NBER WORKING PAPER SERIES

REGULATING INNOVATION WITH UNCERTAIN QUALITY: INFORMATION, RISK, AND ACCESS IN MEDICAL DEVICES

Matthew Grennan
Robert Town

Working Paper 20981
http://www.nber.org/papers/w20981

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2015

Data for this research was provided by Millennium Research Group (www.mrg.net). We owe thanks to Jason Abaluck, Andres Aradillas-Lopez, Steve Berry, Joe Harrington, Robin Lee, Scott Stern, and audiences at the BEA, Chicago, Clemson, FTC, HEC Montreal Summer IO Conference, iHEA, Johns Hopkins, Kellogg Conference on Healthcare Markets, Minnesota, NBER IO and Health Care Summer Institutes, NUS, Ohio State, Penn State, Rice/Houston, UCLA, Wharton, and Yale for helpful comments. The University of Toronto, Wharton Dean's Research Fund and Mack Center for Innovation, and the Indian School of Business provided financial assistance. Ellie Prager, Mike Punzalan, Emi Terasawa, and Christine Yang provided excellent research assistance. Any errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Regulating Innovation with Uncertain Quality: Information, Risk, and Access in Medical Devices
Matthew Grennan and Robert Town
NBER Working Paper No. 20981
February 2015
JEL No. I11,L11,L51

ABSTRACT

This paper examines optimal regulatory testing requirements when new product quality is uncertain but market participants may learn over time. We develop a model capturing the regulator's tradeoff between consumer risk exposure and access to innovation. Using new data and exogenous variation between EU and US medical device regulatory rules, we document patterns consistent with our model and estimate its parameters. We find: without information from regulatory testing, risk shuts down the market; US policy is close to the one that maximizes a measure of welfare derived from our theoretical model and our empirical estimates; EU surplus could increase 20 percent with more pre-market testing; and "post-market surveillance" could increase surplus 24 percent.

Matthew Grennan
Wharton School, U of Pennsylvania
307 Colonial Penn Center
3641 Locust Walk
Philadelphia, PA 19104
matthewgrennan@gmail.com

Robert Town
Health Care Management Department
The Wharton School
University of Pennsylvania
3641 Locust Walk
Philadelphia, PA 19104
and NBER
rtown@wharton.upenn.edu

1 Introduction

Most innovative new products are brought to the market because their makers believe they provide new value. However, once in the hands of consumers, there is always some chance that the product will not operate as hoped. The consequences of this failure range from consumer regret about product choice to death. When this risk matters for welfare, products often must go through pre-market testing and become approved/certified by a formal body before entering the marketplace. The standards that the regulatory body uses to approve products has the potential to fundamentally alter market outcomes. We argue that a key decision the regulator makes in setting its approval criteria is the information the manufacture is required to generate in order for the product to be approved. As first highlighted by Peltzman (1973) in the context of pharmaceuticals, higher informational standards increase product-specific learning and lower consumption risk but also result in delayed access, fewer products, and higher entry costs conditional on approval. Today such certification processes are commonplace and a source of controversy in areas as diverse as electronics, airplanes, automobiles, finance, health care, and toys.[1]

This paper uses new data and exploits exogenous regulatory differences between the US and EU to quantify the tradeoff between access and information for medical devices introduced between 2004-2013. In the US, medical devices are regulated by the Food and Drug Administration (FDA) while in the EU device approval is performed by organizations that contract with the EU called Notified Bodies. Importantly, the different regions apply different standards to medical device approval. Very roughly, the US applies a "safe and effective" standard while the EU only certifies safety of the product. This difference is material. Meeting the "effectiveness" standard often requires manufacturers to generate product performance information through large-scale randomized clinical trials. These trials are costly in both time and expense. As a result, medical device manufactures (many of which are US based) typically introduce products in the EU well before they seek FDA approval, if they decide to enter the US at all. According to the Boston Consulting Group, between 2005 and 2011, the average high risk and likely high value medical device was introduced in the US four years later than in the EU. The differences between the US and the EU in the medical device approval process have led to calls for reform in both regions. In the US, the FDA has faced attacks from both sides, with some claiming that a slower, tougher approval process is crippling innovation; and others claiming that the approval process is too lax, allowing too many dangerous devices into the market.[2] Also, as rising incomes in the developing world lead to both greater incidence of "western" diseases and greater ability to afford the most advanced technologies, the debate on how to regulate medical

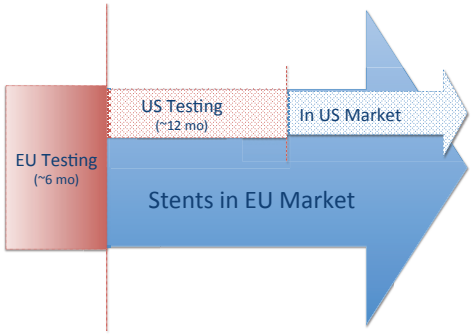
[1] See, for example in electronics "European Environmental Rules Propel Change in U.S.", The New York Times, July 06, 2004; in airplanes "Boeing Acknowledges Tests Underestimated 787 Battery Risks", The New York Times, April 23, 2013; in automobiles "U.S. Sues Chrysler After Auto Maker Refuses to Recall Cars", The New York Times, June 5, 1996; in finance "An FDA for Securities Could Help Avert Crises", Bloomberg, April 2, 2012; in toys "Toy Makers Fight for Exemption From Rules", The New York Times, September 28, 2010.

[2] For an example arguing the FDA is too lax "Report Criticized F.D.A. on Device Testing", The New York Times, January 15, 2009; and too tight "FDA Seeks to Toughen Defibrillator Regulations", The New York Times, March 22, 2013.

devices has taken on global significance, drawing the interest of the UN and WHO.[3]

Despite the importance of information from product testing and the access/risk tradeoff in markets where research and development leads to new products with uncertain quality, empirical research has been limited by two major difficulties: (1) finding exogenous variation in regulatory regimes that can identify the tradeoff between these competing forces; (2) assembling data and a corresponding empirical framework that can quantify the returns from increased information relative to the costs of decreased access. In this study we address the first challenge by exploiting the fact that the EU approval process requires less information from the manufacturers and as a consequences is both faster and less costly than the US process for any given device. This difference is largely due to historical political processes and is uncorrelated with market demand for devices. As depicted in the timeline in Figure 1, this has the dual advantage of allowing us to observe outcomes for the same devices under two regulatory regimes with different pre-market testing requirements; and it also allows us to observe EU outcomes for devices concurrently undergoing US trials and not.

Figure 1: Timeline of EU/US testing and market introduction. All devices enter the EU after safety trials. Some devices that wish to enter the US then run longer, larger efficacy trials (concurrent with being used in the EU market).



To address the second challenge of measuring learning, access, and risk, we acquire monthly data on product-level prices, quantities, and diagnostic procedures in the US and EU. These data are collected at the hospital level which we then aggregate to the geographic area. The data come from Millennium Research Group (MRG), a medical device market research firm. Our analysis focuses on the market for coronary stents. We chose this segment as the coronary stent market is large and important with excellent market data and with constant innovations introduced over time. Coronary stents treat ischemic heart disease—the narrowing of the coronary artery caused by fatty deposits. Ischemic heart disease is the leading cause of global

[3] "UN: Western Diseases a Growing Burden on Developing World," The Wall Street Journal, May 14, 2010. "Global Forum to Improve Developing Country Access to Medical Devices," press release, WHO, September 9, 2010.

death accounting for 7 million fatalities in 2010 (Lozano, 2012). In 2013 total, world-wide sales of coronary stents exceeded $7.9 billion with the vast majority of those sales occurring in the US and the EU.[4] We will be examining the introduction decisions for second and third generation stents, so the costs of introduction will be running the required clinical tests and uncertainty will be about relative quality at the individual product level (as opposed to new-to-the-world products that may face hurdles to reimbursement by payors and uncertainty regarding the fundamental technology itself).

