

# Approval regulation and endogenous consumer confidence: Theory and analogies to licensing, safety, and financial regulation

Daniel Carpenter

*Department of Government, Harvard University, Cambridge, MA, USA*

Justin Grimmer

*Department of Political Science, Stanford University, Stanford, CA, USA*

Eric Lomazoff

*Department of Government, Harvard University, Cambridge, MA, USA*

## Abstract

Safety regulation – in the form of pre-market approval, licensure, screening, and product entry limitations – governs numerous market realms, including consumer finance. In this article, we ask whether the effects of safety regulation go beyond safety and affect consumers' beliefs about the distribution of products they can use. We model “approval regulation,” where a government regulator must approve the market entry of a product based upon observable, unbiased, and non-anticipable experiments. We show that even if regulator and firm disagree about only quality standards, the disagreement induces the firm to provide more information about its product than it would in the absence of regulation. Put differently, purely first-order disagreements in regulation generate second-order consequences (more certainty about product quality). These second-order consequences of regulation are sufficient to generate first-order effects among end-users (more consumption of superior products), even when users are risk-neutral. In other words, even if approval regulation produces little or no improvement in safety or quality, it still aggregates information useful to “downstream” product users; these users will exhibit higher consumption and will more readily switch to superior products. In contrast with libertarian analyses of entry regulation and licensure, the model predicts that entry restrictions may be associated with greater product or service utilization (consumption) as well as with greater price sensitivity among consumers. Because contemporary cost–benefit analyses ignore these second-order effects, they are unlikely to capture the possible confidence effects of approval regulation.

**Keywords:** approval regulation, safety regulation, pharmaceuticals, licensing, financial regulation.

## 1. Introduction

In the name of safety and quality, governments worldwide constrain the development, manufacture, marketing, and utilization of products and services (Warren 2007, 2008). The regimes of governance vary widely, from licensure of professional services to review

Correspondence: Daniel Carpenter, Department of Government, Harvard University, 1737 Cambridge St., Cambridge, Massachusetts, USA 02138. Email: [dcarpenter@gov.harvard.edu](mailto:dcarpenter@gov.harvard.edu)

Accepted for publication 26 September 2010.

of permit applications for construction or infrastructure renewal (such as dam or wetlands construction) and to the extensive pre-approval requirements governing the introduction of drugs, medical devices, and food additives. The ubiquity of these institutions is especially noteworthy. Elaborate regimes of professional licensure characterize not only modern-day economies worldwide, but also the historical past of economic development in North America, Europe, and Asia, among other regions. Mechanisms of pre-market approval for health products are observed not only in the US, the European Union, and Japan, but also with increasing frequency (and rigor) in China, India, Brazil, Saudi Arabia, and other less economically developed nations. In contrast to regimes of price regulation, moreover, institutions of quality and safety regulation have received much less analytic and scholarly attention.

Institutions of product and service regulation have recently attracted greater scholarly and policy attention, not least because scholars in law and economics (such as Elizabeth Warren [2007] and Joseph Stiglitz [2008]) have proposed that a form of “product safety” regulation ought to be applied to financial markets. Among the essential ideas of these proposals are that a government regulator should prevent or limit the introduction of some products due to safety issues that their eventual consumers may not fully perceive. In the wake of the financial crisis of 2008 and 2009, these ideas and related notions and proposals have received considerable attention. With the creation of new national consumer finance agencies – including the Financial Consumer Agency of Canada (created in 2001) and the Consumer Financial Protection Bureau in the US (created in 2010) – these ideas have become institutionally tangible.<sup>1</sup>

The idea of “product safety” regulation for finance and other realms of economic industry is plausible and enticing for some observers. Yet the idea is also untried at the scope and scale for which its proponents are suggesting it. For this reason it deserves further investigation and scrutiny. How does product safety regulation operate? What powers does a government regulator need to have in order to make safety and quality regulation function well? What properties – what particular government powers and what particular sanctions and direct or indirect incentives – does quality and safety regulation entail? Once these powers and incentives are identified, we can then ask which properties of “consumer safety regulation” are appropriable to other realms and which are not. Perhaps most important, when the operating principles (or equilibrium dynamics) of safety regulation are identified, the full complement of regulatory effects can be studied. Do consumer safety regimes influence not only safety, but also other welfare-related variables?

The central claim of this article is that safety regulation can induce a form of consumer “confidence.” Some forms of safety regulation may influence end-users’ beliefs about products they face in the market. And in so doing, safety regulation with some pre-market review of products can reduce consumer uncertainty beyond that which would prevail in an unregulated market. Although forms of safety and quality regulation can be thought to reduce the likelihood of product hazards or to militate against equilibrium fraud, these regulatory regimes do as much or more of their work by providing a certain kind of information about products. Both the quantity and the quality of this information, we contend here, are far greater than would be provided in an unregulated marketplace. Accordingly, in this article we draw upon existing formal models and empirical and historical studies of safety regulation to elaborate the dynamics of pre-market approval regimes. In this model the crucial powers of an approval regulator

rest in two properties: (i) the ability to compel firms to experiment with new products before their introduction into the market and (ii) the ability to veto the introduction of new products or to withdraw such products after their introduction. Property (ii) is a form of regulation common to many forms of “entry regulation.” What makes this approval regulation regime different is property (i), the robust connection between approval and the experimental generation of information used in the entry decision. This connection is observed in many licensure regimes (such that applicants must pass a series of courses or examinations or perhaps a probationary period; Law & Kim 2005), in many permitting institutions (such that the proposed project must pass an environmental impact test or a cost–benefit criterion), and in pre-market approval institutions (such as pharmaceutical and medical device regulation worldwide).

In particular, we present a mathematical connection between a stylized regulatory process and a stylized market. The regulation, the connection, and the resulting market are all modeled, albeit with considerable simplification. In the regulatory process, a firm conducts experiments with its product to satisfy the higher quality/safety standards of a regulator, and the regulator sets an approval policy that skims off a subset of the products developed. We show that, even though the firm and the regulator disagree about only quality, their interaction produces information that reduces the uncertainty of both players. To simulate a market of products, we suppose that this regulation is repeated regularly and many times, creating a single-mode, approximately Normal distribution of product quality that is faced by consumers or users of the product. In the resulting stylized market, a large number of “users” consume the approved products from the distribution induced by approval regulation. In utilizing these products, consumers learn about the products’ quality. In the resulting user equilibrium, *risk-neutral consumers are more likely to consume products (enter the market in the first place), and more likely to engage in optimal product switches once they do, as a direct result of regulation.* First-order disagreements in regulation have second-order informational consequences, and these second-order consequences have first-order implications for consumers whose a priori utilities are unaffected by second-order considerations. If such effects exist, the second-order consequences of approval regulation regimes are unlikely to be captured by modern methods of policy analysis (Viscusi 1993; Boardman *et al.* 1996), which largely focus on the safety effects of regulation alone.

We acknowledge several limitations of our effort, including simplifying assumptions that we discuss in the Implications and conclusions. Yet we pause here to clarify that our models do not serve as a pure rationale for the creation of an “FDA” or “phased-experiment” system of trials for financial products or any other products. This is, not least, because we do not model all of the relevant costs of approval regulation regimes here. The really hard work of applied regulatory theory depends upon the careful translation of theoretical and normative considerations into policy proposals. In part, we intend this work to be undertaken by other scholars. The model elaborated here may be informative for such efforts, as it points to particular features – the type of information generated and the incentives engendered by regulatory veto authority – that a policy proposal might entail.

On the positive side, the models here suggest that safety and quality regulation of the form that is observed in health and food products (and less strictly for licensure) may have broad benefits having less to do with safety and more to do with *ex post* product utilization. Our models offer results that cohere with recent analyses of product quality

and safety regulation and their *ex post* effects in regulated markets (Law 2003; Law & Kim 2005; Law & Marks 2009). They also provide a formalization and a mechanism for the argument that Warren entertains in her essay (2007, 2008), namely that, as Alexander Hamilton (2001 [1769–1804]) wrote two centuries ago in his essays on government and finance, the establishment of confidence should remain a critical focus of government policy.

## 2. A model of approval regulation

Product safety regulation borrows its essential characteristic from political institutions: a *veto*, in this case a veto of the regulator over firms' market entry (or market continuation). We begin by modeling the centrality of this veto in the regulatory process, and in doing so we offer a simple characterization of many features of licensing and product safety regulation in the modern world. What distinguishes our model from usual models of regulation is (i) the explicit modeling of entry vetoes and (ii) the centrality of experiment to the reduction of two-sided uncertainty.<sup>2</sup>

### 2.1. Informational environment and players: "Nobody knows the truth"

There are two players in our model: a firm ("the Firm") and a regulator ("the Regulator"). Both players are imperfectly informed about a parameter  $x$  of a product, which may be thought of as the product's "quality" or "safety." We assume that  $x$  follows a Beta distribution, or, formally, ( $x \sim \text{Beta}(\theta, n)$ ), where  $\theta, n$  are positive integers and  $1 < \theta < n$ . The first parameter of the distribution,  $\theta$ , is the Firm's "type;" that is, the estimated quality of its product.<sup>3</sup> A higher type connotes a better expected product, while a lower type connotes a lower expectation of product quality. The Beta distribution of priors offers a natural interpretation of a set of  $n$  Bernoulli trials, of which  $\theta$  resulted in success and  $n - \theta$  in failure.<sup>4</sup> The distribution is also flexible enough to accommodate a wide variety of "shapes" of the density function, as determined by  $\theta$  and  $n$ . Given  $\theta$  and  $n$ , the Beta distribution implies a prior mean  $\theta/n$  and prior variance  $\theta(n - \theta)/(n^2(n + 1))$ . The uncertainty over  $x$  can be resolved partially through observable experiments. Because the parameters affecting  $E[x]$  change with experimentation, we distinguish between the Firm's "initial" type  $\theta$  and the numerical value of the first parameter of the Beta distribution by letting  $m = \bar{\theta}$ .

