Regulating New Product Testing: the FDA vs. the Invisible Hand

Allan Collard-Wexler*, Matthew Grennan†, Andrew Steck‡

February 12, 2024

Abstract

Product testing plays an important role in the functioning of markets for innovative new products, where uncertainty exists regarding the product's safety, quality, or other attributes. In spite of this, an incomplete understanding of the economic trade-offs and their quantitative welfare implications has contributed to different products facing a wide range of regulatory regimes and private testing incentives (and even similar or identical products facing disparate regulations across geography and time). In this paper, we develop a dynamic model of innovation, testing, and competition between firms to examine the interplay between private incentives to test and regulatory requirements. We calibrate the model using data from the medical device sector and then consider the welfare implications of different regulatory regimes, unpacking the economic forces that drive them. Our results highlight that even in the absence of regulatory requirements, firms have substantial private incentives to conduct their own product tests, and accounting for these is critical to good policy. Optimal regulation weighs the treatment effect of inducing more testing with the selection effect that more testing requirements deter some products from entering the market. Our model also reveals a new and important dynamic "pruning" effect, whereby selection that excludes lower quality products increases market incentives for higher quality products. We also quantify the substantial inefficiency of ex-post quality regulation. By contrast, simple ex-ante minimum testing regulations perform surprisingly well in our calibrated model.

Keywords: product testing, innovation, entry, dynamic games, medical devices
JEL Classification:

*Duke University
†University of California, Berkeley
‡University of Toronto

We thank Bob Town and participants at the UofT IO Brownbag Workshop, EARIE 2023, NYU Stern Miami Micro Workshop, and the Tulane Healthcare Innovation Conference for helpful discussions. Any errors are our own.

1 Introduction

"In god we trust, everyone else must bring data."
- Robert Califf, M.D., US Commissioner of Food and Drugs 2016-17, 2022-current

Innovative new products typically come with some degree of risk that they may not work as intended. In the case of health care technologies such as biopharmaceuticals and medical devices, most countries have a regulatory body (such as the Food and Drug Administration (FDA) in the United States) that mandates safety and efficacy testing before products can be marketed to consumers. The testing required varies widely across the world and over time for similar products (Peltzman 1976; Grennan and Town 2020), as well as within the US for products in different risk classes (Rogers 2023).

This variation in policy stems from what seems to be an incomplete understanding of the full set of economic forces at work as well as a lack of quantitative evidence on more well known forces, such as the tension between the benefits of testing in terms of decreasing uncertainty and reducing risk about the performance of new products versus the costs of testing in terms of money and time. Nevertheless, the stakes can be high and debates can be contentious: For instance, in the debate over the FDA's rules for approving Covid-19 vaccines and diagnostic tests, most of the delay in getting vaccines to market was due to the testing requirements rather than development time.[1] Analyzing regulator policies is further complicated by the fact that firms will likely respond endogenously according to their own private incentives for testing, which in the presence of market power may diverge from the social welfare maximizing policy (Spence 1975).

In this paper, we develop a model in which firms generate new product ideas, invest in testing and development of the new products, and ultimately compete in the market for consumers who care about product quality as well as risk. This is a computational dynamic oligopoly model in the line of work from Pakes and McGuire (1994) or Mermelstein et al. (2020). We calibrate the model to data and estimates from prior research on a real world medical device market (Grennan and Town 2020) and consider policy-relevant variations of those parameters and modeling assumptions. In this environment, we examine the gap between firms' private incentives for testing compared to the social planner's preferred testing policy. We then consider the impact of different regulator policies on market outcomes and welfare in order to develop an understanding of the trade-offs faced by a regulator and how optimal policy making may change based on characteristics of the market.

[1] Famously, the Moderna RNA vaccine was developed in two weeks, but took a year to proceed through clinical trials.

Several quantitative and qualitative takeaways emerge from this analysis. First, in the absence of regulatory requirements, firms still have private incentives to test. Therefore, evaluations of testing policy should depend on the gap between private and public incentives, assessed at the margin. In the monopoly case, the private incentive to test will always be lower that the public value of testing as firms do not recover the entire benefit of testing from consumers in the form of higher profit. However, in the duopoly case, this may no longer be true as much of the impact of testing is to steal market share from the rival. This business stealing externality can give rise to more testing, which can be beneficial, or can even exceed the socially optimal level (Mankiw and Whinston 1986). Further, the welfare benefits of a testing policy may be relatively small, in absolute magnitude, as the products with the highest expected quality also have the greatest private incentives to test. In the parlance of clinical trials, the high expected quality product may be "always takers" in that they would have voluntarily satisfied a testing policy anyways. In our baseline model, we find that competition mostly exacerbates the under-testing problem as the decrease in residual market size dominates the business stealing incentive, and the gap between public and private incentives is large enough that social surplus under a laissez-faire policy is only about 63 percent that under the social planner.

Second, testing requirements change the selection of products in the market. There are fairly low returns to testing a product with low expected quality because the value of additional certainty to consumers is spread over a lower quantity of products sold. As such, low quality products will refrain from entering the market if the cost of testing they need to undertake is too high. In the monopoly case, this selection effect is unambiguously bad because it removes a product from the marketplace. However, in a market with more potential competitors, keeping a product out of the market might encourage future entry of better products—this "pruning" of the product space might be beneficial for future growth. Because it is the low quality products that are selected from the market, the selection effect tends to generate moderate losses to welfare, while the pruning effect can be substantial.

Third, examining the impact of alternative regulatory policies in our baseline model, we find that a simple ex-ante minimum testing requirement achieves 93 percent of the social planner welfare. The optimal minimum testing policy trades off lost entry from selection vs. gains from selection via pruning and the treatment effect of additional testing among the products that enter in equilibrium. We find that it is critical to set such a policy keeping in mind firms' endogenous incentives—a naive "quality control" model that only maximizes the treatment effect of testing without respect for selection only achieves 72 percent of the social planner welfare.

Fourth, we find that the simple minimum testing requirement actually substantially out-

performs an ex-post quality screen (which may more closely mirror FDA policy, at least for clinical trials). The ex-post quality screen only achieves 78 percent of the social planner optimum welfare. This is because an ex-post quality screen induces testing that is decreasing in the ex-ante expected quality of the product, making it difficult to induce meaningful additional testing by high quality products without also inducing too much selection. In our baseline model, the distribution of quality and the learning from pre-clinical testing are such that the ex-ante testing requirement performs better.

Finally, we trace out how differences in market characteristics can affect the performance of different policy regimes. We consider how market size and the degree of initial quality uncertainty affect the social planner's choice of a minimum testing threshold. Unsurprisingly, the larger the market, the higher a minimum threshold will be chosen, as the benefits of testing are larger in bigger markets as are firms profits. Initial quality uncertainty has a more nuanced effect on the regulator's choice. If there is very low uncertainty about a product's quality then testing is not needed. As uncertainty increases, the benefits of testing also increase, leading to higher testing thresholds. This leads to the simple conclusion that testing requirements should be stricter for products where there is considerable uncertainty versus products where there is little uncertainty. However, for products with very large uncertainty, e.g. early-stage experimental treatments, it is very difficult to get consumers to purchase the products in the first place. As such, firms have very low profits, and any testing mandate may lead to firms deciding not to enter the market. Thus, the regulator would choose a very light testing mandate for such "experimental" products.

Related Literature and Roadmap This paper builds on related work that studies the welfare consequences of regulating new product testing with uncertain quality by developing a complete dynamic model of product development, testing, entry, and competition and calibrating the model to real world data. Prior empirical work has examined changes in market outcomes for specific medical products using the the Keffavauer Harris Amendments of 1965 that increased pharmaceutical testing requirements in the US (Peltzman 1976), reclassifications that reduced testing requirements for certain medical devices in the US (Rogers 2023), and differences between US and EU testing requirements for the same devices (Grennan and Town 2020). We build on the Grennan and Town (2020) modeling and estimation of the market and welfare by adding the arrival of new product ideas and endogenous firm testing and entry decisions in a dynamic framework. This allows us to explicitly consider market outcomes under different regulatory regimes (laissez faire, social planner optimal, ex-ante testing requirements, ex-post quality requirements, quality control) and how policy tradeoffs change as product market features such as ex-ante risk and market size change.

In doing so, we are able to make new contributions by quantitatively unpacking some of the economic forces such as how different policies trade off the selection effect of excluding some products that exit due to testing costs versus the treatment effect of increased testing by the products that enter. We also uncover some economics that to our knowledge were previously unknown (or at least under-appreciated) such as the importance of pruning lower quality products to preserve entry and investment incentives for higher quality products and the inherent inefficiencies of ex-post quality regulation.

From a methodological perspective, our model is related to dynamic industry models with endogenous entry and innovation (see Pakes and McGuire (1994); Collard-Wexler (2013); Goettler and Gordon (2011); Igami (2017) and the survey in Aguirregabiria et al. (2021)). The importance of uncertainty and learning in demand builds on a large literature in industrial organization and marketing (see Ching et al. (2013) for a review). The fact that uncertainty is symmetric imperfect information to firms and consumers is similar to Jovanovic (1982), Gentzkow and Kamenica (2011), and Grennan and Town 2020. We argue this is appropriate for certain medical product markets where testing outcomes are verifiable and made available to consumers with expertise and motivation to be informed.

