makes the speculative claims associated with weak signaling unattractive; increases in the substantiation penalty therefore reduce the range of parameters for which weak signaling is possible. Increasing the false claims penalty also discourages weak signaling; however, increasing the prior substantiation penalty reduces the range of parameters for which weak signaling can be supported more quickly.

Proposition 10 *Increases in either the false claims penalty or the prior substantiation penalty increase the range of parameters for which strong signaling is possible and reduce the range of parameters for which weak signaling is possible. Increasing the substantiation penalty shrinks this region more quickly than increasing the false claims penalty.*

The constraints determining the strong signaling boundaries are LC_S^1 and LC_S^2. LC_S^2 does not depend on the penalties. LC_S^1 can be rewritten $t \leq (1-\gamma)[\Delta_1 + \frac{\Delta_2}{1-\gamma} - (\pi_H - \pi_L)]$. Clearly increases in either Δ_1 or Δ_2 tighten this constraint and reduce the range of other parameters for which it is satisfied. It is clear that the range of parameters for which weak signaling is supported shrinks more rapidly for increases in Δ_2 because its effect is multiplied by its division by $1-\gamma < 1$ (while the effect of increases in Δ_1 is not). The constraints determining the region of weak signaling are $IC_W(1)$ and $LC_W(1)$ (note that $IC_W(0)$ holds by assumption throughout this section and does not reference parameters other than π_1 and π_L). $IC_W(1)$ can be written $\pi_1 - \pi_L \geq (1-\gamma)\Delta_1 + \Delta_2$. This constraint is clearly tightened by increases in either Δ_1 or Δ_2, and is tightened more quickly by increases in Δ_2 since Δ_1 is discounted by its multiplication by $1-\gamma$ (whereas Δ_2 is not). $LC_W(1)$ can be written $t \geq (1-\gamma)[\Delta_1 + \frac{\Delta_2}{1-\gamma} - (\pi_1 - \pi_L)]$, for which the argument follows exactly as for LC_S^2 above.

5 Conclusion

This paper makes it clear that a false claims penalty and a prior substantiation penalty are not equivalent. Prior substantiation penalties are not discounted by optimistic firms, and they therefore are more powerful in inducing learning and enabling strong signaling either with modest penalties or among very optimistic firms. However, precisely because they are more powerful in inducing learning, they inhibit weak signaling, which relies on firms not learning product quality but remaining informed of their type only. In general, if the regulator must work with constrained penalties or must devise a penalty to apply across a range of possible parameter values, the ability to use both of these types of penalties is valuable.

A false claims ban alone suffers three shortcomings. First, it supports no learning if penalties are limited to the amount of unfair profits earned through false claims. Second, even with the penalty increased beyond this level, there are always levels of optimism such that a firm will not learn because it discounts the penalty for false claims so completely. Third, further increasing the penalties to ensure learning and strong signaling by more optimistic firms can also induce socially wasteful learning by less optimistic firms. Using mandatory substantiation in conjunction with a false claims ban provides more flexibility and allows the policy-maker to mitigate the weaknesses of the false claims ban. However, it comes at the expense of a reduction in weak signaling. In general, the two policies together outperform either policy in isolation when the regulator must create a policy that applies across a broad set of parameters.

It is worth discussing here how one should interpret in practical applications the model's distinction between product quality and firm type. In the model, the key distinction is that product quality is an objective fact about the product that is testable and verifiable to the policy enforcer, while the firm's type is a subjective assessment, which is not testable and not verifiable to the policy enforcer, that the firm has about its likely product quality. It is this distinction in testability and verifiability that is critical, not the distinction between certainty (H or L) and uncertainty (a probability γ in a range from 0 to 1) that is also present in the model. Thus, a product quality claim could absolutely involve uncertainty–for example, a claim that an herbicide for lawns will kill 80% of weeds on average. While this claim involves uncertainty, it is a testable claim, and it is a claim that could be verified by or to the policy enforcer upon investigation. It could therefore be a high quality claim in the context of the model. A firm's type in this context would be its assessment of the likelihood that its kill rate is in fact at least 80%. Clearly, performing careful tests could allow the firm to learn whether the claim was true, allowing the firm to make that claim with confidence. Or, the firm could make such a claim on a speculative basis, not sure whether testing would bear out this assertion or not. Thus, the distinction between certainty and uncertainty is not central to the application of the model; what is central is the distinction between what is testable and untestable or verifiable and unverifiable to the regulator.

This example does, however, point to a feature of quality claims that is left outside of the present model, in which the standards of high and low quality are exogenous. In practice, the firm has some flexibility in defining the quality claim to be made, and this affects both how the consumers value the high quality product and the exposure of the firm to penalties. To continue the herbicide example, the firm could instead make the claim that its product kills 60% of weeds. While this would probably reduce consumers' valuation of the product relative to the 80% claim (depending on the inferences they make based on the relative credibility of the two claims), the claim is also more likely to be true, which lowers the expected false claims penalty for a firm that was unsure of its product quality. Thus, the endogenous determination of the strength of the product quality claim could be an interesting feature to explore further in a similar model.

This analysis leaves aside many other important considerations by abstracting from the enforcement regime, incentives to innovate, the process of consumer expectations formation, and many other considerations that could be considered in future research. For example, the enforcement of these penalties is exogenous here; the policy-maker simply sets expected penalties for infringement of each prohibition. In fact, levels of enforcement and the endogenously determined expected penalties could hinge critically on the nature of the policy employed. If, for example, the substantiation requirement

lowered government investigation and trial costs by shifting the burden of proof toward the firm compared to a scenario with only a false claims ban, then more stringent enforcement could be expected, meaning that a given nominal penalty might translate into a much higher expected penalty under that policy. Consideration of incentives to innovate would generally raise the social benefit to learning and dissemination of information, making policies that encourage learning and strong separation even more desirable. Similarly, one might doubt that weak separation equilibria are as desirable or as likely in real life as in the model, as they require sophisticated consumers to properly discount claims of product quality to sustain the equilibrium. In the absence of weak signaling as an alternative, learning and strong signaling become even more socially desirable, increasing the importance of the substantiation requirement.

6 References

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

Corts, K., 2011, "Inducing Socially Optimal Information Revelation by Privately but Imperfectly Informed Firms," working paper.

Daughety, A. and J. Reinganum, 2008, "Products Liability, Signaling and Disclosure," *Journal of Institutional and Theoretical Economics* 164 (1), 106-126.

Farrell, J., 1986, "Voluntary Disclosure: Robustness of the Unraveling Result, and Comments on Its Importance," in R. Grieson (Editor), *Antitrust and Regulation* (Lexington, MA: Lexington Books), 91-103.

Grossman, S., 1981, "The Informational Role of Warranties and Private Disclosure about Product Quality," *Journal of Law and Economics* 24 (3), 461-483.

Matthews, S. and A. Postlewaite, 1985, "Quality Testing and Disclosure," *Rand Journal of Economics* 16 (3), 328-340.

Polinsky, M. and S. Shavell, 2006, "Mandatory Versus Voluntary Disclosure of Product Risks," NBER Working Paper 12776.

Viscusi, K., 1978, "A Note on Lemons Markets with Quality Certification," *Bell Journal of Economics* 9 (1), 277-279.