The product's type is private information known only to the Firm. We characterize product type as taking one of two possible values – "high" and "low" – hence  $\theta \in \{\underline{\theta}, \bar{\theta}\}$  where  $\underline{\theta} = \bar{\theta} - 1$ . We will refer to  $\underline{\theta}$  and  $\bar{\theta}$  as the low and high types, respectively. The second parameter,  $n$ , is common knowledge. Let  $p \in (0, 1)$  represent the Regulator's prior belief of the high type ( $\theta = \bar{\theta}$ ).

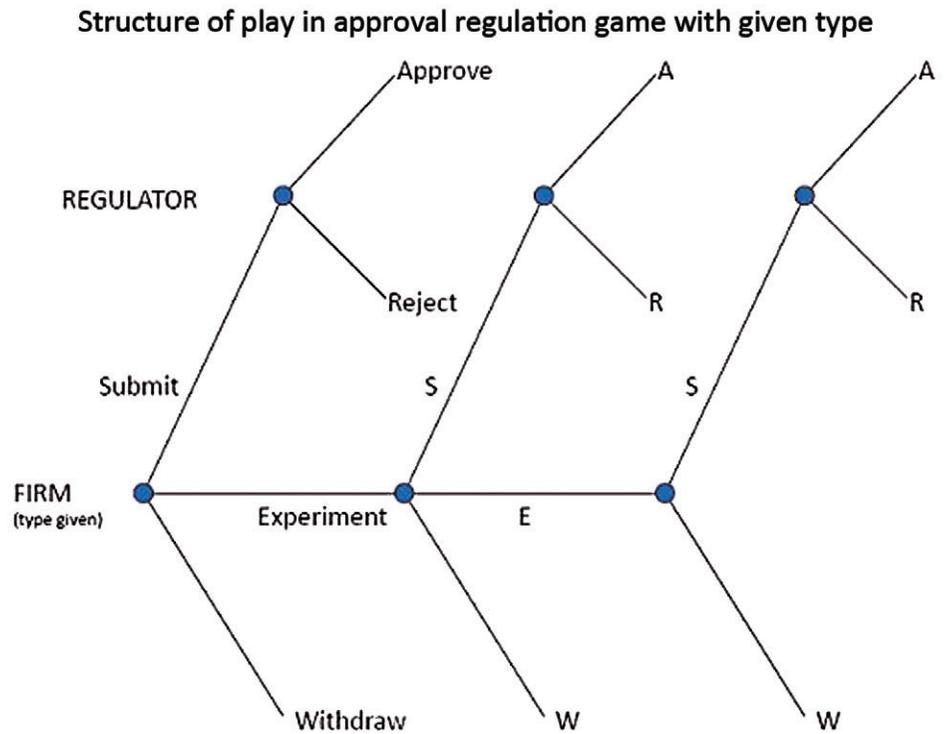
A crucial feature of the approval regulation model, then, is that "nobody knows the truth" about the product's quality in the sense that  $x$  is never known or estimated with full certainty. The Firm and Regulator possess only estimates of the truth, and information is asymmetric only to the extent that the Firm's estimate (its "prior") is better than the Regulator's. The idea that even firms are somewhat uncertain as to the quality and safety of their products is a more accurate representation of regulation (and of other forms of social, political, and economic reality). For this reason, it is, we believe, much more accurate and compelling than the simplistic representation of "incomplete information" that has dominated the regulatory economics literature (Laffont & Tirole 1994).

**2.2. Sequence of play: Development first, then regulation**

The approval regulation game has as many as four periods ( $t = 0, 1, 2, 3$ ) denoted by subscripts. These periods comprise a *development* phase with up to three periods and possibly a *regulatory* phase of one period. The phases are distinguished by the mover: only the Firm moves in the development phase and only the Regulator moves in the regulatory phase, if there is one. The moves for a given type of Firm and Regulator are depicted in Figure 1.

The approval regulation game begins in the development phase, in which the Firm chooses an action  $f_t \in \{S_F, W_F, E_F\}$  at  $t = 1, 2$ .  $S_F$  denotes the firm’s submission for approval, which ends the development phase and commences the regulatory phase in the next period.  $W_F$  denotes the Firm’s withdrawal of a product from consideration, ending the game. Finally,  $E_F$  denotes a Firm-funded experiment to gather more data. An experiment is a single Bernoulli trial, which produces a publicly observable result  $e_t \in \{0, 1\}$  corresponding to failure or success, respectively. An experiment continues the development phase. The Firm cannot experiment past the second period and thus  $f_3 \in \{S_F, W_F\}$ . The final period thus represents something of a “fish or cut bait” choice for the firm (the far-right termini of Fig. 1). For convenience, we let  $e_0 = 0$  and use the term “experimental history” to refer to the set of experiments performed (up to a maximum of two).

At the beginning of the regulatory phase, the Regulator knows the Firm’s actions and experimental results, but not does not know either  $x$  or  $\theta$ . Based on this, she or he makes a review decision  $r \in \{A, R\}$ , where  $A$  and  $R$  denote acceptance and rejection of the Firm’s submission, respectively.



**Figure 1** Approval regulation game tree for given type.

### 2.3. Revelation of information: Calculation of posterior beliefs

Without experimentation, the model reduces to a simple signaling game with the Firm as the sender. With experimentation, both players update their expectations of  $x$ . The assumption of Beta-distributed priors makes the calculation of posterior beliefs very simple when both parameters of the distribution are known. For example, beginning with a prior of  $\beta(\theta, n)$ , two experiments producing  $e_1 + e_2$  successes generate a posterior distribution of  $\beta(\theta + e_1 + e_2, n + 2)$ . Accordingly,  $E[x|e_1, e_2] = \frac{\theta + e_1 + e_2}{n + 2}$  and  $Var[x|e_1, e_2] = \frac{(\theta + e_1 + e_2)(n + 2 - \theta - e_1 - e_2)}{(n + 2)^2(n + 3)}$ .

### 2.4. Utilities

The Firm receives  $x$  if the product is approved and zero for rejection or withdrawal. Each experiment costs  $c_e$ , and a submission costs  $c_s$ , where:

$$c_e \in \left( 0, \frac{m(m-1)}{(n+1)(n+2)(n+m-1)} \right] \quad (1)$$

$$c_s \in \left( 0, \frac{m}{n+2} \right]. \quad (2)$$

These assumptions ensure that the low type is not prevented from experimentation or submission by exogenous costs alone (though it may choose not to do so in equilibrium). They also substantially simplify the analysis by eliminating some trivial equilibria.

The Regulator receives  $x - k$  for an approved product and zero otherwise. The parameter  $k$  is therefore the divergence between the preferences of the Regulator and the Firm. It can be imagined to represent a “certainty equivalent” that a risk-averse or uncertainty-averse society, or a society fearing a product safety disaster, would demand from the Firm in order for its product to be marketed. We assume that  $k$  satisfies:

$$k \in \left( \frac{m}{n}, \frac{m+1}{n+2} \right). \quad (3)$$

These bounds guarantee that some experimentation is necessary to generate a product satisfactory to the regulator but that the two players can disagree over the desirability of marginal products.

We characterize Perfect Bayesian Equilibria (PBE) that satisfy a minor refinement (the online Appendix contains the formal definition of the equilibria, proofs of existence in Propositions 1 and 2, and other demonstrations of interest).

## 3. Experimentation and information provision in the approval regulation game

The main model offers two types of equilibrium – an early submission equilibrium (ESE) and a late submission equilibrium (LSE) – depending on how players react to a successful first-period experiment. For both equilibria, the Regulator’s decision problem occurs

when the Firm submits. Clearly, the Regulator accepts a submission if its expected quality is greater than  $k$ , rejects it if it is less, and is indifferent to it (and so may mix) otherwise.

An ESE is an equilibrium in which The Firm submits “early” (i.e. in period 2) with positive probability. This can occur only if the Firm’s first experiment is successful. There are two possibilities induced by the history with initial experimental success. These depend on the Firm’s quality threshold,  $k$ , relative to an expected quality parameter,  $\bar{x}$ , which is the expected period 2 quality conditional upon both types experimenting successfully in period 1.

If  $k \leq \bar{x}$ , then the prior distribution of product technologies is favorable in the eyes of the Regulator. The Regulator is hence willing, at least in expectation, to accept the set of all successful first-period experimenters (including a product that would later be revealed to be of low type). Because additional experiments are costly and do not improve product quality in expectation, the Firm submits whether its product is of high or low type, and all product submissions are accepted with certainty.

If  $k > \bar{x}$ , then the Regulator would not accept the set of all first-period successes. However, she or he still wishes to accept the high type (by (3)), and can choose a rejection strategy to deter the low type from submitting with certainty. Early submission then requires that an initially successful high type prefers submission to continued experimentation, which results from satisfaction of a particular “early submission condition” (see the Appendix).