We begin the analysis by constructing a theoretical model where products are exogenously invented with uncertain quality, market entry is regulated, and consumers learn about product quality over time. The key feature of our model is that the rate of learning in premarket clinical trials can be greater than the rate of learning after market entry. This introduces a tradeoff where more regulation leads to more information generation, learning and less risk, but also delays access and raises entry costs for new products. The model clarifies patterns in the data that one should expect as a function of the distribution of product qualities invented, the rates of learning, consumer preferences, and regulatory rules.

Our data analysis then documents multiple patterns consistent with the predictions of the model. We show that the EU enjoys greater access to the best new medical technologies, while also bearing greater risk by allowing entry of a wider range of device qualities, earlier in each device's lifecycle. The greater access in the EU is evident in the fact that on average 49 percent of the stents used in the EU are unavailable in the US at that point in time. The greater risk in the EU is suggested by the facts that on average products in the EU experience less usage overall and higher volatility in usage patterns when first introduced, with this usage discount and variance decreasing and stabilizing over the first two years on the market. The US, by contrast, exhibits no such patterns. We employ a series of reduced form analyses to establish that the patterns we observe in the EU market are driven entirely by information spillovers from US clinical trials. By focusing on within-product variation and comparing the same products launched in different regions (the US and EU) and also EU patterns for products that are and are not undergoing US trials we are able to rule out alternative mechanisms such as selection on product quality, non-learning models of product diffusion, learning from observational use vs. learning from clinical trials, and selection on product uncertainty.

In order to develop welfare measures and address policy questions regarding optimal regulation, we estimate the structural parameters of the model. We combine the data with our learning model of product choice to estimate the distribution of product qualities and risk as well as the speed of learning and preferences of consumers in the marketplace. Consistent with the reduced form evidence, the demand parameter estimates indicate that FDA required clinical trials generate significant information while there is practically no observational learning via experience in the marketplace. Furthermore, the estimates indicate that without clinical trials, the stent market would virtually fail with very few patients selecting a stent due to the risk of receiving a low quality device.

[4]Source: BCC Research 2015, "Stents: Technologies and Global Markets."

To estimate optimal regulatory policy, we take a partial identification approach instead of specifying full pricing and dynamic entry/exit models. We develop simple to compute cases that bound total surplus as a function of regulatory policy, and use these bounds to generate a partially identified set of regulatory policies. The results imply that total surplus is maximized when the average premarket clinical trial is at least seven months longer than the current EU requirements. Around the optimal regulatory policy, total surplus is relatively insensitive to the time spent in premarket testing, implying that US regulatory policy is statistically equivalent to the optimal policy.

Some FDA reform proposals advocate for more relaxed pre-market requirements but enhanced post-market surveillance. The logic behind this proposal is straightforward. It would decrease entry costs and entry lags yet would, in principle, maintain the risk reduction of stricter pre-market regulations. We examine the welfare impact of this policy in the context of our model and find that if post-approval learning rates approach those we observe from clinical trials at a comparable cost, the benefits from such a policy change are substantial. In the extreme case where post-approval learning is fully informative and not too costly, the optimal policy is to require no pre-approval trials at all, which would yield a welfare increase of 24 percent.

Because our data collection, research design, and modeling efforts are focused on the issue of information generation, risk, and access, our analysis should be interpreted as holding the other roles of the regulator as fixed. That is, regulatory bodies also set standards for what constitutes acceptable evidence and verifying the information produced in trials. We take the set of new technologies that could enter the market as exogenous and abstract away from the potentially important feedback effects from the device regulatory regime to the incentives to invest in new technologies. In this sense our paper is related to the larger literature on quality information disclosure (e.g. Dranove and Jin 2010), but whereas the focus there is typically on the difference between no disclosure and some disclosure, our focus is on the amount of information required, given a basic disclosure regime. We measure how this disclosure affects market structure and welfare through entry and consumption decisions.

We believe this focus on market structure in our paper is complementary to recent empirical research on other regulatory tools that affect product entry incentives, such as patent breadth and length (Budish, Roin, and Williams 2015) and price regulations (Filson 2012). Whereas the focus of that literature has been on innovative activity with stylized monopoly market structures, we show that the welfare impact of regulation on market structure and buyer decisions can be large as well. There are important complementarities between the value of new medical technologies and the regulatory approval product regime. New medical technologies with uncertain quality can only achieve their welfare potential if the necessary studies to document the product's clinical performance are performed. In our setting, the FDA approval process provides the necessary incentives to do those studies.

More broadly, our work builds on recent empirical research on optimal regulation (Timmins 2002; Seim and Waldfogel 2013 Miravete, Seim, and Thurk 2014) and consumer learning

(Roberts and Urban 1988; Erdem and Keane 1996; Ackerberg 2003; Crawford and Shum 2005; Ching 2010), and to our knowledge is the first to combine these two. This combination is essential in allowing us to build on the pioneering work of Peltzman (1973) where he uses pre-/post-analysis to argue that the 1962 FDA act which require clinical trials for pharmaceuticals prior to their introduction to the market harmed consumers by reducing access to drugs without increasing product information. As we rely on established models and frequently available data, we provide an approach that future researchers might find useful in the area of entry regulation via product approval/certification processes.

Our analysis of the impact of different regulatory regimes not only speaks to the broad questions of the economics of product quality regulation, but also informs policy with potentially large welfare consequences. The amount of economic activity regulated by the FDA and the Notified Bodies is significant. In the US the medical device market sales exceeded $150B in 2010 or 6 percent of total national health expenditures and approximately $130B (7.5 percent) in the EU.[5] Further, the introduction of new medical technologies are responsible for significant reductions in mortality; and in so far as different regulatory regimes affect the availability of these technologies, their welfare impact extends beyond their direct impact on commerce.

The remainder of the paper is organized as follows: The next section discusses the institutional background of medical device regulation in the US and EU. Section 3 develops a general model that captures the tradeoffs involved in regulating market entry of products with uncertain quality and derives testable predictions. Section 4 then tests these predictions in the data, finding evidence in support of the model. Section 5 takes a structural approach, explicitly estimating the parameters of the model. Section 6 derives welfare estimates for current as well as counterfactual regulatory regimes. Section 7 concludes and discusses ways one might think about extrapolating our results to devices beyond those for which we have data and the potential for extending our approach to other products and industries.

2 Medical Device Regulation in the US and the EU

Medical device is a term that applies to a broad set of product categories, ranging from crutches to pacemakers to CT scanners. In this study we will focus on coronary stents, which are themselves a blockbuster device in terms of sales and health impact, but are also typical of implantable devices that are deemed "necessary for the sustainment of life" and thus regulated as Class III devices in the US and EU. It is for this class of devices that EU and US regulatory approaches diverge most widely, creating the variation we leverage in our study.

Before detailing these regulatory differences, however, it is useful to keep in mind some basic facts about the structure of decision making and players in the markets for these important implantables. First, hospitals generate revenue by performing a procedure (such as an angioplasty with stent), and the price for purchasing the device is a necessary cost the hospital must incur. The physician who performs the procedure will typically be compensated either

[5] Donahoe and King, 2012; Medtech Europe, 2013

as a salaried employee of the hospital, or on a fee-for-service basis for the procedure, where in either case importantly the financial benefits to the physician are unrelated to the specific brand of device used. Physicians typically have strong preferences about which brand of device is best to use for a given patient/lesion type (devices in this class are often referred to as "physician preference items") because devices are differentiated in physical characteristics of the implanted device itself (for a stent examples are shape, strength, flexibility, and type of drug/polymer) and also characteristics that effect ease of implantation (for stents unexpanded size and flexibility, and controls and capabilities of the catheters and balloons used in delivery). The supply side of the market is thus a differentiated oligopoly, and prices are typically negotiated between manufacturers and hospitals or hospital systems.