Section 2 discusses the institutional background for the FDA and in particular the coronary stent market to which our baseline model is calibrated. Section 3 starts by laying out the stage game with demand and pricing, the model for testing for both a monopolist, then for an duopolist. Section 4 describes the various policy regimes at the regulator's disposal, and assesses their performance in our calibrated model. A discussion of results follows in Section 5.

2 Background

Our empirical analysis and model details are grounded in the setting of coronary stents 2004-2013, a time when the EU had relatively low testing requirements and the US had more substantial testing requirements, providing the rare combination of data on the distribution of products that would enter in a low testing regime as well as the amount and market value of learning from additional testing. We draw the stage game of our model and baseline parameters from the Grennan and Town (2020) analysis of this setting and data.

The choice of which stent to use is mainly driven by interventional cardiologists who are a relatively small and informed group of expert users, the role of testing in this market is primarily about acquiring more information about the stents for all parties, rather than solving an asymmetric information problem between manufacturer and consumer. In keeping with this, our model of testing will feature symmetric information between all parties. We

further assume that a regulator such as the FDA is available to credibly validate and report any product testing, and as a result the market is one where consumers and firms both face (identical) imperfect information regarding product quality, conditional on any testing, even if that testing is done voluntarily by firms and not mandated by the regulator.[2] We calibrate the model to data and estimates of market primitives from Grennan and Town (2020).

Prices in the stent market are typically set by negotiations between manufacturers and hospitals or groups of hospitals. We follow Grennan (2013) in modeling price formation using Nash-in-Nash bargaining at the hospital level in our stage game.

Entry of new stents into the market depends on a combination of the arrival of new product ideas, testing of those new products, and the satisfying of any regulatory requirements. The stent market has seen the continuous arrival of innovative new products, and we calibrate this arrival process to the arrival rate of new products in the EU during the time of light regulation 2004-13. These products must then make their own decisions about whether to enter the market, given the costs of any testing suggested by their own private incentives as well as any regulatory requirements, and given the expected profits to be earned in the market.

3 Modeling Preliminaries

Before turning to the analysis of dynamics, we outline the stage game elements: demand for stents, competition and pricing, and the testing and learning process.

3.1 Demand

Demand for stent j for consumer i is given by a CARA utility function with a logit idiosyncratic shock term. The consumer's expected utility of consuming stent j can then be represented by the indirect utility function:

$$U_j = \underbrace{Q_j - \frac{\rho}{2}\sigma_j^2}_{\delta_j} + \epsilon_{ij}, \tag{1}$$

where ϵ is a patient-specific logit shock and δ_j is the mean across patients. The parameter ρ governs the risk aversion, the trade-off between the stent's perceived quality Q_j and the un-

[2]We think of such a regime as similar in spirit to financial market regulation. Faced with a similar set of economic issues in the financial sector, the SEC instead requires a certain level of transparency in reporting based on accounting principles and public company disclosures, but otherwise lets firms that create financial products make their own choices regarding how much the products have been "tested" before they can be sold to consumers in the marketplace. One could think of qualified professional investors as akin to specialist physicians making decisions about most medical treatments.

certainty over that perception σ_j^2. The outside option of not consuming a stent is normalized to zero. Note that this utility function assumes that consumers are price insensitive.[3] This insensitivity is not critical for our economic results, but aligns with previous work on demand estimation and welfare measurement for health care products like stents Grennan and Town (2020) and pharmaceuticals, where revealed preference demand estimates can depart from relevant social welfare measures.

The consumer surplus generated given by a choice set of stents $\mathcal{J}$ is the usual so-called inclusive value:

$$\mathcal{CS}(\mathcal{J}) = \frac{1}{\theta^{scale}} \log\left(\sum_{j\in\mathcal{J}} \exp(\delta_j)\right), \tag{2}$$

where we add a scaling factor $\frac{1}{\theta^{scale}}$ to convert consumer surplus measured in utility units into dollars.

With the normalization of the outside option's utility, the logit demand system means that a stent i operating in the market with the other stents in $\mathcal{J}$ has a market share s_i of:

$$s_i = \frac{\exp(\delta_i)}{1 + \sum_{j\in\mathcal{J}} \exp(\delta_j)}. \tag{3}$$

3.2 Testing and Information

At the moment of market entry, prior to sales and revenue, a firm can choose to conduct τ tests, where $\tau \in \{0, 1, \cdots, \bar{\tau}\}$. Each test has a known and constant cost C_τ. Conducting τ tests yields τ signals of the quality of the stent denoted by A, at a cost of τC_τ. The realized signal A is drawn from distribution $\mathcal{N}(Q, \sigma_A^2)$. This signal will lead to an update to the posterior distribution over quality given by the usual Bayesian updating equation with a change in mean estimate of quality:

$$Q' = Q \times \frac{\sigma_A^2}{\sigma_A^2 + \sigma_{prior}^2} + A \times \frac{\sigma_{prior}^2}{\sigma_A^2 + \sigma_{prior}^2}, \tag{4}$$

and the posterior variance is given by:

$$\sigma^2 = \frac{\sigma_{prior}^2}{\sigma_A^2 + \sigma_{prior}^2}.. \tag{5}$$

We restrict the model such that τ is chosen before the first signal is realized, and all τ

[3]Because of the structure of health insurance and delivery, patients and physicians lack of price sensitivity may not represent their true valuation of stents. Thus, we choose θ^{scale} to normalize the total surplus per stenting procedure to \$5,000, which is the approximate median of the estimated dollars in quality adjusted life years from the procedure relative to a coronary artery bypass graph surgery.

signals are realized prior to production, sales, and revenue.[4]

The posterior variance is a deterministic function of the number of tests. This allows us to summarize the ex-post state, the state after testing, by $x = \{Q, \tau\}$. Furthermore, the updating rule yields convexity in the cost of lowering the posterior variance σ^2, which (along with Q) is ultimately what consumers care about.

3.3 Pricing and Profits

Prices are set by negotiations between purchasers and manufacturers: these are given by a static Nash Equilibrium of Nash Bargaining models for each period, following the theory developed in Horn and Wolinsky (1988) and Collard-Wexler et al. (2019).

These approaches assume that the prices set maximize the bilateral Nash product with bargaining weight $\alpha \in [0, 1]$ on the manufacturer's surplus and weight $1 - \alpha$ on consumer surplus:

$$\max_{p_j} \left(Ms_j(p_j - mc_j)\right))^{\alpha} \left(CS(\mathcal{J}_t) - CS(\mathcal{J} \setminus \{j\})\right)^{1-\alpha} , \tag{6}$$

for each $j \in \mathcal{J}$. The term $CS\left(\mathcal{J}_t\right) - CS\left(\mathcal{J}_t \setminus \{j\}\right)$ is the so-called marginal contribution of stent j. In our application we set bargaining weights $\alpha = \frac{1}{2}$, for simplicity and to streamline exposition.

To determine profits per period, we assume that each stent j has a constant marginal cost of production mc_j. Thus stent j in period t earns profits $\pi_{jt} = Ms_{jt}(p_{jt} - mc_j)$, where M denotes the constant market size.

3.4 Entry and Testing Choice

In order to illustrate the economics of the setting, we begin our exposition of the dynamic entry and testing decision by considering the decision of a monopolist entrant who is guaranteed exclusive access to the market after entry.

To solve the model, a few other basic assumptions and pieces of notation are required. We assume each stent has a constant exogenous exit rate of $\chi \in [0, 1]$, that firms, consumers and the social planner all discount future payoffs at a constant rate of $\beta \in (0, 1)$, and (when we move beyond the monopolist's problem) that new stents arrive with a constant hazard of $\lambda \in [0, 1]$.

[4]Models where firms can elect to conduct further tests after the revelation of some information (or changes in competitive conditions) promise to be interesting extensions, though they greatly increase the difficulty of solving for equilibria.

3.4.1 Monopoly Choice

After the firm enters the market, with or without having conducted tests, it earns profits in each period it remains active (until it exits exogenously).

The monopolist's testing policy $\tau^*(x_0) \equiv \tau^*(Q_0, \sigma_0^2)$ can be characterized as the solution to the following value function:

$$\tau^*(x_0) = \underset{\tau}{arg\,max} \;\; W_\tau(x_0) \tag{7}$$

$$= \underset{\tau}{arg\,max} \;\; \beta \sum_{x'} V(x') P(x'|x, \tau) - \tau C_\tau \,, \tag{8}$$

where

$$V(x) = \pi(x) + \beta(1 - \chi) V(x) \,. \tag{9}$$

In Equation 9, $P(x'|x, \tau)$ is determined by equations (4) and (5), and t indexes time. The firm chooses τ to maximize $W_\tau(x_0)$, which we denote testing policy $\tau^*(x_0)$.

3.4.2 Oligopoly Choices

We now provide an overview of the oligopolistic setup. The only difference from the monopolistic case is that more than one firm can enter and serve the market at the same time, but this has ambiguous implications on each firm's optimal testing policy.

With each firm now considering the presence and choices of other firms when making its own testing and entry choice, the optimal testing choice $\tau^*(x_0)$ is affected in two ways. First, the presence of one or more rivals lowers market share s_j, which tends to decrease each firm's returns from testing. Second, there is now the possibility of business-stealing, so the marginal returns to testing may be *higher* when firm j is close to one or more rivals in δ space. The net effect of these two forces is of indeterminate sign in general, though in most states we simulate, the first effect dominates.