An LSE is an equilibrium in which there are no early submissions; that is, none at  $t = 2$ . This may occur either because the expected quality of successful experimenters at  $t = 2$  does not warrant acceptance (i.e.  $k > \bar{x}$ ) or because an initially successful experimenter would prefer to gather more information (i.e. the early submission condition is violated). The existence of the LSE is assured when the ESE does not exist, and furthermore, it does not depend on the values of  $c_s$  or  $c_e$ . The LSE requires only that  $k > \bar{x}$ , while the ESE requires that either  $k \leq \bar{x}$  or the early submission condition hold. Thus when  $k \leq \bar{x}$ , the model uniquely predicts the ESE. And when  $k > \bar{x}$  and the early submission condition does not hold, the model uniquely predicts the LSE. If  $k > \bar{x}$  and the early submission condition holds, then both equilibria exist.

A striking feature of both the ESE and the LSE is that *the average quality of submitted products is usually equal to the Regulator’s reservation value ( $k$ )*, which makes the Regulator indifferent between rejection and acceptance.

### 3.1. Implications of approval regulation equilibria

A critical result from both ESE and LSE is that they result in the production of more information about the product than would occur if the Firm could enter the marketplace without a regulatory veto. This is true despite the fact that both players are risk-neutral and indeed uncertainty-neutral; neither Firm nor Regulator cares directly about information (the option value of experimentation is not informational but instead a “gambling” option value; i.e. contingent upon the possibility of superior results). The intuition here is that the Firm can enter the market only with the Regulator’s assent, and there is no way to gain the Regulator’s assent but to prove the product’s case. Because the Firm’s threshold for a desirable product is lower than the Regulator’s (by the scalar term  $k$ ), the Firm conducts more experiments with its product than it would if it were doing research and development (R&D) on its own. This result is summarized and demonstrated in Lemma 1.

*Lemma 1: Second-order stochastic dominance of regulated-induced product distribution.*

*Denote the unregulated product distribution by  $E_U[x]$  and the distribution produced in the ESE or LSE by  $E_R[x]$ . For all  $k > 0$ ,  $E_R[x]$  has second-order stochastic dominance over  $E_U[x]$ .*

Proof appears in the online Appendix.

The intuitive summary is as follows. The approval regulation model as such offers several useful lessons for mapping into the regulated industry domain. First, information has two dimensions of value (private information and common information). Second, because acceptance with certainty would induce all types to submit, the expected quality of submitted and accepted products will often be exactly the Regulator's reservation value of  $k$ . The equilibrium acceptance rate induces the low type to submit at a point of induced indifference between submitting and experimenting (or, in the final period, between submitting and withdrawing). The Regulator thereby benefits from its gatekeeping power, as  $k > m/n$ . Thus, the approval regulation process allows the Regulator to "skim" the best products from a population that is *ex ante* unacceptable.

#### 4. Cumulation of regulatory decisions: An *ex ante* product distribution for consumers

We now construct a thought experiment by imagining that the approval regulation game in the previous section is repeated many times. These repetitions are independent, such that we are *not* modeling a repeated game. Rather, we construct a distribution of products resulting from many independent plays of the approval regulation game.<sup>5</sup> The repetitions are independent and they also incorporate the possibility that, in each play of the game, a slightly perturbed initial distribution of quality faces the Firm and the Regulator.<sup>6</sup>

Repetition in this way represents the continuous and aggregate operation of regulation, which yields a general distribution of products. End-users of these products (consumers) are faced with the problem of making inferences about a single product drawn from the general distribution. In making inferences about any one product drawn from the distribution, however, consumers will take into account properties of the distribution itself, such as expected values and uncertainty. Our argument, in summary, is that approval regulation shapes this general distribution. The lemma in this section demonstrates how approval regulation would shape the aggregate population of opportunities faced by a consumer.<sup>7</sup>

The basic intuition is that if the approval regulation process is repeated many times over, a representative consumer would be able to treat the aggregate results of the regulatory process as a single-mode distribution of product safety or quality that is approximately normal (Billingsley 1999). The intuition here is that after the products are reviewed, much of the same information that was used by the Regulator to approve the product can also be exploited by the consumer when using it.<sup>8</sup> For example, the results of a clinical trial or experiment, which was required for pre-market approval of the product, can be put into the product's label or advertising. The results of a professional licensing process may convey additional information about the people who are eventually licensed. Accounting regulators may attach grades to licensed accountants (as in California), distinguishing those who have various levels of experience. In the educational process

that is required for legal or medical licensure, some licensees will score high grades and will obtain membership in the Order of the Coif (a law school honors society) or Alpha Omega Alpha (a medical school honors society). Downstream users of health services (patients) may learn that a health professional was “board-certified” in a particular area of practice, while downstream users of legal services (clients) may learn that their attorney is licensed to practice in multiple states, including those where passing the bar exam is quite difficult. The information can be coarse and yet still useful.

For the modeling exercise here, Lemma 2 means that a representative consumer can begin using products about which she or he is uncertain, and that the uncertainty can be represented by a single, continuously valued parameter. This result permits us to model the product utilization process in a very general and robust way, as a multi-armed bandit model with continuously resolved uncertainty.

*Lemma 2: Cumulation of the regulation-induced product distribution into a continuous mixture distribution.*

*Let the product distribution resulting from regulated firm entry be described by  $G_R \equiv G(E_R[x])$ , and let the distribution governed by unregulated entry be  $G_U \equiv G(E_U[x])$ . Consider a countable set of product distributions, each independent and identically distributed (i.i.d.), with  $\zeta \in \mathcal{K}^+$  an index variable demarcating each independent play of the approval regulation game. Each product distribution is perturbed by mixture with another Beta distribution, where the mixture parameter  $\alpha_\zeta$  is expressed as the integer ratio  $a'_\zeta/b'_\zeta$  where  $a'_\zeta \leq b'_\zeta \forall \zeta$ . Assume the moments of  $\alpha$  follow those of a Beta distribution. If  $G_R^\Sigma$  and  $G_U^\Sigma$  are the asymptotic weighted averages of the regulated and unregulated product distributions, respectively, then  $G_R^\Sigma$  and  $G_U^\Sigma$  are approximately normal, and  $G_R^\Sigma$  has second-order stochastic dominance over  $G_U^\Sigma$ .*

Proof appears in the online Appendix, save for the stochastic dominance relation, which appears in the Appendix to this article.

Lemma 2 essentially says that if the approval regulation game were averaged over many independent plays, the resulting distribution of product quality would be continuous and approximately normal. More important, the stochastic dominance relations in Lemma 1 would be preserved under this averaging, which implies that the regulated distribution of product quality under repetition would still have less uncertainty than the unregulated product distribution. Invocation of Lemma 2 permits the result of the approval regulation game to be applied to learning and utilization problems that have continuous representations of uncertainty as well as discrete representations. This can be helpful for when the aggregate consumption or utilization is represented by the decisions of a single, representative agent (e.g. a consumer). Using Lemma 2, this agent will face a single-mode distribution of product safety or quality, one in which uncertainty is represented by a single, continuously valued parameter.

## 5. A generalized dynamic model of utilization under uncertainty in an approval-regulated market

We now turn to model utilization or consumption in the “market” for the products approved by regulation. In so doing, we present a very stylized version of utilization or “consumption” in which a large number of human agents each uses one product at a time, switching among products to find the best one available.<sup>9</sup> The central fact governing

the agents' consumption decisions is uncertainty – the different products can be rank-ordered by the prior estimates of their quality (their “initial appearances”) but these are only best guesses subject to error. Beyond the initial best guesses, human agents' uncertainty about products can be reduced only through experience, such that the regulated product is an experience good.

The uncertainty is expressed as a dynamic stochastic process, and here we use the canonical version of uncertainty or randomness, a Wiener process or “Brownian motion.”<sup>10</sup> One can think of a Brownian motion as an all-purpose random process whose independent movements in continuous time occupy a continuous state space. For more than a century (since Louis Bachelier's *Théorie de la spéculation* [1900]), Brownian motion has been used to model the movement of asset prices, and a general example of its applications to a wide variety of investment decisions occurs in Dixit and Pindyck (1994). Brownian motion is useful for our purposes (and many others) because it offers a simple, additive, and linear expression of uncertainty.

We begin with a single human agent (“the Agent”) who can be thought of as an end-user or consumer of the products produced by the Firm and governed by the Regulator. Let the incumbent product under consideration be indexed by  $i$  (where the countable set of available products is indexed by  $j$ ) and let the consideration time for product  $i$  be given by  $t_i$ . A fallback product with known value is denoted by  $i = 0$ , whereas all incumbents and uncertain products (“arms” of the bandit) are denoted by  $i \geq 1$ . We suppose that each product is characterized by a *quality* parameter. In a health example, this parameter will reflect the product's effectiveness in treating the disease; in a financial example, it will reflect its contribution to the Agent's (the investor's or asset holder's) welfare. A product's quality is a draw from a Normal distribution,  $\beta_i \sim \Phi(b, v_b)$ , where  $\Phi(b, v_b)$  represents the Normal distribution with mean  $b$  and variance  $v_b$ . The actual value of  $\beta_i$  is unknown to the Agent, but is learned from the experience of utilization; only one product can be utilized at a time.<sup>11</sup> A key point of intuition is that, by lemmata 1 and 2, products that have gone through a pre-market process of approval regulation will, *from the start*, be associated with a lower value of  $v_b$  (the prior variance of product quality/safety).