For the purposes of this study the most important features of the stent market to note are the constant innovations over time in terms of both vertical quality advances that make similar yet better products for the mass market, and also horizontally differentiated products designed to address less common niche markets such as small vessel and bifurcated lesions. Interventional cardiologists are a relatively small community who stay engaged with current and upcoming technology developments through journals and several well-attended meetings each year (TCT each October, ACC in March, and ESC in August each year, as well as numerous regional affiliated conferences throughout the year) at which the most recent results of in-progress clinical trials are reported. The result is an active community that both cares and knows about the most recent technologies and evidence for these technologies.

2.1 Similarities and Differences in US and EU Regulation

Medical device regulation in the US began with the passage of the Medical Device Amendments Act of 1976. This law established the regulator pathway for medical devices in the US, placing oversight authority within the Food and Drug Administration (FDA). The criteria the FDA is mandated to use is "safe and effective." Prior to the passage of the Act, there was little regulatory oversight of the medical device sector. The Act established three classification of devices (I, II and III) which are assigned based on the perceived risks associated with using the device. Class III devices are defined as those used in in supporting or sustaining human life, of substantial importance in preventing impairment of human health, or presents a potential unreasonable risk of illness or injury. Class I and Class II devices are lower risk devices for which there is a sufficient body of evidence demonstrating a performance standard for the design and manufacturing of the device.

There are two basic regulatory pathways within the FDA to bring a device to market: Pre-Market Approval (PMA) and the 510(k). The PMA process applies to Class III devices, while the 510(k) process generally applies to Class II and some Class I devices. Under the 510(k) process the manufacturer needs to demonstrate that the device is 'substantially equivalent' to a predicate device. Generally, bench testing data and perhaps a very small clinical study is all that is necessary for a device to demonstrate equivalency. While there is no standard timetable for 510(k) clearance, a straightforward clearance can typically be obtained within

several months.

However, the approval process is much more complicated and costly for PMA devices. Approval of a PMA device generally requires the sponsor to provide data from a pivotal study. These are large, multi center, randomized clinical trials. These studies involve hundreds to thousands of patients and cost tens of millions of dollars to complete. In 2012, only 37 PMAs were approved by the FDA.

In the EU the device approval process for Class III devices is very different than in the US.[6] Medical devices are regulated by three EU Directives. The main directive is the Medical Devices Directive which passed in June, 1993 and has been adopted by each EU member state. A medical device is approved for marketing in the EU once it receives a 'CE mark' of conformity. The CE mark system relies heavily on third parties know as "notified bodies" to implement regulatory control over devices. Notified bodies are independent, commercial organizations that are designated, monitored and audited by the relevant member states via "competent authorities." Currently, there are more than 70 active notified bodies within the EU. A firm is free to choose any notified body designated to cover the particular type of device under review.[7] To obtain an CE mark a Class III medical device needs to only demonstrate safety and performance. Compliance with this standard usually can be demonstrated with much simpler and cheaper clinical trials than required by the FDA. In both the US and EU, new-to-the-world devices may face the additional hurdle of gaining payor reimbursement, but the devices we study are second and third generation products, so coverage determination has already been made prior to their introduction.

The differences in the two regulatory regimes is largely a consequence of different histories that lead up to the passing of the primary medical device legislation in the two regions. The Medical Device Directive, the centerpiece of the EU medical device regulatory framework, was passed in 1993 when there was keen interest in a new approach to harmonizing regulatory frameworks across the member states. The EU had just undertaken a long and frustrating harmonization process for food and drugs. This new approach sought to avoid detailed and bureaucratic government approval processes, particularly duplicative approvals. This framework was also applied to other products including toys, pressure vessels and personal protective equipment. In contrast, the US medical device regulatory framework was established after the Dalkon Shield injured several thousand women which garnered significant public outcry. The FDA already had oversight on some aspects of medical devices and expanding that role was the most viable political option. At that time, a non-governmental approach to device regulation was never seriously considered by the Congress.

The gap between the two regulatory systems is the focus of a number of consulting, lobbying, and government reports. For example, a series of Boston Consulting Group reports shows that there is no difference in recalls between devices that are marketed in both the US and the EU.

[6] Actually, there are four different classes of medical devices in the EU(Classes I, IIa, IIb and III). Class III devices in the EU closely map into Class III devices in the US.

[7] See *Guidelines Relating to Medical Devices Directives, http://ec.europa.eu/health/medical-devices/documents/guidelines/.*

Of course, as we show below, the mix of devices that are introduced into the US is different and thus it is unclear what this study says about the impact of counterfactual regulations on device safety. In fact, the FDA countered the BCG report with their own case study of 10 devices that were approved in the EU, not approved by the FDA, and lead to significant adverse events in patients. The FDA study only focused on the negative consequences of the EU's relatively lax regulatory standards and does not acknowledge the benefits of greater access to devices in the EU.

While the consequences of the different regulatory regimes has generated significant policy debate, what is less controversial is that there are significant lags between the US and the EU in device introduction. Conditional on entry into both the US and the EU, BCG documents that medical devices are introduced into the US approximately four years after the EU.[8] In the next section we develop a theoretical framework for assessing the trade-offs inherent in the different regulatory approaches. A notable advantage of our model is that the key parameters can be directly estimated from commonly available data, and thus the welfare of counterfactual policies can be assessed.

3 A Model of Quality Uncertainty, Learning, Entry Regulation, and Consumer Choice

We now develop a model that captures the tradeoff between risk and access involved in regulating market entry of products with uncertain quality. In our model, products are developed with uncertain quality; this uncertainty is potentially resolved over time via exogenous signals (e.g. from clinical trials or other research); a regulator restricts entry by requiring costly premarket clinical trials to accelerate learning; and risk-averse consumers choose from the available products in the market at a point in time.

Our model captures many of the salient features of medical device markets and the role of the regulator. However, the medical device sector is complicated and there are notable institutional features that we purposefully de-emphasize in order to keep the model tractable and parsimonious. As we have modeled, medical device quality is uncertain, but this uncertainty is symmetric among manufacturers, regulators, and consumers. If manufacturers are differentially informed about their devices quality, device regulation could solve a lemons problem (Leland 1979). At the extensive margin of whether to have any regulation at all, the lemons problem is surely relevant. However, our focus is on the appropriate standards of that regulation not on whether the regulation should exist at all. The variation that we exploit aligns with this focus. The EU is more lax in their standard relative to the US yet we are unaware of any significant evidence that the device market in the EU 'unravels' more than in the US. In fact, the presence of many more device offerings in the EU suggests that the variation in regulations between the US and EU is not a margin that would induce a lemons type market failure. This may be due to the fact that the relatively small number of interventional cardiologist who use the devices

[8] BCG (2012) *Regulation and Access to Innovative Medical Technologies.*

we study tend to stay well-informed about the most recent clinical updates for the products on the market (two days each at the three major conferences are devoted to reporting such data).

We also do not model the possibility that the regulator will reject a device. We do, however, allow for manufacturers' optimal entry decisions in the face of clinical trial and entry costs conditional on developing the device. This amounts to an implicit assumption that no firm would enter with a product the regulator would want to reject, which is exactly the case under our symmetric information assumption.[9]

The next several subsections lay out the model. Section 3.1 describes how market participants learn about product quality over time, Section 3.2 describes consumer behavior and how it is affected by uncertainty about product quality, Section 3.3 turns to supplier pricing and entry, and finally Section 3.4 lays out the role for a regulator to affect total surplus via information requirements and their effect on risk and access.

3.1 Consumer Learning

In our framework, consumers (and manufacturers and regulators) are uncertain about the quality of newly developed stents. Information on the quality of a given a stent accrues to the market over time from two sources, and the quality of the information from each source is potentially different. First, products undergo clinical trials, and information from these trials diffuses to physicians through updates reported at major cardiology meetings throughout the year and published articles in medical journals. Second, usage of the product may generate observational learning which is then shared and diffuses to the market.