For the remainder of this paper, except where noted, we focus on the duopoly case. To be explicit, an entering firm i chooses $\tau^*(x_0)$:

$$\tau_i^*(x_{0i}) = \underset{\tau_i}{arg\,max} \;\; \beta E\left(V(x_i', x_j')|x_{0i}, x_j\right) - \tau_i C_\tau \,, \tag{10}$$

where

$$V(x_i, x_j) = \pi(x_i, x_j) + \beta E\left(V(x_i', x_j')|x_i, x_j, \tau^*\right) \,. \tag{11}$$

On a technical note: when solving for the equilibrium strategies, we restrict attention to symmetric pure-strategy equilibria.

4 Optimal Regulation

Our goal is to look at regulation of products where there is symmetric uncertainty about product quality. The first-best outcome (henceforth called the Social Planner solution), is for the social planner to specify a testing level τ at each point in the state space ($\tau(x)$), that is for each ex-ante quality level Q_0 and rival state x_{-i}.[5] This testing level will maximize the expected net present value of consumer and producer surplus conditional on the use of Markovian regulation, that is strategies which do not feature the ability to commit.[6]

The social planner's optimal policy – requiring a different amount of testing from each product, depending on the state of the product and of the market (and without respect for firms' positive expected profit constraint) – is likely infeasible and almost certainly does not describe regulator policies observed in the real world. This motivates us to explore a number of potential policies that correspond to those we observe or that would seem feasible in practice. In addition to their interest per se, examining these policies also helps to clarify some of the economic tensions faced by real world regulators of new product testing when firms respond endogenously to that regulation.

The main policy we consider is the minimum testing policy that specifies that a firm must test its product at least $\underline{\tau}$ periods. There are of course alternatives. For instance, the FDA could regulate the quality of products sold in the market, which we call the ex-post product quality given by δ_j. In addition, to illustrate some of the tensions in setting a minimum testing policy, we also consider the case of requiring a minimum ex-ante quality level Q_0.

Finally, we consider what a regulator would choose if she had a misspecified model of firm behavior. Specifically, suppose a regulator was comparing the value of a stent testing for τ periods versus no testing at all, that is $\tau = 0$. This would ignore both the endogenous testing incentives of firms, in that firms may choose higher levels of testing than zero, as well as the entry choices of firms faced with a testing requirement. This misspecified model, which we call the quality control model, corresponds to the case where a regulator ignores firm responses, and seems to line up with much of the rhetoric in policy discussions surrounding the FDA and similar regulators.

4.1 Public and Private Incentive

Before diving into the optimal regulation problem we first look at the reasons for a divergence between private and public incentives for testing. It is this wedge that forms the basis for

[5] All stents start with identical σ_0^2, so the only heterogeneity in x_0 comes from Q_0.

[6] See Mermelstein et al. (2020) for more discussion on Markovian regulation in the context of merger policy.

regulation to potentially outperform "laissez faire" policy without any regulator intervention (other than to act as a certifying body for firm clinical trial results).

As is common in models of monopoly Spence (1975), the seller will not capture the entire social surplus of the value of its products. This creates a gap between the social incentives for testing and the private incentives for testing.

The planner's per period welfare function (after sunk testing costs) TS given a set of stents $\mathcal{J}$ is given by the difference between consumer surplus and production costs:

$$TS(\mathcal{J}) = CS(\mathcal{J}) - \sum_{j \in \mathcal{J}} Ms_j mc_j \tag{12}$$

The social planner has a choice specific value function defined similarly to 7, but replacing firm profits with the TS metric and accounting for the value from future entry. Note that in the bargaining solution we are using with $\alpha = \frac{1}{2}$, the manufacturer receives one half of the surplus, and thus, half of any incremental surplus from testing one more period. As such, this will lead the manufacturer, at least for the monopoly case, to have under-powered incentives to test.

Figure 1 shows testing choices as a function of initial quality draw Q_0: the private choice (in blue), social planner (in red), and minimum testing (in gold) policies. In panel 1a, we show the case where there is no other firm in the market, while in panel 1b there is one competitor in the market who has initial quality $Q_0 = 3.5$ and has tested for ten periods (a high quality competitor).

Starting with the private policy in panel 1a, firms have greater incentives to test with higher initial quality. This is rationalized by the fact that a decrease in the after-testing uncertainty, σ_j^2, has a multiplicative effect on expected profits in that both this uncertainty and the expected quality level, Q_j, determine both the market share and price. Thus the marginal returns to testing are higher for high expected quality products. The social planner's testing choice mirrors the private one in that better products are tested more. However, the social planner has strictly higher returns to testing than a private firm due to the fact that firms do not capture the entire increase in consumer surplus to testing. This is an ubiquitous feature of models of regulation with market power.[7]

The fact that social returns to testing are greater than private returns is key for the possibility of a welfare-enhancing intervention by the social planner to mandate testing.

The social returns to testing are strictly greater in the case of no-competition compared to the case of competition. For the private return to testing, the effect of competition is

[7]Notice that the number of periods of testing is always below $\bar{\tau}$ (the upper bound in the state space on the number of tests). This is due to the incremental decrease in variance from testing falls with more tests.

more nuanced. However, in virtually all the simulations we run, private testing falls with the quality of the rival product.[8]

The fact that it is more difficult to incentivize testing when there is competition in the market foreshadows our discussion of the social planner having incentives to exclude certain low quality stents from the market.

4.2 Minimum Testing

The gold line in both panels of Figure 1 discussed in the previous section shows the testing choices of firms if the regulator sets the (optimal) minimum testing threshold of $\underline{\tau} = 12$. Given that the private choices of testing are always below 12, in this specific case, firms will never choose to test above this minimum threshold. In addition, firms may choose not to test and thus not enter the market. Thus, a minimum testing rule will select some of the possible stents out of the market.[9]

In general, the regulator chooses an optimal minimum testing policy $\underline{\tau}$ to maximize:[10]

$$V^{SP}(x) = TS(x) + \beta \sum_{x'} V^{SP}(x') P(x'|x, \underline{\tau}) \,.$$

In the monopoly case, we can fill in some details on the transition probabilities:

$$\begin{aligned}
P(x'|\emptyset) &= \lambda \sum_{Q} 1\{W_{\underline{\tau}}(\tau_Q^*) > 0\} P(x'|Q, \tau_Q^*) P(Q) \\
P(\emptyset|\emptyset) &= (1-\lambda) + \lambda \sum_{Q} 1\{W_{\underline{\tau}}(\tau_Q^*) \leq 0\} P(Q) \\
P(\emptyset|x) &= \chi P(\emptyset|\emptyset) \\
P(x'|x) &= (1-\chi) 1\{x' == x\}
\end{aligned}$$

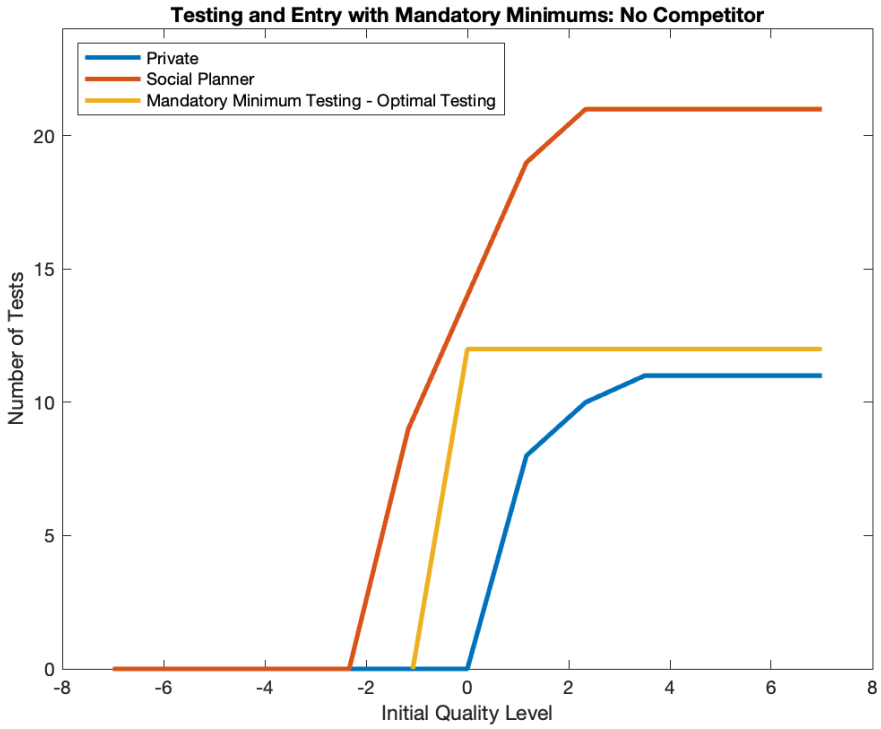
In these equations $\emptyset$ denotes the absence of a firm from the market.