### 5.1. Continuous-time evidence of quality

The Agent collects continuous-time evidence about a product's quality according to Brownian motion with drift, where the drift is determined by the (unobserved) quality,  $\beta_i$ , of the case. Formally, the Agent observes and experiences the realized value of the incumbent product,  $X_i$ , which evolves according to the following stochastic differential equation (“Law of Motion”):

$$dx_j(t) = \beta(x_j(t))dT_j(t) + \xi(x_j(t))dz_j(T_j(t)); t > 0$$

where  $T_j$  is the learning or utilization time for the  $j$ th product and  $z_j$  is a standard Normal distribution with mean zero and variance  $t_j$ .

#### 5.1.1. Intuitive explanation of the “Law of Motion”

The “Law of Motion” for an observed product may appear complicated, yet it can be simplified by walking term-by-term through the equation. On the left-hand side of the equation are *movements in quality* ( $dx_j(t)$ ), and these are governed by two forces on the

right-hand side of the equation: the *movement induced by the underlying quality itself* ( $\beta$ ) over the interval  $dT$  (hence  $\beta \times dT_j(t)$ ) and *movements in a random variable*  $z$  over the interval  $dT$  ( $\xi \times dz_j(T_j(t))$ ). The parameter  $\xi$  encodes the amount of information the “Law of Motion” equation contains for the regulator: if  $\xi_i = 0$ , then the regulator can immediately infer the quality of the case by examining the slope of the quality time-series and as  $\xi \rightarrow \infty$  the SDE contains no information about a product’s quality.

### 5.2. Estimating quality from evidence

Given that the Agent observes only  $X_i(t)$ , we first prove that the learning problem is identified (Chernoff 1968): the Agent is able to disentangle the contribution of the quality of the case to  $X_i(t)$ .

Let  $\Pi \in (U, R)$  denote whether the product in question is regulated or unregulated. Then the Agent can calculate a posterior mean and posterior variance as follows:

$$\text{Posterior mean} \equiv E_{x_i}(\beta_i) = \hat{\beta}_i = \frac{b/v_b^\Pi + x_i/\xi^2}{1/v_b^\Pi + t_i/\xi^2} \quad (4)$$

$$\text{Posterior variance}(\hat{\beta}_i) \equiv V_{\hat{\beta}_i}^\Pi(t_i) = \frac{1}{1/v_b^\Pi + t_i/\xi^2} = \frac{1}{[v_b^\Pi(\theta^\Pi, n^R)]^{-1} + \xi^{-2}t_i}. \quad (5)$$

### 5.3. Filtered evidence and value functions

The Agent seeks to define an optimal continuation rule for the *filtered evidence process* found by combining Equations 4 and 5, the Agent faces a convex function,  $\hat{\beta}_i(t) \times t \mapsto \Psi(\hat{\beta}_i(t), t_i)$ , that is twice differentiable with respect to both  $\hat{\beta}_i(t)_i$  and  $t$ . This function is a map from the current state of the filtered evidence process and time to the value experienced by the Agent. For any particular product, the Agent wishes to maximize

$$E \int_0^\infty e^{-\delta t_i} \bar{h}(i(t_i), x_{i(t_i)}, q_i) dt_i$$

where  $\delta$  is a discount factor interior to the unit interval and  $\bar{h}(x_i)$  is a running payoff function or flow value, realized when observed quality is  $x_i$  and  $q_i$  is the per-unit-time ( $dt$ ) cost of utilization, with  $\frac{dq_i}{dt} \equiv 0$ ,  $\frac{d\hat{\beta}_i(t_i)}{dq_i} \equiv 0$ , and  $\frac{\partial \bar{h}(x_i)}{\partial q_i} < 0$ . For the moment (until consideration of Proposition 6 below), assume that utilization cost is identical across products:  $q_i \equiv q_j \equiv q$ ,  $\forall i, j$ .<sup>12</sup> In the health example,  $\bar{h}$  can be considered as the patient’s realized and experienced health; in many investment and financial examples, it is construed as a dividend. For the following analysis we will replace  $x_i$  with  $\hat{\beta}_i$ , without loss of generality due to the scale-invariance property of  $X_i(t_i)$ .

### 5.4. Utilization/consumption objective

The Agent can utilize only one option at a time (the “incumbent” option) but faces a countable sequence of alternatives, including one that has certain value (a “fallback” option). For each incumbent, the Agent’s objective is to define an optimal rule to stop  $\hat{\beta}_i(t_i)$  in order to maximize her or his expected reward, which is given by

$$J(x, M; \mathfrak{S}, \tau) = E_x \left[ \int_0^\infty e^{-\delta t} \bar{h}(i(t), \hat{\beta}_{i(t_i)}^*(t_i^*)) dt + M_i e^{-\delta \tau} \right]$$

where  $M$  is an idiosyncratic termination payoff that is static, positive, and known with certainty, and  $\tau$  is a stopping time at which point the Agent hypothetically switches to the known fallback option (described below) or a product with its certainty equivalent. The quantities  $\beta_i^*$  and  $t_i^*$  are the case's quality and current time, respectively, under the optimal learning policy.

The solution to a problem of this nature is well known and entails the human Agent's calculation of a dynamic allocation index (DAI) or "Gittins index." Intuitively, the Gittins index of an uncertain product is the minimum certain reward that an Agent would choose over that product, given everything that the Agent knows about the uncertain product. For any given  $\beta_0$ , let  $M = E \int_0^\tau e^{-\delta t_i} \beta_0 dt_i$  stand for the expected terminal reward associated with choosing the fallback option ( $i = 0$ ) until  $\tau$ . The Agent's optimal strategy is to establish a "follow the leader" policy associated with Gittins index  $Q_i$ , defined as:

$$Q_i(\hat{\beta}_i(t_i)) = \inf \left\{ \kappa \in \mathfrak{R}^+ \mid \sup_{\tau > 0} \frac{E \int_0^\tau e^{-\delta t_i} \bar{h}(i(t_i), \hat{\beta}_{i(t_i)}) dt_i}{E \int_0^\tau e^{-\delta t_i} \beta_0 dt_i} = \kappa \right\}$$

where the supremum is over all  $\mathfrak{S}_i$ -stopping times that are positive "almost surely," and at least one is not necessarily finite. Given a calculable Gittins index for any option, the Agent's optimal strategy is to choose the option with the maximum Gittins index (Karatzas 1984).

**5.5. The optimal stopping rule and its properties**

Because each product is associated with a prior  $\beta_j^0 = \hat{\beta}_j(t_j = 0)$ , the battery of products can be ranked by their prior value. As has been amply demonstrated in the field of mathematical statistics (Karatzas 1984; Mandelbaum 1987), the Agent's optimal policy is to utilize the product with the highest initial estimated quality and pursue an optimal switching policy from there. (Ties have measure zero and the Agent can be assumed to randomize among tied products.) The only additional assumption required to pin down the model is the existence of a *revisitation penalty*,  $\chi > 0$ , for any abandoned product, such that once the  $j$ th option is abandoned, its prior for any possible future consumption will be reduced by  $\chi$ .<sup>13</sup>

The Agent's sequential problem is equivalent to following the leader according to an optimal stopping policy: Begin utilizing the product with the highest  $Q_i$ , then abandon this incumbent if and only if, and when and only when, an optimal stopping criterion is satisfied. The Agent's optimal switching policy will be to observe the first passage of the evidence process,  $\hat{\beta}(t)_i$ , through a border that encodes the tradeoff between continuation of experience with the incumbent product and the value of switching (with partial irreversibility because  $\chi > 0$ ) to the next best product, where "next best" simply encodes the best prior among the non-incumbent options:  $\sup_{1 \leq j \leq d} \beta_j^0$ . The form of the barrier is described in Proposition 3.

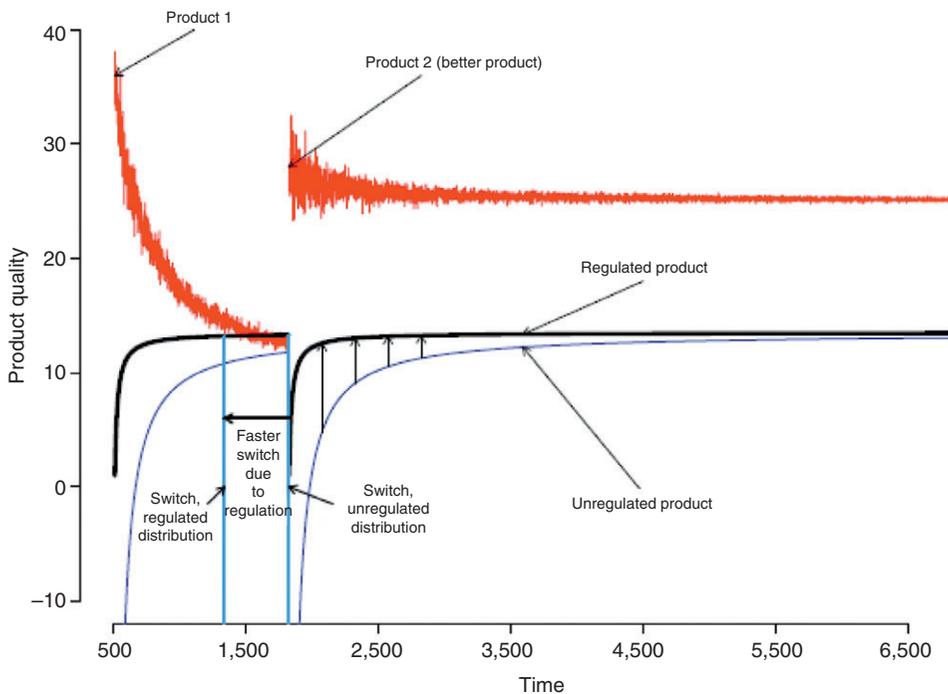
**Proposition 3:** *Optimal stopping barrier for each incumbent product.*