Specifically, we assume innovative new devices j are each developed with quality Q_j according to a distribution $F_t(Q)$:[10]

$$Q_j \sim F_t(Q) := N(\overline{Q}_t, \sigma_Q^2). \tag{1}$$

where the subscript t allows for technological advancement over time.[11]

Over time, unbiased but noisy signals A arrive regarding the product's quality as new information from ongoing clinical trials and real world usage are generated, released, and diffused into the market. Letting age a denote the time since product j was introduced into

[9]The FDA does not report data on the number of PMA applications that are rejected. However, there are frequent communications between the FDA staff and the device manufacturer over a given application and the evidence the FDA needs to see prior to granting approval. Thus, our impression is that there are few devices that submit a full PMA applications that are rejected as the process is costly for the manufacturer and they have a relative precise estimate of the likelihood of approval prior to submission. This differs from the pharmaceutical context where there is often disagreement between the manufacturer and regulator regarding the weighting of any positive treatment effect vs. potential negative side effects. Due to their local and mechanical nature, devices typically have few if any severe side effects.

[10]For simplicity, we assumed the prior and signal process to be normally distributed. In principle, we could relax this assumption. However, sample size limitations make taken a more non-parametric route undesirable, and we find that the simple normal model fits the data quite well with a small number of parameters.

[11]We abstract away from any feedback effects from the regulatory regime on the incentive of device manufacturers to invest in developing new products and assume that the quality distribution is exogenous.

the market (not calendar time), A_{ja} is given by:[12]

$$A_{ja} = Q_j + \nu_{ja} \qquad \text{where} \quad \nu_{ja} \sim \begin{cases} N(0, \sigma^2_{A^c}) & \text{if in clinical trials} \\ N(0, \sigma^2_A) & \text{if not} \end{cases} \tag{2}$$

Given these signals, beliefs about product quality are updated via Bayes' rule, and due to the normally distributed prior and signal, posterior beliefs are also distributed normal with mean:

$$Q_{ja+1} = \frac{\sigma^2_{ja}}{\sigma^2_{ja} + \sigma^2_{A^{ja+1}}} A_{ja+1} + \frac{\sigma^2_{A^{ja+1}}}{\sigma^2_{ja} + \sigma^2_{A^{ja+1}}} Q_{ja} \tag{3}$$

and variance:

$$\sigma^2_{ja+1} = \frac{\sigma^2_{A^{ja+1}}}{\sigma^2_{ja} + \sigma^2_{A^{ja+1}}} \sigma^2_{ja}. \tag{4}$$

With this uncertainty and learning as a backdrop, the regulator must decide on the required length of clinical trials, trading off the costs of later access versus the benefits of more information and thus reduced risk. Once a product has been subjected to the required clinical trials, it is released to the market and consumers (doctors and patients) make decisions about which product to use given the current available choice set and information. Because the regulator weighs the implications for total surplus in its decision, we begin with the consumers' problem and work backwards.

3.2 Consumer Choice

Given beliefs regarding a stent's quality and the uncertainty over that quality, we assume consumers select the stent that yields the highest expected utility. For each patient/doctor combination i the indirect expected utility function from using device j at time t (where the t subscript refers to the calendar month, which will be associated with different product age a for different products) takes the form

$$u_{ijt} = Q_{jt} - \frac{\rho}{2}\sigma^2_{jt} + \epsilon_{ijt}, \tag{5}$$

where ρ is the coefficient of risk aversion, and ϵ_{ijt} is an i.i.d. error term capturing the deviation of doctor preferences and/or patient appropriateness for device j relative to the population average. In our empirical exercise, we do not find price to be a statistically or economically significant determinant of demand. This is not surprising since the vast majority of patients receiving a stent are insured and physicians generally do not have an incentive to consider device cost in their product selection decision. For this reason, we leave price out of the demand specification.[13]

[12] We assume that information release from a clinical trial accrues to the market with a consistent signal each month. This fits with the regular release of interim results at major meetings and in journal articles. We have examined and found no evidence of a discrete demand response in the EU upon US trial completion or FDA approval.

[13] See Appendix D for regressions with price. Prices for stents are generally set via bargaining between the relevant hospital authority and the device manufacturer, and thus a pricing equilibrium exists even though

Assuming consumers choose the product j that maximizes expected utility from the set of products available $\mathcal{J}_t$, the set of patients for whom a doctor chooses product j (in month t) is then $\mathcal{A}_{jt} := \{i|j = \arg\max_{k\in\mathcal{J}_t} u_{ikt}\}$. Then expected quantities are then given by the market size M_t and the choice probabilities:

$$q_{jt} = M_t s_{jt} = M_t Pr[j = \arg\max_{k\in\mathcal{J}_t} u_{ikt}] = M_t \int_{\mathcal{A}_{jt}} f_t(\epsilon)d\epsilon = \frac{e^{Q_{jt}-\frac{\rho}{2}\sigma_{jt}^2}}{\sum_{k\in\mathcal{J}_t} e^{Q_{kt}-\frac{\rho}{2}\sigma_{kt}^2}}, \tag{6}$$

where the last equality obtains from the standard "logit" assumption that ϵ is distributed i.i.d. extreme value type I with unit variance. The choice set always includes an outside option $j = 0$, with utility normalized to zero, representing the best non-stent treatment for that patient.

Total surplus per patient (relative to the best non-stent alternative; not including fixed costs) can then be obtained by summing over patient utility:

$$TS(\mathcal{J}_t) = \int_{\mathcal{A}_{jt}} u_{ijt} f_t(\epsilon)d\epsilon = \ln\left(\sum_{j\in\mathcal{J}_t} e^{Q_{jt}-\frac{\rho}{2}\sigma_{jt}^2}\right) , \tag{7}$$

where the final equality obtains from the logit distributional assumption on ϵ.

3.3 Bounds on Supply Effects on Total Surplus

We are interested in total surplus as the object that the regulator should seek to maximize in its choice of the optimal length of clinical trials T^c. Total surplus is a function of the choice set $\mathcal{J}_t$, which in turn is a function of supplier entry behavior, given costs of trials required for entry $\phi^e(T^c)$ and expected profits after entry, conditional on the expected behavior of other firms. A fully specified supply model requires models of pricing and entry/exit dynamics. In the case of our analysis, both of these modeling efforts would entail a combination of approximating assumptions and work at the frontier of both the business-to-business contracting and dynamic oligopoly literatures. For pricing, we would need to either construct an approximate expected price function at our region-time level of analysis, or build on the hospital-time level analysis in Grennan (2013, 2014) to allow for strategic choice in who contracts with whom (Lee and Fong 2013), an important feature of the EU market during the time frame we study. For dynamic entry and exit, we face a problem with a large and continuous state space, requiring approximations of the type explored in recent papers such as Ifrach and Weintraub (2014). Because constructing such models involves a substantial refocusing of the contributions of this paper to arrive at point estimates that would be caveat to a number of additional modeling and approximating assumptions, we instead take a bounds approach to partially identify the answers to our policy questions of interest.

We construct bounds on the total surplus generated under any regulatory policy $TS(T^c)$ that rely on weak assumptions on supply side behavior:

consumer demand is perfectly inelastic. See Grennan (2013) (which also finds an economically small response of demand to price) for more discussion on the role of bargaining in determining stent prices.

Supply Assumption 1 (EU Entry Costs): EU entry costs are low enough such that all products developed with positive expected profits after EU testing enter.

Supply Assumption 2 (Entry Policy): The equilibrium entry policy of firms is increasing in own expected profits: $\pi_j > \pi'_j \Rightarrow Pr[Enter_{jt}|\pi_j] > Pr[Enter_{jt}|\pi'_j]$

Supply Assumption 3 (Pricing): Prices are bounded by the marginal contribution of the product: $p_j - c_j \leq TS(\mathcal{J}) - TS(\mathcal{J} \setminus \{j\})$.

Under these three assumptions, we can construct upper and lower bounds for total surplus. The upper bound is given by the case where there are no direct fixed costs of longer clinical trials, so all firms enter in equilibrium, and the only impact of increasing trial length on market structure is to delay access to the newest technologies (in addition to increasing learning). The lower bound is given by assuming that the cost of trials is \$1.6M per month, but with firms' entry decisions made under the assumption that other firms enter as if entry costs are zero. Less firms will enter than in a full equilibrium model because this case doesn't allow expected market shares and prices to increase as fixed costs increase and the market becomes more concentrated. Full proofs for both of these bounds are provided in Appendix A.1.