The first equation states that the probability of a state x' in the next period is a function

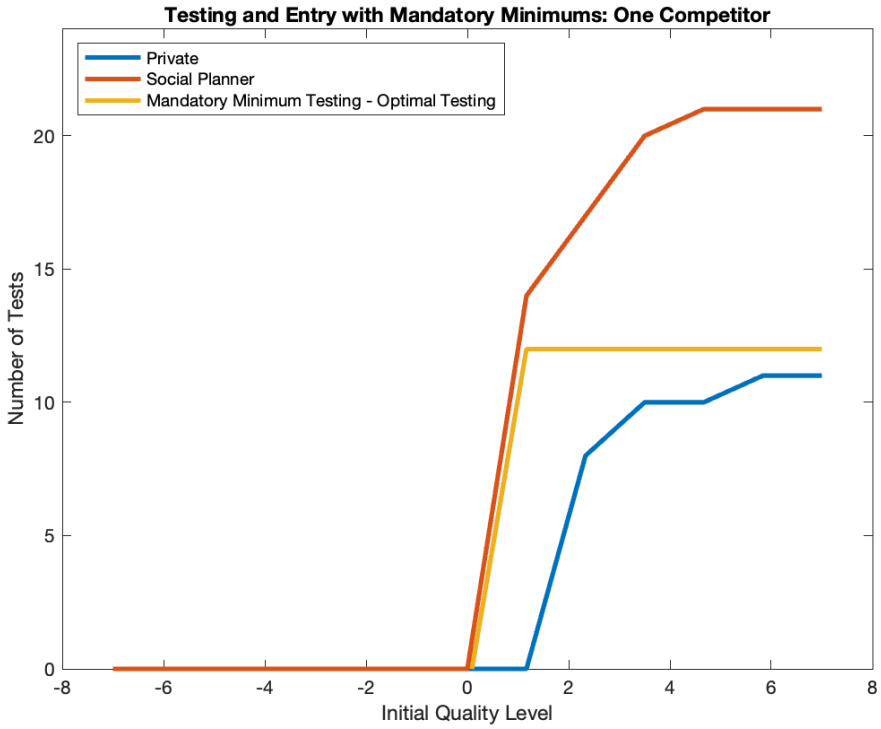
[8]Note that in a duopoly setting, testing need not be decreasing in the quality of the rival. Indeed, in the case where products are undifferentiated, that is $\epsilon = 0$ in our demand system, we would revert to homogeneous Bertrand competition, and the return to out-testing a rival would be quite high. This possibility result, that testing could be greater in duopoly than monopoly is confirmed by our simulations for very specific parts of the state space.

[9]Given that our model does not feature any fixed costs other than the costs of testing, firms always find it profitable to enter if they will not test.

[10]Notice that we assume the regulator makes a choice to maximize welfare from the perspective of a market which has no incumbent products, or alternatively choose a policy before firms show up. An alternative would be to to maximize welfare weighted by the stationary distribution, such as is done in Pakes and McGuire (1994). Given this paper is interested in new products, we think it is more appropriate to think of the discussion of regulation happening before these products arrive.



(a) Testing Policy with no competitors



(b) Testing Policy with one competitors ($Q_0 = 3.5$, $\tau = 10$)

Figure 1: Initial Quality Level and Testing Policy

of the arrival rate λ, the testing policy τ_Q^*, and the probability of testing moving to a new information state given by $P(x'|Q,\tau)$. The indicator function that the choice specific value function be positive is where the selection effect arises.

The second equation states that the probability of an empty market remaining empty is the sum of the probabilities of no new firm arriving, and a new firm arriving and choosing not to enter because the testing requirement is too costly in expectation.

The third and fourth equations specify the probability of a market moving from served by a firm to empty, and the probability of a non-empty market remaining non-empty.

We find it useful to think about regulator policies in terms of their *selection* and *treatment* effects. By *selection*, we refer to the fact that fulfilling regulatory requirements will often make entry unprofitable in expectation for some firms, and those firms will endogenously select out of the market. By *treatment*, we refer to the effect of regulatory requirements on how much testing firms actually conduct. Importantly, the baseline for product testing is not typically zero – firms have incentives of their own to test – and thus the treatment effect is the additional testing induced by the regulator.

This sets up an interesting balance in regulator's choice of an optimal testing policy. On the one hand, using a minimum testing requirement to get the private testing choices closer to the socially optimally one is valuable. However, because the regulator cannot condition the testing policy on the quality of the product, the policy will have the greatest treatment effect on low quality products that would choose not to test much absent the regulation. Moreover, the minimum testing policy would induce the exit (non-entry) of products below some quality threshold, but these products were creating the least value to begin with, so this loss in product variety may tend to be small.

Figure 2 decomposes these treatment and selection effect on total surplus (vertical axis) as we change the minimum periods of testing required (horizontal axis). The blue line shows the total welfare, while the red line shows welfare if only the policy's selection effect is considered, and the gold line shows welfare if only the treatment effect is considered.

[**Andrew: dropping in decomposition equations for now**]

$$\tau^{selection} = \tau^{private} 1\{\tau^{mintest} > 0\} + \tau^{mintest} 1\{\tau^{mintest} \leq 0\}$$
$$\tau^{treatment} = \tau^{mintest} 1\{\tau^{mintest} > 0\} + \tau^{private} 1\{\tau^{mintest} \leq 0\}$$

The minimum testing requirement of 0 corresponds to laissez-faire policy. As the minimum testing required increases, the selection effect becomes positive as some of the lowest quality products select not to enter, but the treatment effect remains zero because the firms that do enter have private incentives to test more than the minimum requirement. As the

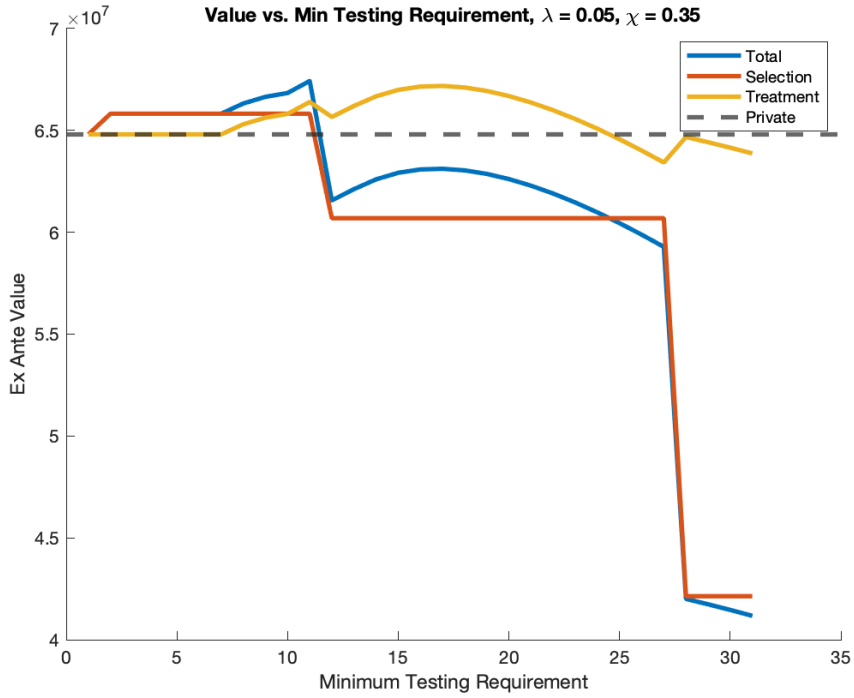


Figure 2: Treatment and Selection Effects of Minimum Testing

minimum testing required increases, the treatment effect becomes positive and continues to increase as some products are induced to extra testing by the policy. Eventually, the minimum policy reaches a level such that sufficiently high quality firms choose not to enter, turning the selection effect negative. The optimal policy is at 12 periods of testing, just before that drop in the selection effect.

Note that the treatment effect is maximized at 16 periods of testing, which is greater than the optimal. Thus considering the selection effect restrains policy relative to a policy that considers treatment alone.

Notice as well that the selection effect is positive for small values of testing requirement, as it is above the laissez-faire value. This reflects the fact that the regulator has an indirect "pruning" incentive to use this selection effect – it reserves space in the market for better products and preserves testing incentives for the products that do enter.

4.3 Alternative Policies: Ex-Ante and Ex-Post Quality Regulation

Regulating the the number of periods of testing is not the only way to regulate testing. Instead, one could target the quality of a product after testing. We call this "ex-post quality regulation" that operates on the mean utility δ_j = term. This relates to actual FDA policy