*The Agent switches products when and only when, and if and only if,  $\hat{\beta}(t)_i$  passes for the first time through the optimal stopping barrier established by the availability of the next-best product,*

$$\gamma^*(t) = \delta Q_j + \frac{1}{2\xi^2} \Psi_{\hat{\beta}, \hat{\beta}}(\hat{\beta}(t), t) V_{\hat{\beta}_i}^{\Pi}(t_i)^2 \quad (6)$$

where  $\Psi_{\hat{\beta}, \hat{\beta}}(\hat{\beta}(t), t)$  is the second partial derivative of the value function  $\Psi$  with respect to the filtered state variable  $\hat{\beta}$  given a realization of  $\hat{\beta}$  at time  $t$ . For the terminal product with known value, replace  $Q_j$  with  $M$ .

This border represents the optimal tradeoff between continuing with the incumbent product (and delaying the utilization of the next-best product) and switching instantly. If the Agent delays utilization of a case, she or he receives more information, reducing the value of  $V_{\hat{\beta}_i}^{\Pi}(t_i)$ , which under the optimal stopping policy declines quadratically with the length of the Agent's experience. Figure 2 shows the filtered evidence process and the optimal stopping barrier from Proposition 3. The horizontal axis is time, the vertical axis is the utility provided to the Agent, the random (gray) series at the top of the graph is the evidence process for one product, and the bold and non-bold black lines represent families of optimal stopping barriers for the set of alternative products. The border slopes upward as the value of more information decreases over the course of the regulatory history. A product is abandoned only if its evidence crosses the boundary, which occurs at the left-hand side of Figure 2. Figure 2 demonstrates how regulation of products can lead to agents moving more quickly toward a better product. Notice that regulation shifts the stopping barriers upward. As a direct result of this upward shift in the barrier, the Agent would switch from the inferior first product to the better second better product earlier (all things being equal). The results depicted in Figure 2 illustrates how the second-order effects of regulation translate into first-order effects on how consumers use the products.



**Figure 2** Approval regulation induces consumers to more quickly abandon inferior products.

We are now in a position to advance our principal theoretical result. Before doing so, we restate two central features of our elaboration so far – that the payoff structures of the approval regulation game and the market are linear, additive, and entail only the “first moment” or expected value of the random variables involved. In other words, not one actor in the models we have presented – Firm, Regulator, or Agent (end-user) – is “risk-averse,” and none of these actors directly values higher moments or orders of the uncertainty (such as the variance or skew) they experience.<sup>14</sup>

**Proposition 4:** *Higher switching rate to superior products by agents under a regulation-induced quality distribution when the unregulated product has superior priors.*

Consider any two alternative products ( $j = 1, 2$ ) where the first arrives to the market via approval regulation (is regulated; R) and the second arrives without approval regulation (is unregulated; U). Assume

- 1 that the products are of identical value and that both are superior to the incumbent product, such that  $[\beta_{j=1,R}^{\hat{\beta}_R(0)} \equiv \beta_{j=2,U}^{\hat{\beta}_U(0)} > \beta_i^{\hat{\beta}_i(0)}]$ ,
- 2 that the alternative products generate identical coefficient histories,

$$H(X_{1,R}(t_j) | \beta_{1,R}^{\hat{\beta}_R(0)}) \equiv H(X_{2,U}(t_j) | \beta_{2,U}^{\hat{\beta}_U(0)}), \forall t_j$$

- 3 and that the initial quality estimate of the unregulated alternative product is higher than the initial quality estimate of the regulated alternative ( $\hat{\beta}_U(0) > \hat{\beta}_R(0)$ ).

Then the Agent still switches more quickly to the regulated product (R) at  $t_j = t_{switch}$  unless the prior difference is sufficiently large that

$$\hat{\beta}_U(0) - \hat{\beta}_R(0) > \frac{\xi^2 + \nu_b t_{switch}}{\xi^2} [V_{1,R}^\Pi(t_{switch})^2 - V_{2,U}^\Pi(t_{switch})^2].$$

This result is actually quite general because it relies upon the equivalence of stochastic histories – the Agent’s accumulated data on the regulated and unregulated products – and because the stopping times used in the proof are arbitrarily chosen (and, given the equivalence of histories, substitutable across any two products). More complicated (but no less rigorous) proofs are available using stochastic integration.<sup>15</sup> Moreover, the equivalence of stochastic histories can be relaxed slightly and analysis of the model yields a similar result (see Lemma 3 in the Appendix). We are now in a position to state a more general result, which follows as a corollary from Proposition 4.

**Proposition 5:** *Higher switching rate to superior products by agents under a regulation-induced quality distribution.*

Consider any two products ( $j = 1, 2$ ) satisfying assumptions 1 and 2 of Proposition 4. The Agent operating under the optimal dynamic allocation policy will always switch more quickly and with higher probability to superior products under regulation than under an unregulated product distribution.

Propositions 4 and 5 offer one other lesson that has some mathematical interest but is also of relevance to more practical policy considerations and is, furthermore,

testable, perhaps both experimentally and observationally. Even small reductions in the prior uncertainty of the set of potentially utilized products have great effects because the reduction of uncertainty behaves according to quadratic convergence.<sup>16</sup> So even a small amount of induced confidence from regulatory institutions can have a great effect upon utilization of the product, because initial data have the greatest marginal effect.

### 5.5.1. Implications for consumption and price sensitivity of agents

The results in Propositions 4 and 5 have an interesting corollary for the role of price in a regulated market. The enhanced pre-market information about a product may induce a behavior where consumers are more price-sensitive because they are less sensitive to a priori uncertainty about product safety or quality. A full investigation of this dynamic would require the technical construction of the price elasticity of demand in the bandit model; we instead proceed with the marginal change in the dynamic Agent's willingness to pay for the incumbent product for another period ( $dt$ ) for each change in price.

**Proposition 6:** *Regulation and price sensitivity of switching rate to superior products.*

*Consider again any two products ( $j = 1, 2$ ) satisfying assumption 2 of Proposition 4, but now the underlying quality of the incumbent is not different from that of the alternatives ( $\beta_{j=1,R}^{\hat{\beta}_R^{(0)}} \equiv \beta_{j=2,U}^{\hat{\beta}_U^{(0)}} \equiv \beta_i^{\hat{\beta}_i^{(0)}}$ ), and let  $q_j \equiv \underline{q} < q_i \forall j$ . Let  $q$  contribute to  $\bar{h}$  by a strictly monotone decreasing and twice-differentiable function,  $v(q)$ , so that  $\underline{q} < q_i \forall j$  is sufficient to yield  $\bar{h}_j > \bar{h}_i$  for  $j = 1, 2$ . Then the Agent utilizing products according to the optimal dynamic allocation policy will always switch from the incumbent to the lower-priced alternative more quickly under the regulated product distribution.*

The intuition of Proposition 6 is as follows. An Agent in the unregulated market will hold on longer to an inferior (more costly) product given that her or his knowledge about the distribution of (potentially superior) alternative products facing her or him is subject to greater uncertainty and given that abandonment costs are strictly positive ( $\chi > 0$ ). Price differentials are alone sufficient to generate switches, and these switches occur more quickly under approval-regulated product distributions than under distributions without approval regulation.

## 6. Implications and conclusions

### 6.1. Consistency with studies of approval regimes and other forms of regulation

Interpreted accordingly, Propositions 4 and 5 establish a correspondence between the predictions of the current model and the interesting empirical results of Law (2003), Law and Kim (2005), and Law and Marks (2009). Law (2003) adduces panel regression evidence that state food regulation was associated with increased consumption of regulated foodstuffs. To be sure, state food regulation did *not* impose the approval-regulation scheme upon food products, but it did create quality-based entry barriers and in numerous states, laboratory and (animal) pharmacological testing of food additives, food preparation and processing processes. These experimental regimes were a critical accompaniment to the regulatory institutions of the time. Along with the US Department of Agriculture's capacities, such chemical and pharmacological testing of food and drug substances in the states was far progressed beyond other regulatory arrangements in the world or even in other realms of regulation (Law 2006). More

relevant to the current model because licensure often involves a feature of pre-market testing, Law and Marks (2009) show benefits for minorities from professional licensure regimes because revealed information about worker quality can more easily trump discriminatory assumptions about minorities' abilities or training. As Law and Marks argue, "in those occupations in which minorities are underrepresented but for which information about worker quality is costly to obtain, licensing can reduce statistical discrimination" (Law & Marks 2009, p. 364).