The advantage of these particular bounds is their simplicity of computation. Expected profits and thus entry decisions can be computed directly from the data, and then total surplus can be computed using the estimated demand model for the set of products that enter. How informative these bounds will be depends on the size of trial costs relative to the distribution of product qualities in the data. The bounds will be equal to each other and the full equilibrium model at $T^c = 0$ (the case of no US trials beyond EU requirements), then diverge as increasing entry costs drive a wedge between the entry assumed in the bounds and the entry that would obtain in a full equilibrium.

3.4 Modeling the Regulator's Tradeoffs

The total surplus equation (7) captures the primary tradeoff between access and risk: the longer time T^c that products spend in premarket clinical trials, the lower the risk from uncertainty about product quality in the market σ_{jt}, but the less new technologies available in the consumer choice set $\mathcal{J}_t$ at any point in time and greater costs of entry. This tradeoff can be formalized by writing total surplus as a function of time spent in premarket clinical trials and considering the marginal return to increasing the amount of time spent in premarket testing to $T^c + 1$ over any time period $t = 1, ..., \tau$:

$$TS(T^c + 1) - TS(T^c) = \sum_{t=1}^{\tau} \ln \left(\frac{\sum_{j \in \mathcal{J}_t(T^c+1)} e^{Q_{jt} - \frac{\rho}{2}\sigma^2_{jt}(T^c+1)}}{\sum_{j \in \mathcal{J}_t(T^c)} e^{Q_{jt} - \frac{\rho}{2}\sigma^2_{jt}(T^c)}} \right) - \phi^e \left| \mathcal{J}^e_t(T^c + 1) \setminus \mathcal{J}^e_t(T^c) \right| , \tag{8}$$

where $\mathcal{J}^e_t(T^c)$ is the set of firms who enter in period t, given testing requirements T^c. From the social planner / regulator's perspective, the optimal length of a clinical trial sets (8) to zero.

One way to very clearly see the tradeoff between access and risk as a function of trial requirements is to consider the simplest scenario where there is no observational learning once a product enters the market and no direct cost of premarket testing. In this case, the per-period marginal return to increasing premarket testing simplifies (proof of this special case and simplification in Appendix A.2) to:

$$\frac{TS(T^c+1) - TS(T^c)}{\tau} = \frac{\rho}{2}(\sigma^2_{T^c} - \sigma^2_{T^c+1}) - \frac{1}{\tau}\ln\left(\frac{\sum_{j\in\mathcal{J}_\tau(T^c)} e^{Q_{jt}}}{\sum_{j\in\mathcal{J}_0(T^c)} e^{Q_{jt}}}\right) . \tag{9}$$

The first term captures the per period utility gain from decreased risk (and is determined by the unconditional uncertainty in product quality σ_Q and the rate of learning in trials σ_{A^c}). The second term captures the total surplus generated by the rate of technological improvement in product quality over time (which will be determined by the rate of drift in product quality over time $\overline{Q}_t$ and the expansion of variety in the choice set $\mathcal{J}_t$).

3.5 Model Predictions to Take to the Data

The model has several testable implications that we can take to the data. In order to map the model into the data, we make a revealed preference assumption and match the choice probabilities implied by utility maximization to the market share data, and invert the system as in Berry (1994) to recover the mean utility estimates for each product in each month:

$$\ln(s_{jt}/s_{0t}) = \delta_{jt} := Q_{jt} - \frac{\rho}{2}\sigma^2_{jt} . \tag{10}$$

The model then implies the following:

Prediction 1 (Time and Monetary Costs of More Stringent Regulation): If the US requires longer clinical trials ($T^{c,US} > T^{c,EU}$), this implies that for the set of products introduced in both the EU and US, US entry should lag EU entry. Further, if trials have non-negligible direct costs, then the US should experience less entry than the EU.

Prediction 2 (Learning): Learning implies that product-specific quality estimates converge with age ($|\delta_{ja} - \delta_{ja+1}| \searrow^{a\to\infty} 0$).

Prediction 3 (Risk Aversion): If consumers (doctors making decisions on behalf of their patients) are risk averse ($\rho > 0$), then product usage (on average for a given age to remove signal noise) should increase as learning occurs ($\frac{1}{J_a}\sum_{j=1,a=a}^{J_a}\delta_{ja} \nearrow^{a\to\infty} \overline{Q}$).

The summary statistics and reduced-form analysis in the next Section will use these predictions in exploring how much more stringent the US regulatory policy actually is relative to the EU, and the implications for learning, risk, and access in the market.

4 Data and Preliminary Analysis of Access/Risk in US and EU

In this Section we introduce the data on product entry, usage, and pricing that we use in our analysis. We then proceed to document the basic patterns in the data as they relate to our model and the mechanisms behind product learning, risk, and access.

The data used in this study consists of quantities and prices at the product-hospital-month level, collected by Millennium Research Group's (MRG) *MarketTrack* survey of hospitals across the US and EU from 2004-2013. This survey—covering approximately 10 percent of total market activity—is the main source of detailed market intelligence in the medical device sector, and its goal is to produce representative estimates of the distribution of market shares and prices by region. Though we use the hospital level data for some relevant summary statistics, for the majority of our analysis we aggregate the data to the region-month (US and EU) level in order to obtain accurate measures of market entry and overall usage of each device within a region, which is the relevant unit of observation for this study. In principle we could perform the analysis at the country level. However, because of MRG's sampling hospital sampling strategy, country level measures of market shares would likely contain significantly more measurement error than the regional measures. The demand models we estimate are not well suited to data with significant measurement error in the market share information and primarily for that reason the unit of analysis is the region.

In addition to the detailed market data, we also collected clinical trial data from various journal articles, conference abstracts, press releases, and product catalogs for a subset of products for which we could locate such data online. This data provides further evidence regarding the size and length of trials required for US versus EU entry.

4.1 The EU has Access to More, Newer Technologies

Our model and qualitative institutional knowledge predict that the EU enjoys greater access to a broader variety of devices and these devices are available earlier than in the US. However, EU consumers will posses less information on the quality of these products and hence face greater risk from lower quality products. Below we show that these predictions are borne out in the summary statistics.

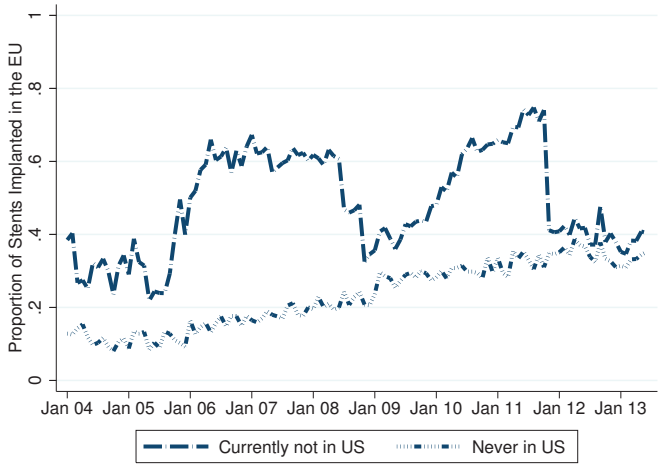
The top third of Table 1 presents summary statistics for our clinical trial data. We were able to find such data for 48 percent of the products entering the US and 16 percent of the products that enter the EU only. Conditional on publishing a clinical trial online, the informational content of the EU-only products is much less than products that are introduced in the US: On average, US clinical trials enrolled over 1,200 patients and lasted almost 3 years, while the EU-only products enrolled a third of the patients and the trials were 66 percent shorter in duration.

The bottom two thirds of Table 1 and Figure 2 show how these pre-market testing requirements are correlated with market structure and product usage in the US and the EU over our sample period. The EU has over three times as many manufacturers and products in the

Table 1: US and EU differences in clinical trial size and length, and resulting differences in market structure. The US has longer, larger clinical trials, less manufacturers and products, and later entry dates than the EU for the subset of products that enter the US.