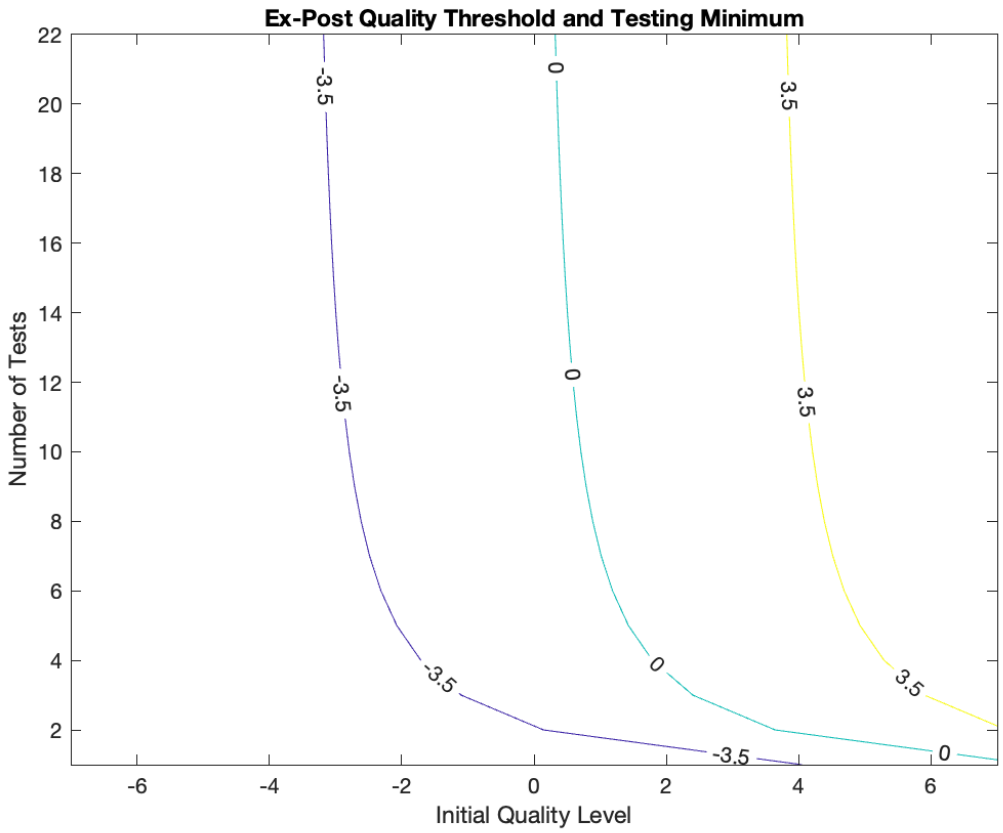


Figure 3: Number of Tests Required to Pass Ex-Post Screen

that requires safety and efficacy above some threshold for product entry.

Figure 3 shows how regulation on the ex-post quality operates in our model. On the horizontal axis we plot the product's initial mean expected quality level Q_O. On the vertical axis, we plot the number of tests required in expectation to achieve a given ex-post quality level $\delta = Q - \frac{\rho}{2}\sigma^2$, and we plot indifference curves for ex-post quality levels of -3.5, 0, and 3.5. What is important to notice is that an ex-post standard is more lenient on a high initial quality product than a low initial quality product—a product with $Q_0 = 1$ would need to test 10 periods to achieve a $\delta = 0$, while a product with $Q_0 = 2$ would only need 4 periods of testing. Given that the social planner would like high-initial quality product to test *more* than low initial-quality ones, an ex-post standard targets additional testing at the wrong products. Thus, an ex-post standard will do worse than an minimum testing threshold than has a uniform requirement on all the products in the market.

Another alternative is to have products regulated by initial product quality Q_0. This is equivalent to a minimum testing policy that can only impose a selection effect (zero treatment effect on testing). As such, we would expect this ex-ante quality regulation to perform worse

than a minimum testing policy.

4.4 Quality Control Model: Regulation without considering firm responses

So far we have looked at the choices of a regulator that correctly understands firm responses to the regulation. However much of the policy debate on testing revolves around a more economically naive model of firm behavior, one which assumes that firms choose testing at the regulated level and would not test at all absent the regulation. We call this the "quality control" model. The choices of a regulator that uses this misspecified model of firm behavior are useful to assess the value of the testing model proposed in this paper.

4.5 Comparing Policies

Table 1 presents the welfare and market outcomes of the different regulatory regimes. We consider the social planner solution, the optimal minimum testing threshold, laissez-faire, ex-ante and ex-post quality regulation, and the choices of the regulator using the quality control model of firm behavior. For each of these policies we compute the expected social surplus, and the component of this surplus going to producers and consumers, as well as expenditures on testing. We also show the expected number of products and product quality in the market, as well as a few statistics about the testing choices of firms.

Table 1: Welfare and Testing Regulation

	Social Planner	Quality Control	Optimal Minimum Threshold	Laissez Faire	Quality Screen	Ex-Post Quality Screen
Regulator Policies						
Minimum Percentile Quality with > 0 Testing	0.32	0.0	0.0			
Tests by Median Quality Product	13.6	15.0	12.0			
Tests by 90th Percentile Quality Product	18.8	15.0	12.0			
Market Outcomes						
Minimum Percentile Quality with > 0 Testing	0.32	0.93	0.68	0.68	0.68	0.68
Tests by Median Quality Product	13.6	0.0	12.0	0.3	4.5	4.5
Tests by 90th Percentile Quality Product	18.8	15.0	12.0	7.8	7.8	7.8
Expected Social Surplus	2.95e8	2.131e8	2.75e8	1.87e8	2.346e8	2.293e8
Expected Producer Surplus	2.51e7	3.601e7	3.86e7	3.427e7	4.129e7	3.919e7
Expected Consumer Surplus	2.12e8	1.585e8	2.021e8	1.398e8	1.749e8	1.706e8
Firm Expected Quality (percentile)	-0.3236	1.484	0.7074	-0.4001	0.6671	-0.5383
Social Surplus Value of Testing	3.522e8	2.307e8	3.081e8	1.988e8	2.518e8	2.476e8
Cost of Testing Performed	-5.814e6	-1.862e6	-3.428e6	-1.292e6	-1.841e6	-1.951e6
Expected Products in Market	0.2614	0.1512	0.2086	0.2614	0.2112	0.2614

First, notice that the optimal minimum threshold achieves 93 percent of the first-best surplus achieved by the social planner. In contrast, laissez-faire only achieves 63 percent of the first-best surplus. Thus, there are large returns to regulating product testing in this calibrated model. Looking at an ex-ante quality screen, which only operates on the selection

margin, shows that this achieves 79 percent of the first-best surplus, and thus both the treatment and selection effects of the minimum testing policy are important for welfare. An ex-post quality screen does slightly worse than an ex-ante one, and far worse than minimum testing, achieving 77 percent of first-best surplus.

The quality control model leads to the regulator choosing a policy which is too strict with a minimum threshold of 15, versus 12 for the regulator with correct beliefs on firm behavior. This leads to fewer expected products in the market in each period, 0.15 instead of 0.20, and lower welfare as well, achieving only 73 percent of first-best surplus.

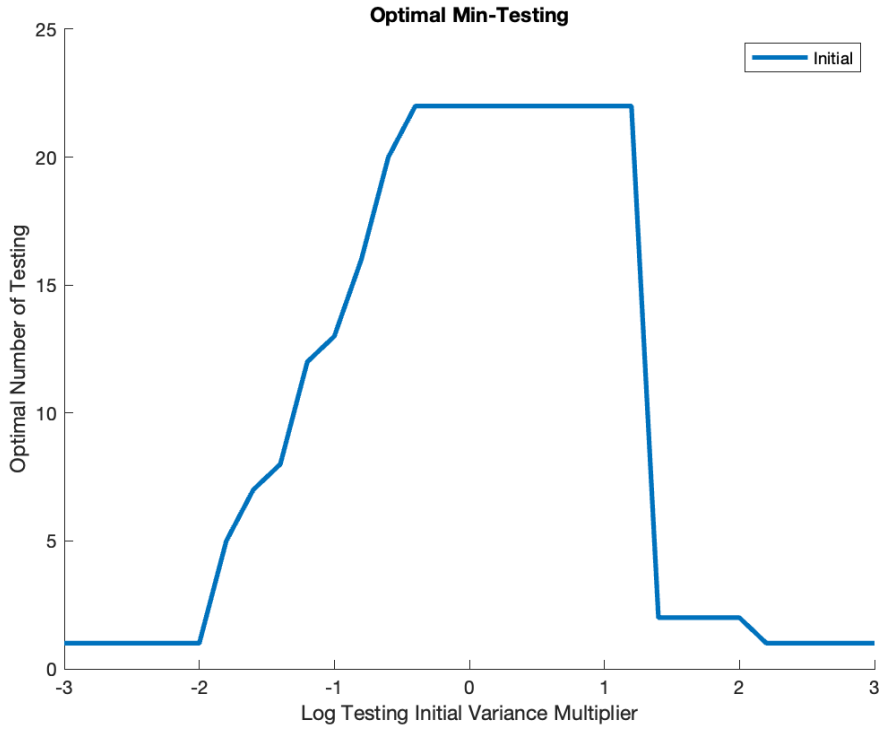
4.6 Regulations tailored to market characteristics

So far we have analyzed the choice of testing regulation in the context a single market. However, the FDA, as well as their equivalent in other countries, need to devise testing regulation for a variety of different product types with different underlying characteristics. Even if these markets have a similar symmetric learning process as the one we have tailored for the market for stents, they will not share the same parameters.