Our results are also consistent with recent arguments in law and economics scholarship that approval regulation regimes may not blunt innovation but enhance it (Eisenberg 2007; Katz 2007). As Katz argues, regulation and innovation may not be at odds, mainly because "drug regulation provides certification of drug quality" (p. 1) (see also Carpenter 2010). This certification mechanism "may not be easily achieved by private market-based mechanisms," and by regulation, it "prevents the market from becoming a market for 'lemons' " (p. 1). From this point, Katz's argument relies upon a mechanism that the expected returns to innovation (higher quality products) will be enhanced under US Food and Drug Administration (FDA)-like regulation, and that "rather than decreasing the expected returns to innovation, this aspect of regulation contributes to the value of new drugs and may actually encourage innovation" (p. 1).

Finally, our model provides a formal account consistent with recent evidence for consumption increases following data disclosure regulations at the municipal and state level (although these are not regimes of approval regulation strictly speaking). Bollinger *et al.* (2010) used detailed data from Starbucks stores to examine responses to the calorie posting law in New York in 2008, and found that "There is no impact on Starbucks profit on average, and for the subset of stores located close to their competitor Dunkin Donuts, the effect of calorie posting is actually to *increase* Starbucks revenue." Jin and Leslie (2003) found similar evidence in a study of restaurant hygiene ratings. These studies and others provide evidence consistent with the claim from the optimal product utilization model of the market "downstream" from regulation: reduced uncertainty may be associated with increased consumption and more probable abandonment of inferior products.

## 6.2. Institutional analogies for financial regulation

Could regimes of approval regulation be applied to consumer financial products? With recent legislation creating a Consumer Financial Protection Bureau in the Federal Reserve, there are many regulatory models available. Consider the following remarks.

Why not create a Financial Products Safety Commission charged with responsibility to establish guidelines for disclosure, collect and report data about the uses of different financial products, review new products for safety, and require modification of dangerous products before they can be marketed to the public? . . .

An FPSC would *promote* the benefits of free markets by ensuring that consumers can enter financial services markets confident that the products they purchase meet minimum safety standards. (Warren 2008)

A financial products safety commission could help fill in the gap, particularly in relationship to products being produced by and invested in by regulated entities. Each product would have to have a stated objective (e.g. in what ways was it helping manage and mitigate risk; what was the risk profile for whom the product was intended). Its risk characteristics would be identified, using conservative models

which paid due attention to the failures previously noted. The Financial Products Safety Commission would evaluate whether products provided significant risk mitigation benefits of the kind purported by the product. (Stiglitz 2008)

The mechanics of such institutions would demand particular care and abundant thought.<sup>17</sup> One can imagine a regulator that would take new financial products and subject them to experiments. These experiments could be conducted before market entry, as in the model presented here. Alternatively, the tests could be “roll-out” experiments that were conducted as the product entered use, with the added possibility of regulatory withdrawal if the experiments revealed severe risks that were not compensated by, say, the benefits of completing a market.<sup>18</sup>

What might these experiments look like in practice? Consider several possibilities. One is that a group of psychologists and behavioral economists conduct laboratory experiments with new products, examining properties of consumer learning and consumer choice with the new product relative to learning and choice with products already on the market. Subjects could be randomized to the new product or the old products, or to trial versions of the new product with different terms of disclosure. Field experiments might also be possible, although contamination of field experiments by active marketing is a real possibility. Alternatively, a regulator might conduct a kind of “financial epidemiology” by examining patterns of product use as the market grows and as observable problems (e.g. mortgage defaults, personal bankruptcies) arise. Finally, a regulator might conduct simulations with new products, where the simulations consider economic conditions of varying adversity, ranging from systemic shocks to cyclical declines (Candell 2010). These simulations could perhaps be informed by the results of more rapid laboratory experiments. Armed with this information, an approval regulator might have veto power over market entry (as with the FDA and licensing regulators), might be able to study products while they entered limited markets in a test phase, or might be able to compel withdrawal of products deemed to have been shown especially hazardous.

Yet the potential problems with such a system would be many, and they would require institutional design and, in all likelihood, continuous institutional adaptation. For one, it is not clear what a “new financial product” or even a “financial product” is. This difficulty might well be insuperable, although there are two reasons to think it may not be. In the field of insurance regulation, many states have begun to adopt “file-and-use” systems in which new entrants to the market notify a state regulator of any new insurance product as defined by a substantial change in the terms of the contract. These regulatory regimes do not have pre-market approval or regulatory veto, but the novelty of the product is defined by changes to the contract (including, but not limited to, changes in rates). Because most consumer financial products (e.g. mortgages, credit cards) exist as contracts, one could plausibly imagine a similar set of rules governing new financial products. The second reason for guarded optimism is that a new product category, just like other regulatory and legal categories, would likely get shaped over time in administrative rulemaking and in legal decisions. The initial difficulty in defining regulated products might not prevent the longer-term emergence of a “conceptual equilibrium” in which most innovations could be clearly categorized as to whether or not they fell under the auspices of a regulation and its administering agency.

One might also worry that approval regulation for consumer financial products would induce consumers to be too friendly to new mortgages, loans, and other instru-

ments. To be precise, this is not a problem with our model per se, as it renders a more compact prediction – not so much that consumption in general will increase, but that consumers will switch more readily to products that are in fact superior for them. Yet in general, it is a plausible worry that a regulatory floor may induce behaviorally limited consumers to attach too much faith to new financial products.

One might wonder, finally, whether there are simply cheaper ways to achieve the second-order effect of regulation that is described here. From the standpoint of our model, there would be two responses to the concern that confidence can be achieved more cheaply with non-regulatory mechanisms. First, safety regulation does have first-order effects as well as second-order effects, and so if our model is right (on average), not only confidence will improve under approval regulation, but so will equilibrium safety and quality themselves. Second, it is debatable whether society or private firms will be able to credibly commit to a rigorous, high-quality experimental program in the absence of a regulatory veto that compels them to do so in order to gain market entry. So if there is a cheaper way to generate (public good) information about product quality, it would have to be accompanied by incentives for the production of high-quality information such as that obtained from randomized, controlled trials.

### 6.3. Policy implications: Veto-strong regulators, institutional confidence, and diffusion of regulation-induced information

We have extended the analogy between safety regulation and financial regulation by focusing on a critical property of institutions of safety regulation and a critical property of modern markets for complex products: (i) the regulatory veto and its induction of experiment and (ii) the central role of uncertainty in utilization decisions over time. Warren (2008) and others argue that product safety regulation ought to be a model for financial regulation in many forms. The argument here demonstrates that the critical mechanism that renders safety regulation powerful is the *forcible veto* (or removal power) of the government regulator that oversees the marketplace. It is this negative power that induces companies to provide more information about their products, and it is this negative power that induces firms to provide a particular kind of information (quasi-experimental information, where the evidence is a truly random and representative draw from an underlying parameter of the product, with independent and identically distributed errors that permit valid inference). In the modern health care world, the phased clinical studies system of the FDA ratified in 1963 serves this important public function (Avorn 2004, Carpenter 2010).<sup>19</sup>

At its core, then, our model suggests that *some of the most important benefits of safety regulation have nothing to do with safety*. Approval regulation leads to a superior distribution of products, most likely, but regardless (and independent) of this result, approval regulation also leads to the provision of more information, and information of a higher quality, than would be provided in the absence of approval regulation institutions. *Information provision alone is sufficient to induce risk-neutral as well as risk-averse consumers to switch more readily to the products that will in fact benefit them.*

Of course, our article points to the crucial feature of public confidence in a regulator (Carpenter 2010). In the model, if the Agent's prior variance is not reduced, then the confidence effects that generate more optimal switches among products do not materialize. If then the regulatory process is doubted, or the quality of the information produced by that process is subject to known bias, the confidence benefits of approval regulation

may be dashed. In recent years many regulators (particularly the FDA) appear to have suffered a material loss of public confidence (Carpenter 2010, ch. 12). While a loss of regulatory confidence is not in and of itself sufficient to dash all of the benefits of approval regulation institutions (even poorly trusted regulators can prevent bad products from coming to market and thereby improve the quality of utilization), vital benefits of approval regulation are forgone when agents believe that the Regulator is being duped, is colluding with the Firm, or is otherwise negligent.

#### 6.4. Limitations and extensions

As with all such theoretical efforts, our model has some basic limitations. First, what we have not yet done is to think about safety and the very real possibility that someone can get defrauded (the “exploding toaster” problem in Warren [2007, 2008]). Second, the model relies upon simplifying assumptions throughout, most notably the existence of a single firm at the regulatory stage and the exogeneity of consumers’ strategies to regulation.<sup>20</sup> While this separation of regulation from the utilization strategies of consumers may seem unduly constraining, it is instructive for several reasons. First, the connection drawn between regulation and consumer utilization is one that rests on the communication of information produced in the regulatory process to “downstream” users who might benefit from that information (such as doctors and patients who might benefit from the extensive knowledge produced in clinical trials conducted for FDA approval of a drug or medical device). This connection is accomplished in part through *labeling and other warning mechanisms*, which are little studied in political economy. Our model makes it clear why these little-studied mechanisms of information transmission may in fact be so important. Second, our theoretical separation of regulation from consumption corresponds to the *de jure* separation of these phenomena in actual regulatory institutions. Many quality and safety regulators do not officially or explicitly consider downstream price and demand effects in their approval or licensure decisions (see for instance the provisions of the Federal Food, Drug and Cosmetic Act of 1938).