	US	EU
Clinical trial data:		
Mean clinical trial size (patients)	1252	471
Mean clinical trial length (months)	32	11
Market structure data:		
Mean manufacturers in market	4	21
Mean products in market	11	39
Total products in market (2004-13)	21	109
Mean months from EU to US entry	10	-
Mean months from EU to US entry (DES)	17	-

Figure 2: EU market share of products not available in US. On average over the sample period 49 percent of stents used in the EU were not *currently* available in the US; and 23 percent were *never* available in the US.



market. For those products that eventually enter the US, the average lag time between EU and US introduction is 10 months (17 months for the more technologically advanced DES). Many of the products to which the EU has greater access are important, frequently-used products. In the average month, 49 percent of the stents used in the EU are unavailable in the US at that point in time, and 23 percent will never be available in the US. Our subsequent analysis will show this is a mixed blessing, with some greater access to new high-quality products, and some usage of uncertain products that turn out to be low quality ex-post.

4.2 The EU Grants Access to More Technologies with Lower and More Uncertain Quality

The data on clinical trials suggests that less information is generated for devices in the EU. We also documented that the EU has access to more devices. In the context of our model, these two observations imply that this greater access enjoyed by the EU comes along with greater risk in the form of more low quality devices and more uncertainty regarding device quality at the time market access is granted. Below, we document the presence of these patterns in our data.

Figure 3: EU vs. US. Left panel (a) plots mean across products of the mean utilities $\frac{1}{J_a}\sum_{j=1}^{J_a}\ln(s_j/s_0)$ by age since introduction into each region. Right panel (b) plots mean absolute differences $\frac{1}{J_a}\sum_j|\ln(s_{ja}/s_{0a})-\ln(s_{ja+1}/s_{0a+1})|$ by age, which should be larger with more uncertainty, and converge toward zero with learning.

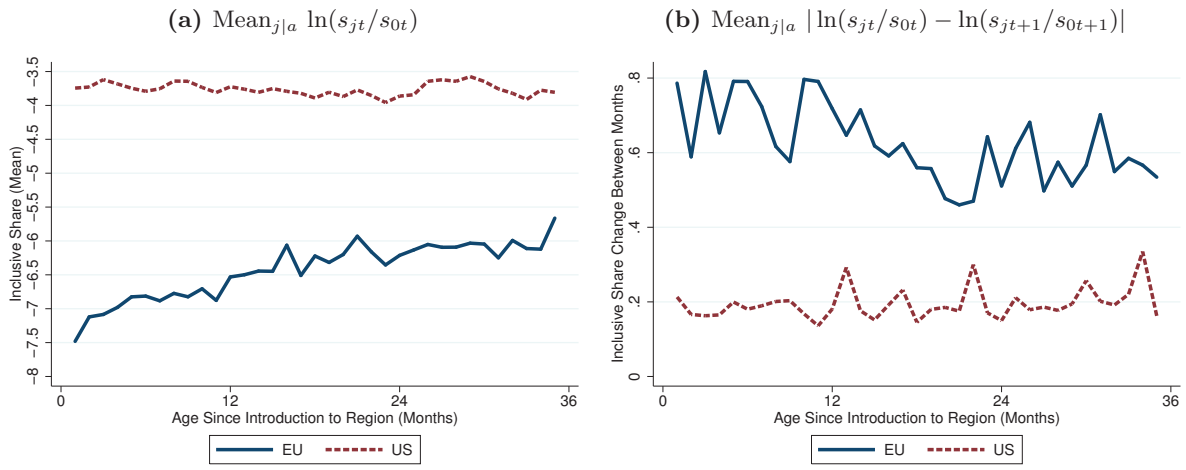


Figure 3 begins to explore the idea that EU consumers bear more risk than those in the US by introducing a larger number of devices earlier in their life cycles with less information imparted about the quality of those devices. Panel (a) shows that in the EU the mean across products of a given age of the mean utility estimates $\frac{1}{J_a}\sum_j\ln(s_{ja}/s_{0a})$ is lower upon introduction and gradually increases with age, plateauing after approximately two years in the market. The pattern in the US is different. There the mean of the mean utility estimate does not vary with product age and is higher on average.

Panel (b) plots mean absolute differences of the mean utility estimates over time $\frac{1}{J_a}\sum_j|\ln(s_{ja}/s_{0a})-\ln(s_{ja+1}/s_{0a+1})|$, which should asymptote toward zero as learning moves the quality estimate, Q_{ja}, toward true quality, Q_j. Again, this statistic is decreasing in the EU and constant over the product lifetime in the US.

These inclusive share patterns are consistent with greater uncertainty regarding product quality early in the product lifetime in the EU which is gradually resolved over time via learning. The fact that the mean inclusive share in the EU is lower early on aligns with our model if consumers are risk averse and are discounting products whose quality is more uncertain.

However, these plots with the raw data leave open several alternative explanations, which are explored further in the next section.

4.3 Evidence of Learning Versus Diffusion or Selection

While the patterns in the raw data in Figure 3 are consistent with our model of learning, they are also potentially consistent with several alternative mechanisms: drivers of product diffusion other than learning, observational learning flowing from the EU to US, and product selection on uncertainty or on quality via early exit. Figures 4 and 5 rule out these alternatives in favor of the mechanism of EU learning through US clinical trials.

All of the patterns in these figures are constructed after subtracting product means, so that all patterns are driven by within-product variation over time. In particular, this rules out any composition effect whereby increasing usage with product age is driven by worse products exiting the market at a younger age.

Figure 4: Same products, EU vs. US. Left panel (a) plots mean across products of the mean utilities $\frac{1}{J_a}\sum_{j=1}^{J_a}\ln(s_j/s_0)$ by age since introduction into each region. Right panel (b) plots mean absolute differences $\frac{1}{J_a}\sum_j|\ln(s_{ja}/s_{0a})-\ln(s_{ja+1}/s_{0a+1})|$ by age, which should be larger with more uncertainty, and converge toward zero with learning.

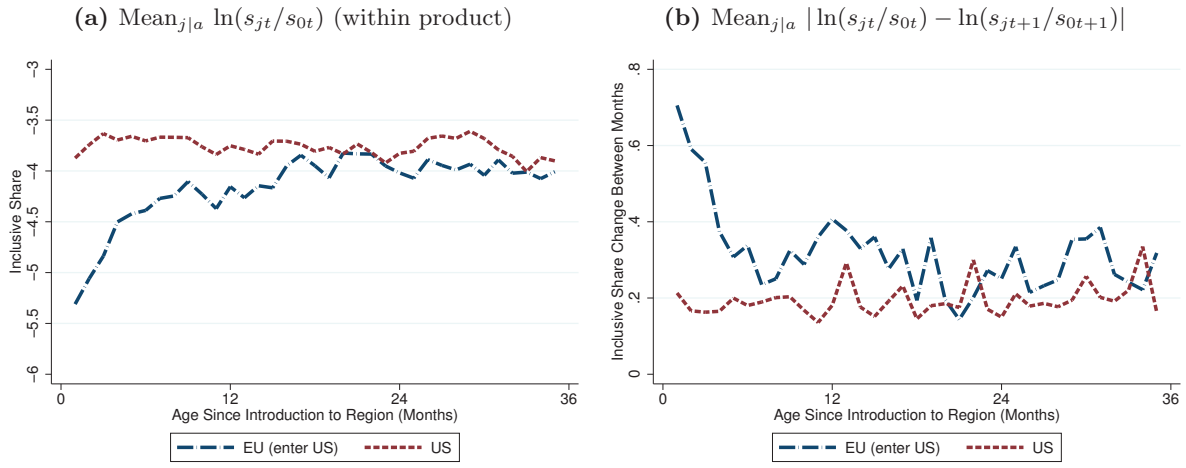


Figure 4 compares patterns for the exact same products in the EU vs. US. The fact that mean of mean utility estimates in panel (a) is flat in the US rules out that the increasing usage with age is driven by marketing, sales or distribution efforts that are specific to product launch in a new region—if it were such a non-learning diffusion story the US would exhibit a similar pattern (and the learning effect could then be estimated as the difference between the shapes of US and EU curves). Panel (b) plots mean absolute differences $\frac{1}{J_a}\sum_j|\ln(s_{ja}/s_{0a})-\ln(s_{ja+1}/s_{0a+1})|$, which which in the EU starts near 0.7 logit utils and converges fairly rapidly over the first 6-12 months on the market before leveling off around 0.2 (that it does not asymptote to 0 suggests there is some unobservable source of noise in the product usage data that we have not modeled). By contrast, in the US this same quantity stays level at 0.2 the entire time after

introduction. That the volatility and mean usage in the EU converge to US levels when we restrict the sample to the same products in both regions is consistent with learning occurring during the time these products are in the EU, and risk-averse users responding to this learning. Figure 5 explores whether the source of this learning is from observational usage in the EU or from US clinical trials.