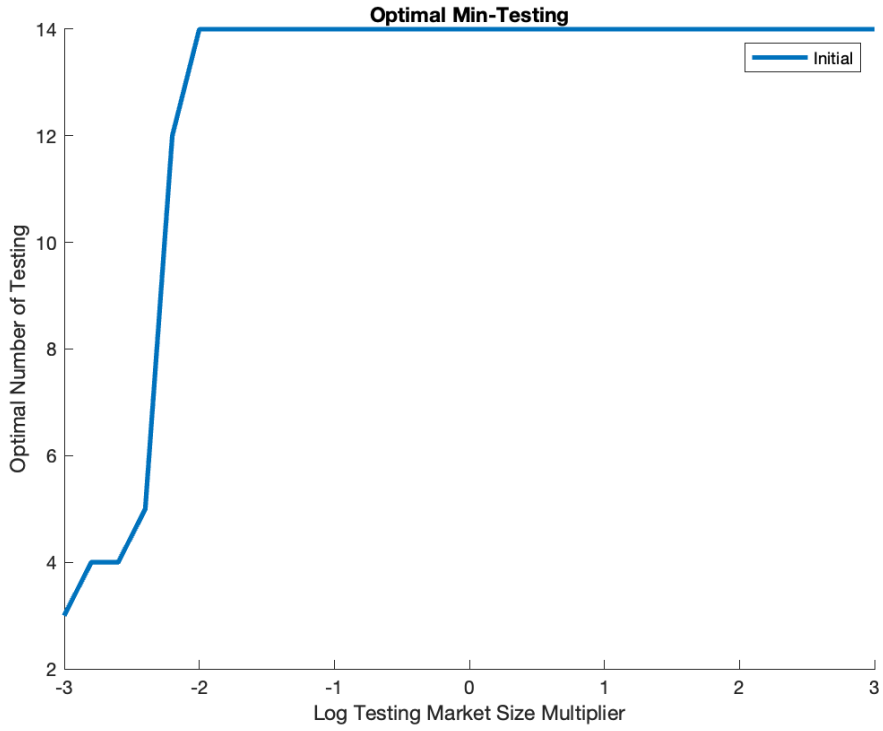
Figure 4 show how the optimal minimum testing policy would change as we either change market size (Figure 4b) or the variation of ex-ante uncertainty over product quality (Figure 4a). Notice that the horizontal axis is a term that scales market size and initial variance by between 0.04 ($\log(-3)$) and 20 ($\log(3)$), with zero ($\log(0) = 1$) indicating the value used in the rest of the paper.

Unsurprisingly, Figure 4b finds that the regulator would choose a more stringent testing minimum in larger markets. This is not surprising as both the social value of testing and the private value of testing are larger in bigger markets. As a result, the treatment value of additional testing is greater, and the selection effect takes longer to turn negative as the larger market justifies more testing.

Figure 4a shows the more nuanced effect of ex-ante uncertainty on optimal minimum testing policy. If there is no uncertainty on product quality, there is no need to test to begin with, and thus for low values of uncertainty, the regulator does not require testing at all. As uncertainty increases, so does the value of testing. However, at a high level of ex-ante uncertainty, then the mean utility $\delta_j = Q_j - \frac{\rho}{2}\sigma_j^2$ will be quite low. If consumers do not want to consume the product, then the private and social returns to testing will be very small. Thus, the regulator will also choose very low testing requirements at the top end of the distribution of uncertainty.



(a) Initial Variance of Quality



(b) Initial Variance of Quality

Figure 4: Optimal Policy with Different Market Characteristics

5 Conclusion

In this paper we show that a minimum testing rule outperforms other regulations for testing new products, such as having screens based on ex-post quality, achieving 93 percent of the first-best welfare. Having the correct model of endogenous firm behavior is critical for this exercise, as a model where the regulator believes they can tell firms exactly how much to test without considering either exit or firms' private incentives to test absent the policy leads to poor outcomes, achieving only 78 percent of the first-best. Related, ex-post quality regulation induces too much testing by low quality products and too little testing by high quality products, making it an inferior to a flat minimum testing policy in our calibrated model.

More generally, this shows the value of having a fully formed economic model when designing regulations of new product testing. Indeed, some commonly raised claims that the FDA has an overly strict attitude to product testing are difficult to assess in the absence of a model.

We also note that there are policies that we have not considered, such as having the regulator also set prices, similar to what is done by the UK's NICE, It would be interesting to see how a combination of regulatory levers may improve outcomes in future research.

Finally, some product categories may have more noise in choice process, such as individuals making errors in terms of which products they choose, say due to incorrect perceptions about product quality. These types of markets increase the pruning incentives of regulators because keeping bad products out of market helps consumers make better choices. We think this is a fruitful avenue for future research that will complement the symmetric information model we explore here.

References

Aguirregabiria, Victor, Allan Collard-Wexler, and Stephen P Ryan. 2021. "Dynamic games in empirical industrial organization." In *Handbook of Industrial Organization*, Volume 4. 225–343, Elsevier.

Ching, Andrew, Tulin Erdem, and Michael Keane. 2013. "Learning Models: An Assessment of Progress, Challenges, and New Developments." *Marketing Science* 32 (6): 913–938.

Collard-Wexler, Allan. 2013. "Demand fluctuations in the ready-mix concrete industry." *Econometrica* 81 (3): 1003–1037.

Collard-Wexler, Allan, Gautam Gowrisankaran, and Robin S Lee. 2019. ""Nash-in-Nash" bargaining: a microfoundation for applied work." *Journal of Political Economy* 127 (1): 163–195.

Gentzkow, Matthew, and Emir Kamenica. 2011. "Bayesian Persuasion." *American Economic Review* 101 (6): 2590–2615.

Goettler, Ronald L, and Brett R Gordon. 2011. "Does AMD spur Intel to innovate more?" *Journal of Political Economy* 119 (6): 1141–1200.

Grennan, Matthew. 2013. "Price Discrimination and Bargaining: Empirical Evidence from Medical Devices." *American Economic Review* 103 (1): 145–177.

Grennan, Matthew, and Robert J Town. 2020. "Regulating innovation with uncertain quality: information, risk, and access in medical devices." *American Economic Review* 110 (1): 120–161.

Horn, Henrick, and Asher Wolinsky. 1988. "Bilateral monopolies and incentives for merger." *The RAND Journal of Economics* 408–419.

Igami, Mitsuru. 2017. "Estimating the innovator's dilemma: Structural analysis of creative destruction in the hard disk drive industry, 1981–1998." *Journal of Political Economy* 125 (3): 798–847.

Jovanovic, Boyan. 1982. "Selection and the Evolution of Industry." *Econometrica* 50 (3): 649–670.

Mankiw, N Gregory, and Michael D Whinston. 1986. "Free entry and social inefficiency." *The RAND Journal of Economics* 48–58.

Mermelstein, Ben, Volker Nocke, Mark A Satterthwaite, and Michael D Whinston. 2020. "Internal versus external growth in industries with scale economies: A computational model of optimal merger policy." *Journal of Political Economy* 128 (1): 301–341.

Pakes, Ariel, and Paul McGuire. 1994. "Computing Markov-perfect Nash equilibria: Numerical implications of a dynamic differentiated product model." *The Rand Journal of Economics* 25 (4): 555.

Peltzman, Sam. 1976. "Toward a more general theory of regulation." *The Journal of Law and Economics* 19 (2): 211–240.

Rogers, Parker. 2023. "Regulating the Innovators: Approval Costs and Innovation in Medical Technologies." *Working Paper*.

Spence, A Michael. 1975. "Monopoly, quality, and regulation." *The Bell Journal of Economics* 417–429.

Appendices

A Parameters

In this section we detail the parameters used in the simulations and discuss their interpretation; many are drawn directly from estimates in Grennan and Town (2020).

We begin by discussing the primitives of the demand system. For reference, we reproduce Equation 1 here:

$$U_j = \underbrace{\mu_j - \frac{\rho}{2}\sigma_j^2}_{\delta_j} + \epsilon_{ij} \, .$$

B Model

Two core elements of the economic model:

1. Does firm decide to test: how much? (or just do/don't?)

    - Firm definitely needs a draw of its quality before deciding to test, but this can be publicly known (similar to the paper with Bob, we would assume there is still basic safety testing and reporting, and we are talking more about phase three style study to try to "prove effectiveness")

2. Implications of level of testing for public confidence / adoption (and safety / errors outcomes)

Further stuff:

1. Might also consider **how pricing takes level of evidence into account**. This is how Aduhelm adoption has been slowed in that insurance doesn't want to cover it (at a high price manufacturer wants to charge). This seems realistic, and prices should matter in economics. But it seems like this can muddy the waters a bit in that we don't want insurer to be sort of "backdoor regulator" here. This seems worth discussing.

2. **Implications for investment and advancement of technology.** Again with the Aduhelm example, one of the rationales was that it would help encourage further innovation if investors thought the bar to get to market might be lower. This also seems worth discussing. We may have a lot going on even with this part kept exogenous.

3. Does it matter for economics that everything condensed into a **single vertical "quality" index vs. more horizontal elements**?

    - If effectiveness and side effects are vertical, doesn't seem to be any loss.
    - If either effectiveness and/or side effects are horizontal – and it is possible to learn something about horizontal match – that would be different for demand and info modeling. Though still not clear how much actual economics are different?

C Model Testing

C.1 Monopoly Entrant

There is a pattern emerging here of how testing behavior is driven by a combination of marginal incentives and level effects. [Are there clear policy implications of this?]

[Would it be more illuminating to plot the benefits and costs of testing, with number of tests on the horizontal axis, for different quality levels (and parameters)? Would that help better illustrate what is going on under the hood? I think I need some intermediate step to more clearly see why these comparative statics in optimal testing policy work the way they do.]

C.2 Duopoly Entrants

Here think about how we can validate the model and also start to unpack any new effects from competition.