#### 6.5. Is a veto necessary? The promise and perils of limited licensing

Critical readers might wonder whether a market-entry veto is a necessary or an optimal way of achieving the confidence benefits of safety regulation. More generally, one might suggest that the veto is only a limiting case of a more general mechanism that imposes market-entry costs that are correlated with the accumulated evidence for product quality. An important limitation of our effort is that our model examines only this limiting case – where the regulator-imposed costs of market entry are infinite (or sufficiently large that no entry occurs). Hence we can claim from our model only that, under plausible conditions, consumers are better off in a regime with a veto than without any veto at all. Yet one might wonder whether more intermediate measures would suffice to generate the confidence benefits described here.

One possibility is limited licensing or what for the pharmaceutical market Manski (2009) has called “adaptive partial drug approval.” Consider the following intuitive example. In a market of potentially a million customers, the regulator might proceed according the following rule. If the evidence for quality is less than fair, no sales are allowed; if the evidence is fair but plausible, the sponsor gets a license to sell to 50,000 customers; if the evidence is pretty good, the sponsor gets a license for 250,000 customers;

if the evidence is good the license is for 500,000; and if the evidence is very good the sponsor gets a license to sell to 750,000 to 1,000,000 customers. The resulting policy would have a *monotonicity property* that links better experimental evidence to an increased size of the allowed market.

Limited licensing offers interesting possibilities. Indeed, limited licensing suggests one interpretation of recent regulatory tools adopted by US drug regulators, notably the Risk Evaluation and Mitigation Strategies (REMS) embedded in the Food and Drug Administration Amendments Act (FDAAA) of 2007. In its September 2010 decision on the diabetes drug rosiglitazone (Avandia; GlaxoSmithKline), the FDA responded to evidence of cardiovascular risk not by removing the drug from the market, but by limiting its prescription to those diabetic patients who have exhausted other alternatives and who can attest in appreciable paperwork requirements that rosiglitazone is the only alternative to them. Prescribing other drugs under this regulatory control strategy will be much easier and hence rosiglitazone's market is expected to decline significantly to a small fraction of its former breadth.

Yet alternatives to a veto do not come without problems. One is that, in order for the partial approval license to work, there must be stringent regulatory control of production; someone, somewhere, has to be able to enforce the contract that the company can produce no more than (say) 250,000 units. This requires not only legal authority for market size limits, but also the capacity to estimate the size of the market in the first place. Perhaps most daunting is that, in a large marketplace, production control may necessitate substantial government expenditures through the employment of inspectors and monitors. In a world of free advertising, moreover, a company can convert a small market into a much larger one – the past decade's experience with large-market pharmaceuticals in the US certainly suggests this possibility. The difficulty of production control combined with incentives to inform by advertising mean that, *without a regulatory veto, companies will have a difficult time making a credible commitment to further experimentation once their foot is in the door of the market*. Because experimentation and advertising are often substitute ways of informing consumers, companies will often prefer advertising to experiments once a small market has been established. Again, the experience of American pharmaceutical markets is informative, as few companies have complied with FDA requirements for post-approval clinical trials (often called Phase IV trials). Of the thousands of commitments that drug companies have made to carry out these studies in recent years, only a third of the required trials have even been started, much less fully executed and reported (Carpenter 2010, p. 609).

The properties of entry vetoes in approval regulation and partial licensing demand further investigation. One could examine various types of errors and health or welfare consequences, both theoretically and empirically. Our analysis suggests, however, that inducing a credible commitment to experiment remains a serious issue for consideration of any alternative to veto-based institutions.

## **6.6. Implications for cost–benefit analysis, and empirical and historical agendas**

We think that critical research agendas lie in the empirical and historical extension of “confidence” effects in regulation. Historians and historically oriented scholars may wish to ask a variety of questions: (i) How is trust created among wide and dispersed populations of human agents in response to consumers? (ii) What are the historical mechanisms and contingencies by which regulation affects consumers' beliefs about the

set of alternative products they are facing? (iii) Because there is no such thing as a single, undifferentiated consumer – instead there are heterogeneous audiences – how do differing perceptions of the regulator shape the confidence effects of institutions?

Empirical analyses, too, are needed, but our theory suggests that they should be focused not only on the avoidance of drastically bad outcomes but also on the quieter but no less vital outcome of everyday utilization of products. The empirical literature on regulation is shot through with attempts at weighing the safety- or externality-related “benefits” of the policy with its imposition-related “costs” (see the representative efforts of Viscusi [1993] or Boardman *et al.* [1996]). Combined with the studies of Law (2003), Law and Kim (2005), and Law and Marks (2009), our model points to systematic patterns of regulatory effect that are not captured by this literature. At a minimum, many efforts in contemporary cost–benefit analysis of regulation are arguably affected by omitted variable bias.<sup>21</sup> Consumer and public perceptions are far more than a matter of public opinion and survey research; they are integral to optimal administration of, and proper measurement of the benefits and costs of, a wide variety of regulatory policies.

Libertarian analyses of approval regulation – such as licensure, permitting, the governance of consumer products by the Consumer Product Safety Commission (CPSC), and the approval of new therapeutic products by agencies like the FDA – tend to predict a restriction of supply and a bevy of poor policy outcomes: a corresponding increase of commodity prices, reduction of product consumption, and in general a limitation of freedom. The theory presented here (and the set of related models that can be built upon it) points to a quite different dynamic, one that is better attuned to the fact of learning (even imperfect and partially irrational learning) among human agents. Institutions of regulation that limit market entry but do so by compelling forms of experimentation or knowledge generation can often create a more predictable marketplace. Both risk-neutral and risk-averse human agents will more readily enter the marketplace created by approval regulation and will more readily rely upon quality data to switch to the products that present them with the most value.

## Acknowledgments

Earlier versions of this article were presented at the Tobin Project Conference on Economic Regulation, White Oak, Florida, April 2009, and the Conference on Consumer Financial Regulation at University of Pennsylvania Law School, November 2009. We thank Tom Baker, Eric Budish, John Y. Campbell, Danny Goroff, Richard Herring, Howell Jackson, Brigitte Madrian, David Moss, Pablo Spiller, Michael Ting, Susan Wachter, and Elizabeth Warren for comments on previous versions. For research support we acknowledge the Tobin Project, the Alfred R. Sloan Foundation and Russell Sage Foundation Program on Behavioral Economics and Consumer Finance, and the National Science Foundation (SES-0076452).

## Notes

- 1 In particular, Warren (2007) proposes a concrete analogy between a form of regulation that is widely known and familiar in one set of markets (e.g. medical products, licensure) and the absence of such institutions in financial and lending markets. She also introduces the concept of “safety” to discussions of financial regulation.

- 2 The elaboration follows Carpenter and Ting (2007), and many lateral proofs are presented there. For a decision-theoretic analysis of product approval in the context of multiple firms and sequential submissions, consult Carpenter (2004).
- 3 Because firms submit only one product in the model, we equate the type of Firm with the type of its product. It is possible to separate quality and safety and consider  $x$  to be their additive (hedonic) sum (e.g. “quality – safety”) but a more compelling separation of the two variables demands a much more complicated model with dual experimental strategies and possibly dual regulatory strategies. We leave this endeavor to another effort.
- 4 The reputation games of Calvert (1987) and Alt *et al.* (1988) use a similar technology.
- 5 Future inquiry into a repeated game that builds on the model in Sections 2 and 3 is a desirable task, as firms may act to build and preserve reputations with the regulator, and regulators may seek to do the same with firms. However, these issues involve a material step away from the focus of the present article and, just as important, they raise serious issues of mathematical tractability.
- 6 The perturbations are, of course, not so large that the basic assumptions in the model are violated. The proof of Lemma 2 relies upon quite general mixture distributions to represent each perturbation, and so limiting the perturbations by a certain bound holds as a special case.
- 7 Put differently, Lemma 2 allows us to construct a mathematical map between the regulatory process and a distribution of consumed products, but where we do not model both jointly. This remains as a research agenda, but see the Implications and conclusions for a discussion of very real barriers of mathematical tractability for doing so, and see the Introduction for the difficult but pointed research questions that arise from thinking about this real-world map as a next modeling and empirical step.
- 8 In order for the results of the model to hold, it is *not* the case that all regulation-induced information must be used; even a small amount of reduction in the prior variance of product quality/safety will suffice to induce the consumer behavior described.
- 9 We eventually assume that there are many human agents, such that the market is “large,” or at least sufficiently large that coordination among agents becomes prohibitively difficult. This lack of ability to coordinate – say, for consumers to conduct their own tests of products – can be thought of as contributing to the need for a regulator.
- 10 We aim eventually to embed a multi-armed bandit model in this Levy process framework.
- 11 As long as convex bundles of products impart some randomness relative to previously utilized components, the results of this one-product-at-a-time model will generalize easily.
- 12 The utilization cost is a commodity price, and we are assuming its exogeneity to user behavior and regulation here. Endogeneity to user behavior would necessitate a dynamic stochastic equilibrium as in Jovanovic (1979), who does not model regulation or an alternative battery of products/jobs.
- 13 Thus assumption prevents infinite cycling between two arbitrarily close alternatives, which would not affect the results of the model or the Propositions, but would complicate simulations. The assumed homogeneity of  $\chi$  is necessary to maintain the computability of the Gittins index (Banks & Sundaram 1994). Because the revisitation penalty is the same for all arms, the value of the Gittins index depends only on the present arm; there is no cost experienced or expected upon a switch.
- 14 Of course, the Agent in the diffusion bandit model does embed second-order considerations in dynamic utilization, but this is induced because  $\chi$  is strictly positive.
- 15 We say no less rigorous because the equivalence of stochastic histories implies a control for the measure-zero event “all other things held constant.”
- 16 The quadratic form of the convergence ensues functionally from Ito’s Lemma in the present model. More broadly, it is a specific instance of the law of the iterated logarithm (Karatzas & Shreve 1991, pp. 111–114).