Figure 5: EU only, products in trials vs. not. Left panel (a) plots mean across products of the mean utilities $\frac{1}{J_a}\sum_{j=1}^{J_a}\ln(s_j/s_0)$ by age since EU introduction. Right panel (b) plots mean absolute differences $\frac{1}{J_a}\sum_j|\ln(s_{ja}/s_{0a})-\ln(s_{ja+1}/s_{0a+1})|$ by age, which should be larger with more uncertainty, and converge toward zero with learning.

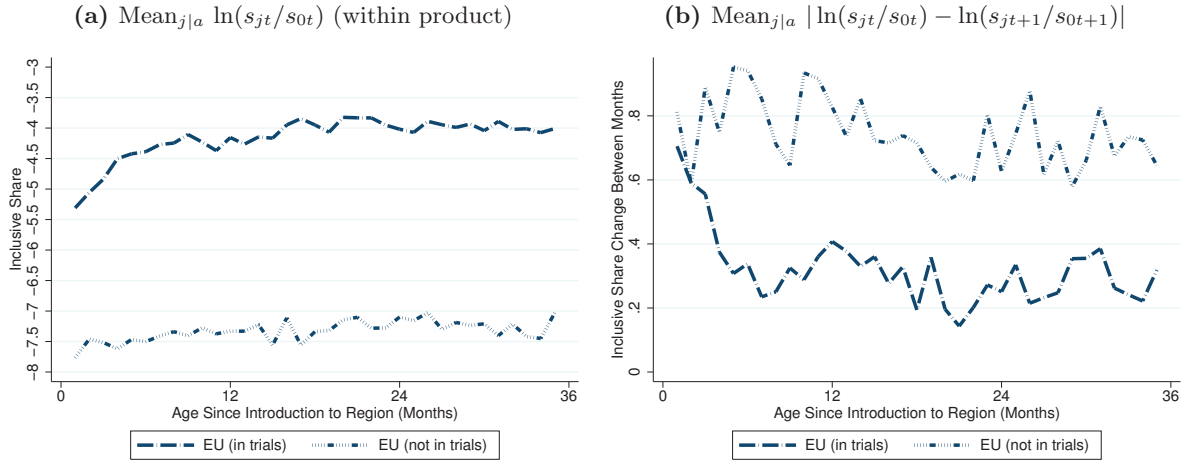


Figure 5 compares products in the EU: (1) that undergo clinical trials for US release (a set that is the same as the EU products that eventually enter the US shown in the previous figure, plus a few products that undergo US trials but are not introduced to the US); and (2) that never undergo trials beyond those required for EU introduction. The mean of mean utility estimates in panel (a) reveal that all of the evidence of learning is driven by those products in clinical trials. The curve is flat for products not in trials, which not only rules out observational learning for these products, but also rules out diffusion driven by marketing, sales, or distribution of new-to-the-world products. The mean absolute differences in panel (b) reconfirm that the learning in the EU is driven by the products undergoing clinical trials. Importantly, they also refute the argument of selection on uncertainty, as the EU products not in trials begin near 0.7 (the same as the products in trials), but remain flat near 0.7 over time, suggesting that there is plenty of uncertainty for these products, but no learning.

Thus the evidence from the data is consistent with our model in which there is uncertainty about new product quality, learning occurs over time, and risk-averse decision makers factor uncertainty about quality into their product choice. The data also suggests that there is very little observational learning outside of clinical trials. This is perhaps not surprising as the interventional cardiologist who selects and inserts the stent is generally not performing any long-term monitoring. Thus, they are not well positioned to accumulate information about a stents

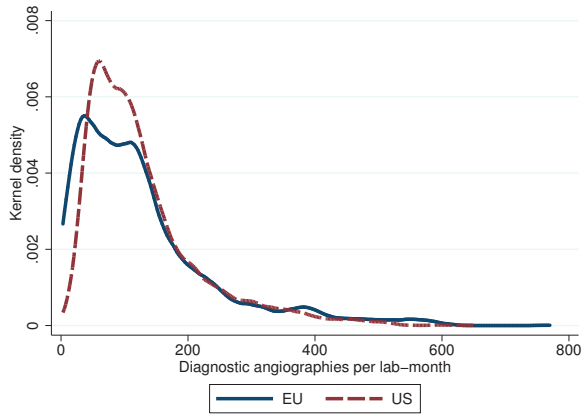
performance from their own experience or other interventional cardiologists' experience. We will return to the potential value that might be created if this type of "post-market surveillance" were increased.

4.4 Regulatory Differences Are Not Driven by Differences in Disease Incidence or Treatment Preferences

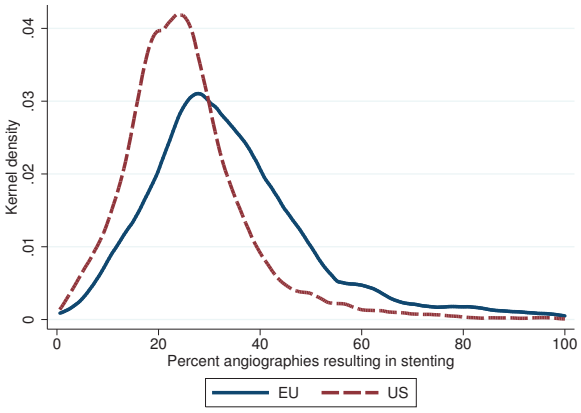
In theory it could be that the differences in usage patterns between the US and EU are driven by differences in disease incidence, preferences for angioplasty and stents, or variation in price setting regimes between the US and EU. However, all the evidence that we have been able to gather indicates that these explanations do not plausibly explain the patterns in the data described above. For example, the average ischemic heart disease mortality rate is very similar between the US and the EU, suggesting that the disease incidence is also similar. The 2010 mortality rate in the US for ischemic heart disease was 126.5 deaths per 100,000; and the corresponding figure for the EU is 130.0 per 100,000.[14] This modest differential seems unlikely to account for the stark differences of entry rates between the two regions.

Figure 6: Comparison of diagnostic procedure patterns, EU vs. US. Left panel (a) plots the distribution of number of diagnostic procedures across hospitals—the US and EU are nearly identical. Right panel (b) plots the distribution across hospitals of the probability that a diagnostic procedure results in stenting—the EU is shifted slightly to the right of the US, with a mean of 32 versus 28 percent.