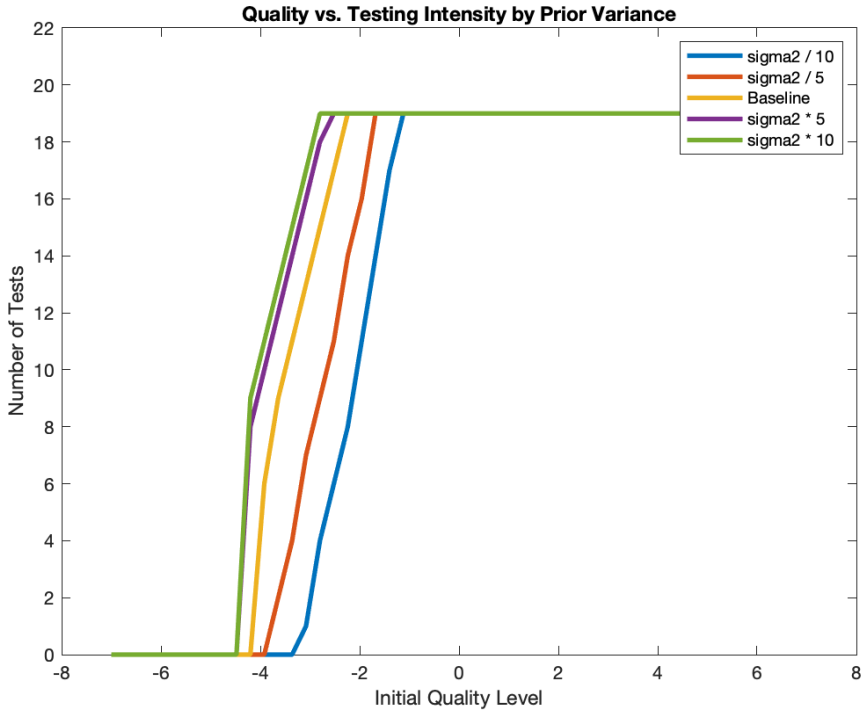


Figure A1: Figure A1

Notes: As the prior variance – the level of uncertainty faced by all new products – increases, products of a given initial quality estimate will choose to test more. This happens in a stochastic dominance type way. [However, it seems that there may be a cutoff initial quality level such that no product with quality below that will conduct tests, and increasing the prior variance asymptotes to the case where this cutoff would separate types that would do no testing vs types that would do the maximum amount of testing (for any given maximum amount of testing).]

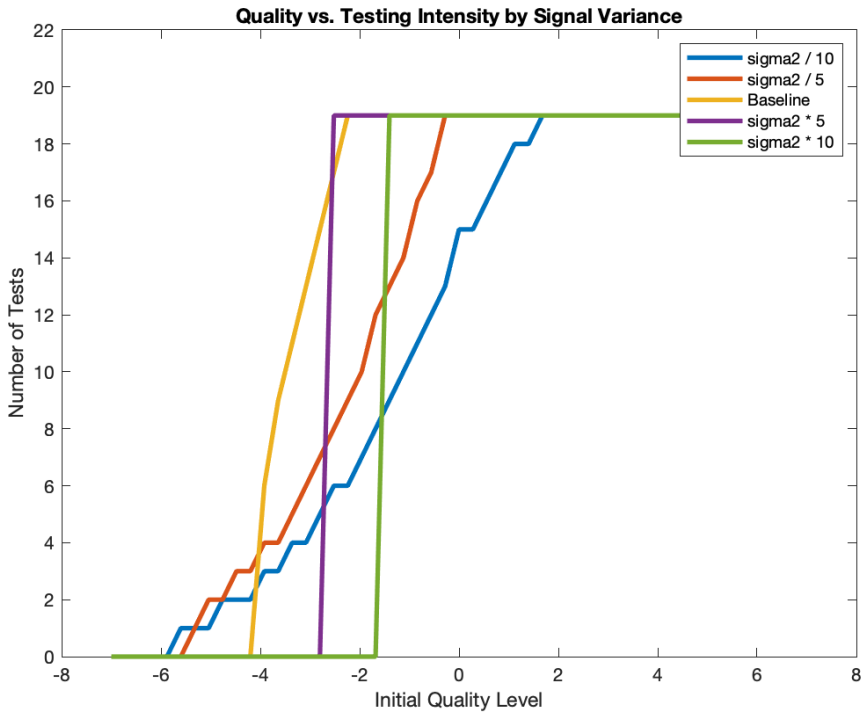


Figure A2: Figure A2

Notes: As the signal variance – the amount of noise in a test – decreases, this rotates the testing policy with respect to initial quality estimates. The testing policy becomes less "steep" – whereas it moves rapidly from no testing to a high level of testing when signal variance is large, the increase in testing with quality is more gradual when signal variance is low. Relatedly, as signal variance decreases, products of lower quality begin to do *some* testing with higher probability. [As the amount of noise in a test increases, it appears that this too may asymptote to a cutoff initial quality level that tests (again for a given max amount of testing).]

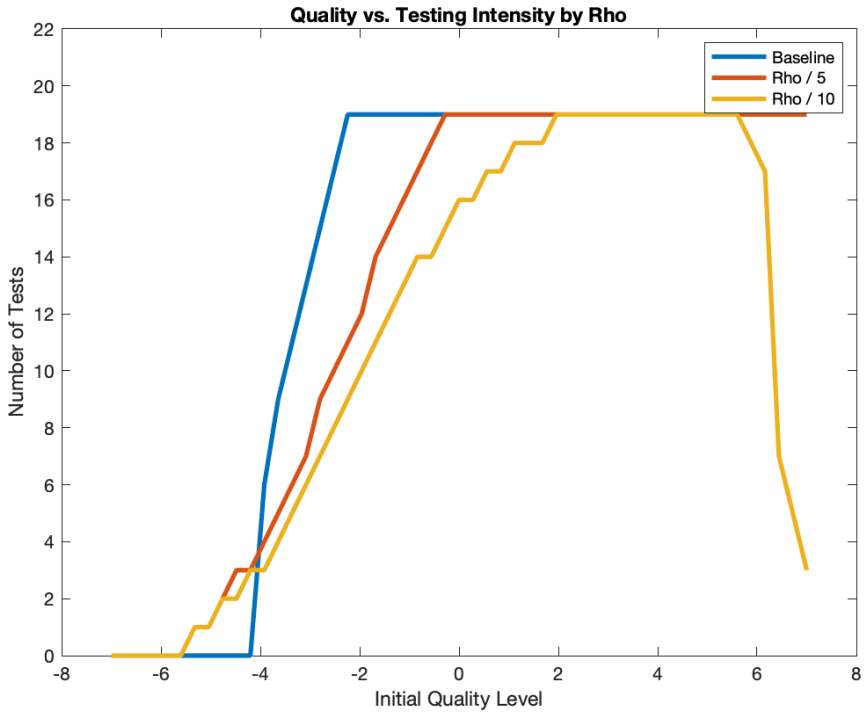


Figure A3: Figure A3

Notes: As the coefficient of risk aversion – the extent to which consumers dislike uncertainty about product quality – decreases, this rotates the testing policy with respect to initial quality estimates, similar in nature to a decrease in signal variance. However, this moves in the opposite direction of the rotation that occurs when signal variance increases. This is because while a decrease in ρ does decrease the marginal return to a test (similar to a signal variance increase), it also decreases the overall level of the uncertainty effect. Thus, at any level of uncertainty (in particular at the prior variance), lower risk aversion means that uncertainty has not moved quality as far from zero (and the steep part of the logit demand curve).

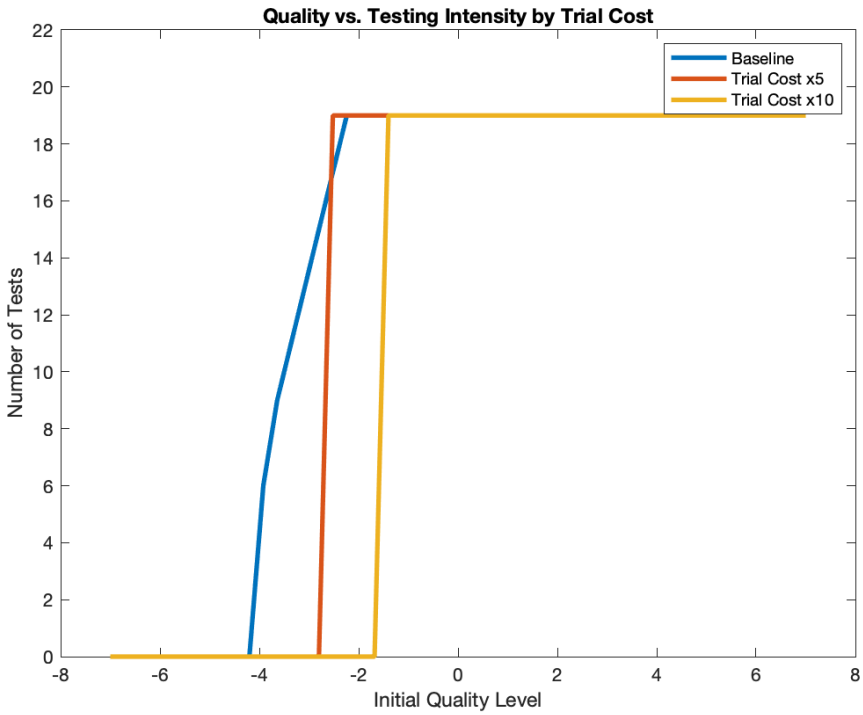


Figure A4: Figure A4

Notes: As the trial costs – cost per period of testing – increases, the testing policy rotates and shifts so that: (1) the initial quality level at which any testing occurs increases; (2) the policy shift becomes very steep, with any testing quickly becoming max testing.