- 17 We are deeply grateful to Howell Jackson for discussions on this mapping and for his criticisms. We avoid any discussion of approval regulation of “systemic” risk, such as bank-to-bank loans and investments.
- 18 It is well known by observers and scholars of consumer credit firms that companies selling credit cards conduct both experimental and observational studies of their products, both before and after market entry. So at least for some consumer financial products, the feasibility of these experiments is not likely a problem; the question is whether aggregate and summary statistics from this information would have characteristics of a public good.
- 19 The question of whether these rules and other forms of approval regulation are beneficial to society relative to their costs is of course an empirical proposition. Yet the model here shows that cost–benefit analyses of safety regulation that focus only upon “safety” benefits will generate systematically biased estimates of policy impact and will lead to invalid inferences (for one among many examples, see Peltzman 1976). In other words, it is only when cost–benefit analyses examine the confidence effects of regulation that such exercises can even approximate the validity that public policy demands.
- 20 This latter assumption is in part enabled by the large extent of the market, hence any minor deviation by a single consumer from derived equilibria would have vanishing weight in the regulators’ calculations.
- 21 Katz (2007) argues that the possible innovation benefits of regulation have “largely been absent from most cost–benefit analyses of drug regulation, yet without [them] any discussion of the merits of regulation is incomplete.”

## References

- Alt J, Calvert R, Humes B (1988) Reputation and Hegemonic Stability: A Game-theoretic Analysis. *American Political Science Review* 82, 445–466.
- Avorn J (2004) *Powerful Medicines: The Benefits, Risks, and Costs of Prescription Drugs*. Knopf, New York, NY.
- Banks JS, Sundaram R (1994) Switching Costs and the Gittins Index. *Econometrica* 62, 687–694.
- Billingsley P (1999) *Convergence of Probability Measures*, 2nd edn. Wiley, New York, NY.
- Boardman AE, Greenberg DH, Vining AR, Weimer DL (1996) *Cost–Benefit Analysis: Concepts and Practice*. Prentice Hall, Upper Saddle River, NJ.
- Bollinger B, Leslie P, Sorensen A (2010) Calorie Posting in Chain Restaurants. Working Paper, Stanford University, Stanford, CA.
- Calvert R (1987) Reputation and Legislative Leadership. *Public Choice* 55, 81–119.
- Candell LN (2010) What the Financial Sector Should Borrow: A Military Approach to Keeping the Economy Safe. *Harvard Business Review* 88(1), 46–48.
- Carpenter D (2004) Protection without Capture: Product Approval by a Politically Responsive, Learning Regulator. *American Political Science Review* 98(4), 613–631.
- Carpenter D (2010) *Reputation and Power: Organizational Image and Pharmaceutical Regulation at the FDA*. Princeton University Press, Princeton, NJ.
- Carpenter D, Ting MM (2007) Regulatory Errors with Endogenous Agendas. *American Journal of Political Science* 51(4), 835–853.
- Chernoff H (1968) Optimal Stochastic Control. *Sankhya – The Indian Journal of Statistics Series A* 30(3), 221–252.
- Dixit A, Pindyck R (1994) *Investment Under Uncertainty*. Princeton University Press, Princeton, NJ.
- Eisenberg R (2007) The Role of the FDA in Innovation Policy. *Michigan Telecommunications and Technology Law Review* 13(1), 345–388.
- Hamilton A (2001 [1769–1804]) *Writings*. Ed. Joanne Freeman, The Library of America, New York.
- Jin G, Leslie P (2003) The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards. *Quarterly Journal of Economics* 118(2), 409–451.

- Jovanovic B (October 1979) Job Matching and the Theory of Turnover. *Journal of Political Economy* 87(5), 972–990.
- Karatzas I (1984) Gittins Indices in the Dynamic Allocation Problem for Diffusion Processes. *The Annals of Probability* 12(1), 173–192.
- Karatzas I, Shreve S (1991) *Brownian Motion and Stochastic Calculus*, 2nd edn. Springer-Verlag, New York, NY.
- Katz A (2007) Pharmaceutical Lemons: Innovation and Regulation in the Drug Industry. *Michigan Telecommunications and Technology Law Review* 14(1), 1–41.
- Laffont JJ, Tirole J (1994) *A Theory of Incentives in Regulation and Procurement*. MIT Press, Cambridge, MA.
- Law M (2003) The Origins of State Pure Food Regulation. *Journal of Economic History* 63(4), 1103–1130.
- Law M (2006) How Do Regulators Regulate? Enforcement of the Pure Food and Drugs Act, 1907–38. *Journal of Law, Economics, and Organization* 22(2), 459–489.
- Law M, Kim S (2005) Specialization and Regulation: The Rise of Professionals and the Emergence of Occupational Licensing Regulation. *Journal of Economic History* 65(3), 723–756.
- Law M, Marks MS (2009) The Effects of Occupational Licensing Laws on Minorities: Evidence from the Progressive Era. *Journal of Law and Economics* 52(2), 351–366.
- Mandelbaum A (1987) Continuous Multi-armed Bandits and Multiparameter Processes. *The Annals of Probability* 15(4), 1527–1556.
- Manski C (2009) Adaptive Partial Drug Approval: A Policy Proposal. *The Economist's Voice* 6(4). [Last accessed 27 Sep 2010.] Available from URL: <http://www.bepress.com/ev/vol6/iss4/art9>.
- Peltzman S (1976) *Regulation of Pharmaceutical Innovation: The 1962 Amendments*. American Enterprise Institute, Washington, DC.
- Stiglitz J (2008) The Financial Crisis of 2007/2008 and Its Macroeconomic Consequences. Working Paper, Columbia University.
- Viscusi WK (1993) The Value of Risks to Life and Health. *Journal of Economic Literature* 31(4), 1912–1946.
- Warren E (2007) Unsafe at Any Rate. *Democracy: A Journal of Ideas*. Issue 5, Summer 2007.
- Warren E (2008) Product Safety Regulation as a Model for Financial Services Regulation. *The Journal of Consumer Affairs* 42(3), 452–460.

## Appendix

### Proof of preserved stochastic dominance relations in Lemma 2

Note first that the product and summation operators in the integral distributions  $G^{\zeta}$  preserve the continuity of countably additive functions of  $\pi_{\zeta}(x)$  [mixtures of distributions  $G$ ], and  $G^R$  and  $G^U$ . Take any four distributions drawn from the exponential family (which include the Beta and normal),  $G_1$ ,  $G_2$ ,  $G_3$ , and  $G_4$ . By additivity of independent variances, if  $G_1$  SOSD  $G_2$ , then for  $0 \leq \alpha \leq 1$ ,  $\alpha G_1 + (1-\alpha)G_2$  SOSD  $G_2 \equiv \alpha G_2 + (1-\alpha)G_2$  and  $\alpha G_1 + (1-\alpha)G_3$  SOSD  $G_2 + (1-\alpha)G_3$ . Then for the  $\zeta^{\text{th}}$  and  $(\zeta - 1)^{\text{th}}$  terms of any quadratic mixture as defined above, the following stochastic dominance relation holds:

$$[\zeta^{\text{th}} \text{ mix}]: \alpha_{\zeta}(\alpha_{\zeta-1}G_1 + (1-\alpha_{\zeta-1})G_3) + (1-\alpha_{\zeta})(\alpha_{\zeta-1}G_3 + (1-\alpha_{\zeta-1})G_4)$$

$$\text{SOSD } \alpha_{\zeta}(\alpha_{\zeta-1}G_2 + (1-\alpha_{\zeta-1})G_3) + (1-\alpha_{\zeta})(\alpha_{\zeta-1}G_3 + (1-\alpha_{\zeta-1})G_4).$$

Because  $\zeta$  is arbitrary,  $\sum_{\zeta} \{\pi_{\zeta}^R(G^R)\} / N_{\zeta}$  SOSD  $\sum_{\zeta} \{\pi_{\zeta}^U(G^U)\} / N_{\zeta}$ . But by Fatou's lemma,

$$\int_{\mathfrak{R}} \liminf_{\zeta \rightarrow \infty} \sum_{\zeta} \{ \pi_{\zeta}^R(G^R) \} / N_{\zeta} \leq \liminf_{\zeta \rightarrow \infty} \int_{\mathfrak{R}} \sum_{\zeta} \{ \pi_{\zeta}^U(G^U) \} / N_{\zeta}.$$

Because second-order stochastic dominance simply invokes expectations of the variance of  $G^{\Sigma,R}$  and  $G^{\Sigma,U}$ , preservation of SOSD relations falls out as a special case for all finitely valued  $G^{\Sigma}$ . The assumed square-integrability of  $G^{\Sigma}$  is sufficient, with Fatou's lemma, to preserve the finiteness of all expectations of  $G^{\Sigma}$ .