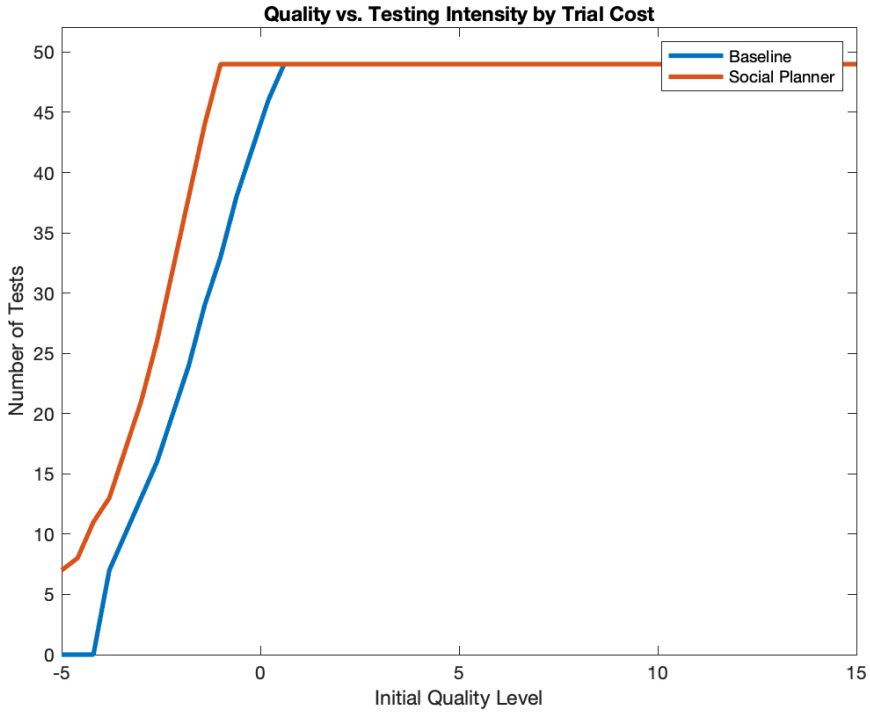


Figure A5: Figure A5

Notes: The social planner always prefers to test more than a monopolist firm. This is because of the standard result that the monopolist captures only a fraction of the social surplus created from testing. At a cost per test of $1.6M, the social planner will test to some degree for all of the quality levels shown here, whereas the monopolist does not begin testing until a sufficient quality level, and tests less at any quality level. [The max testing makes it hard to tell, but does the gap narrow as initial quality increases?]

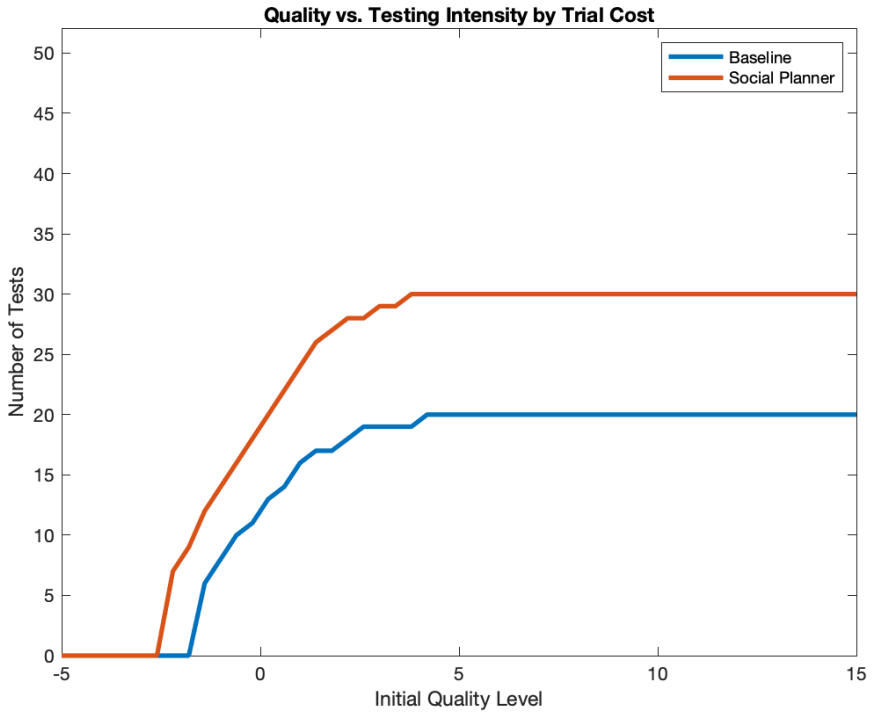


Figure A6: Figure A6

Notes: With (an order of magnitude) larger testing costs of $16M per test, we see less testing overall. We also see that the gap between monopoly and the social planner increases with quality. [Why do both asymptote to a testing level?]

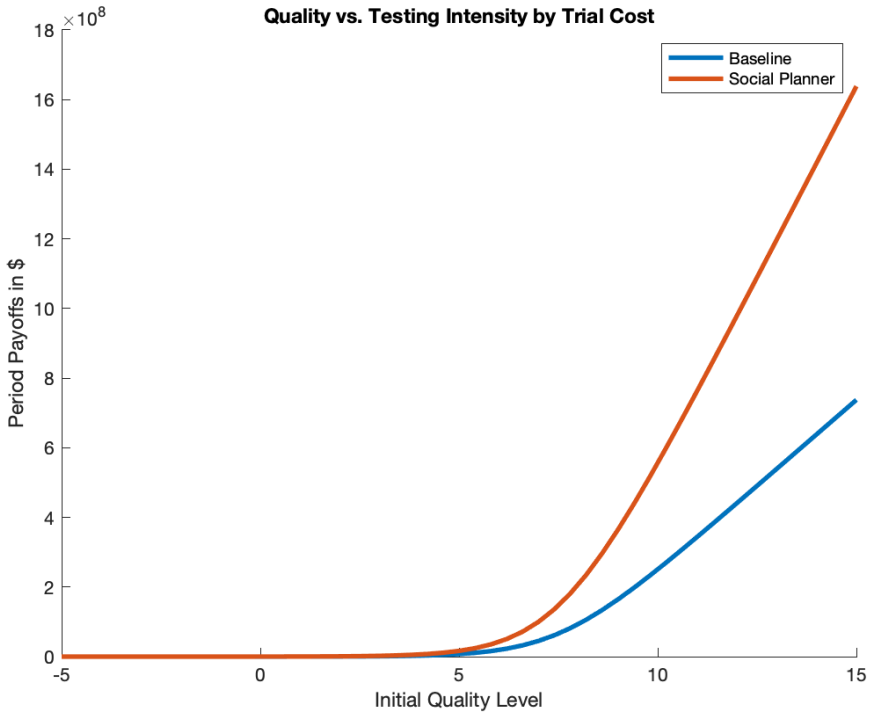


Figure A7: Figure A7

Notes: payoff social planner private – [is this consistent with what we have seen so far?]

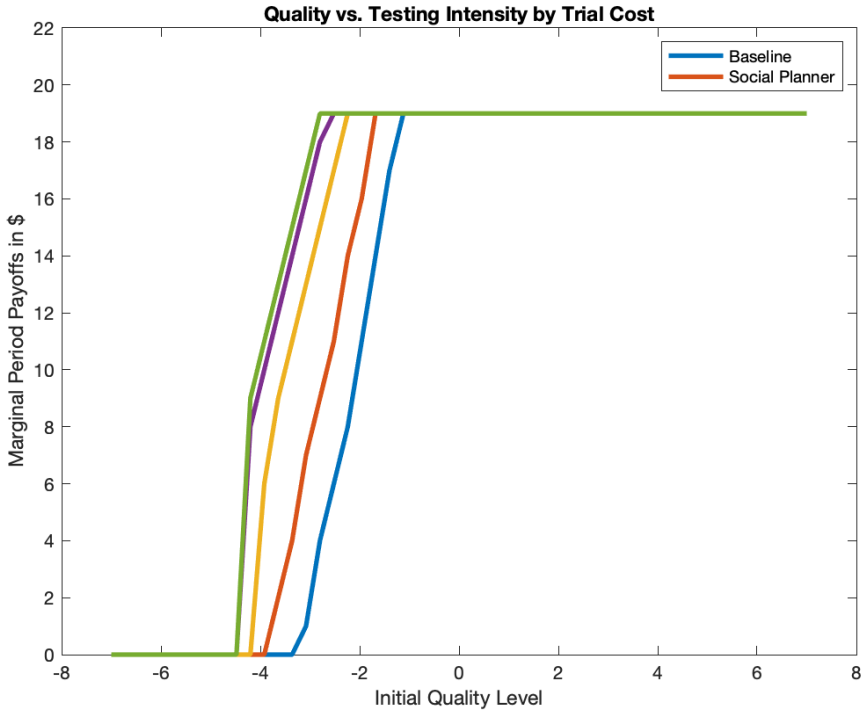


Figure A8: Figure A8

Notes: payoff social planner private mp – [this seems like wrong figure]