(a) Diagnostic Angiographies



(b) Percent Resulting in Stenting



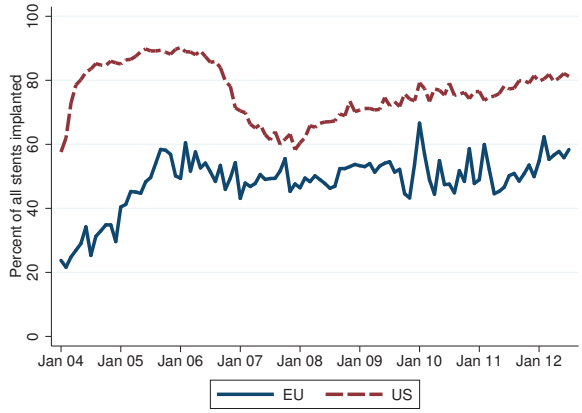
Prior to performing an angioplasty in which a stent may be inserted, the patient must undergo a diagnostic angiography. In this procedure, the blood flow through the coronary artery is visualized and this information is used to determine whether the patient should receive a stent or some other medical intervention. If the difference in the number of stents available between the EU and the US was driven by higher demand for stents, then it should show up in the data with the EU performing a larger number of angiographies or having a higher rate

[14] OECD *Health at a Glance, 2013.*

of stenting conditional on the angiography rate. Figure 6 documents the distributions of the number of diagnostic angiographies performed across the hospitals in our data and percent of those diagnostic procedures resulting in a stenting procedure across hospitals in the US and EU samples. The distributions are close to identical statistically, with the EU having a few more small volume hospitals and hospitals that are more likely to place a stent conditional upon a diagnostic procedure. In the EU, 32 percent of patients received a stent conditional on an angiography while in the US that figure was 28 percent. Like the evidence on heart disease prevalence, this small difference seems unlikely to explain the large disparity in entry rates between the two regions.

Figure 7: Comparison of usage and price patterns EU vs. US. Left panel (a) plots the percentage of stents used that are DES over time—the US uses DES 72 percent of the time on average, while the EU averages 49 percent, but both follow the same qualitative pattern over time. Right panel (b) plots quantity-weighted mean prices for DES and BMS over time—all prices fall over time, but EU prices for both technologies are on average 60 percent of those in the US.

(a) DES Share of All Stents



(b) Average Price Paid

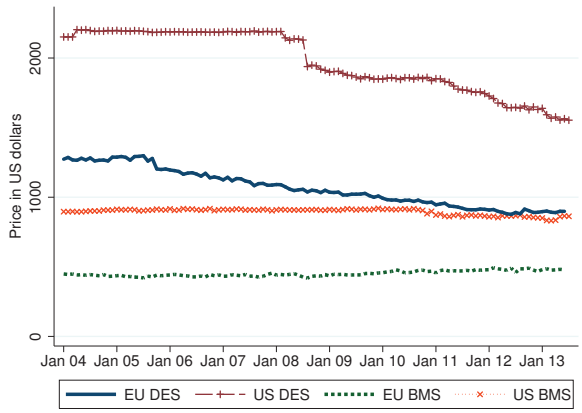


Figure 7 documents that DES usage as a percentage of all stents used is lower in the EU but follows similar patterns to the US over time. If the increased DES entry in the EU was driven by higher demand, we would expect the opposite pattern. Figure 7 also shows that the prices and hence profits per stent sold are lower in the EU. This is true for both BMS and DES and is true over our entire sample period. Both of these patterns are likely the result of lower reimbursement levels for stent procedures overall, lower DES reimbursement levels in particular, and more competing devices in the EU market. These findings suggest that conditional upon FDA approval, average variable profit in the US is higher making it a more attractive entry target than the EU. This, in turn, suggests that the differential entry rates is driven by differences in regulation and not underlying demand.

5 Identification, Estimation, and Results

The statistics presented in the previous Section 4 align with the predictions of the model of regulation and learning developed in Section 3 and suggest that the EU is indeed less stringent than the US in regulating the entry of new medical devices. In this Section we estimate the parameters of our model in order to better understand and quantify the broad impact of product regulation policies on welfare. Using the quantity date from the EU 2004-13, we estimate the distribution of product quality for innovations that could be introduced in the US and EU, the rates of learning over time, and risk aversion. We then use the parameter estimates to explore the economic and policy implications of our model. Specifically, we quantify the welfare generated under different premarket clinical testing requirements (including those observed in the EU and US) and under a proposed alternative policy that would relax premarket requirements but increase the rate of observational learning through increased post-market approval data collection and reporting.

The parameters of the utility function—and by extension the parameters of the device quality distribution and learning process—can be estimated by a revealed preference assumption and data on device market shares in each month. Matching the choice probabilities implied by utility maximization and the market share data, and inverting the system as in Berry (1994) to recover the mean utility parameters gives

$$\ln(s_{jt}/s_{0t}) = \delta_{jt} := Q_{jt} - \frac{\rho}{2}\sigma^2_{jt} := Q_j - \frac{\rho}{2}\sigma^2_{jt} + \xi_{jt} \ , \tag{11}$$

where the unobservable ξ_{jt} in the final equation includes any errors in the current expected quality estimate Q_{jt}. The main challenge here is that none of the variables on the right hand side of this equation are directly observed in the data. Our strategy will be to use variation over time and across products to estimate the product qualities Q_j, the mean $\overline{Q}$ and variance σ^2_Q of the product quality distribution, the signal variances σ^2_A and $\sigma^2_{A^c}$, and the risk aversion parameter ρ.

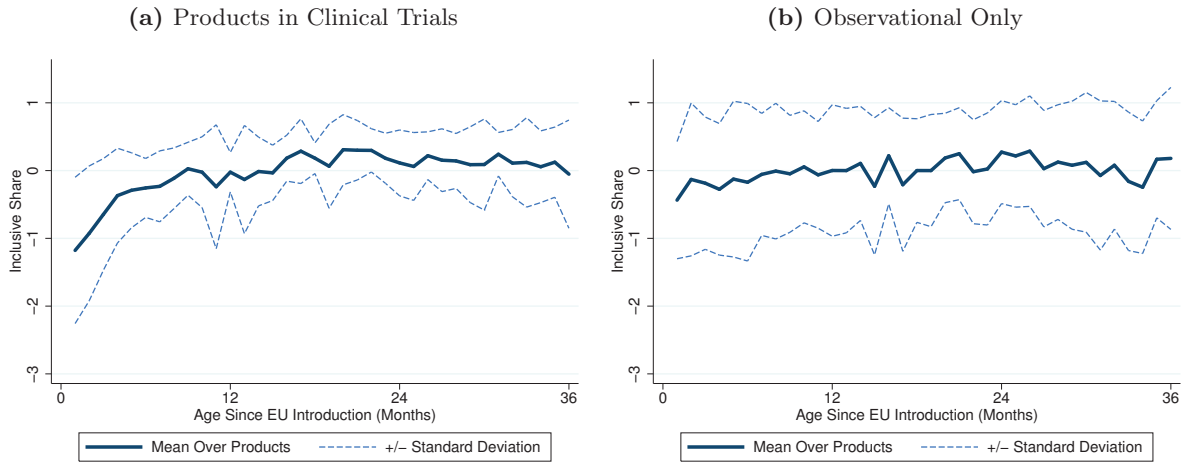
5.1 Identification and Estimation of Demand and Learning

We estimate the parameters via a generalized method of moments algorithm (detailed in Appendix C). A simple and semi-parametric way to estimate Equation (11) would be to regress the inclusive shares $\ln(s_{jt}/s_{0t})$ on product and age fixed effects (age fixed effects interacted with whether a product is in clinical trials or not to allow for differential learning rates from trials and observation). The age fixed effects would then capture the combined effect of learning and risk aversion on utility. However, because we are interested in questions that involve market reactions to different learning rates and levels of risk, we need to add structure via the learning model to disentangle these forces. Comparison to the fixed-effect model provides a useful benchmark for assessing the fit of the more parsimonious and parametric learning model.

As with most empirical learning models, the identification of the signal precision depends on fitting the model to the *shape* of how choice behavior changes with the age of the product. The

risk aversion parameter is then identified as the multiplicative shifter that best fits that shape to the observed choices. In our simple learning model, identification is even clearer because learning is identified by the fact that product-specific quality estimates converge over time. Risk aversion is then identified by how choice probabilities increase (or don't) as learning decreases uncertainty. This can be seen in the first panel of Figure 8, where the distance between the light blue dotted lines—which are each standard deviation of inclusive shares for a given age (net of product fixed effects) away from the mean—decreases with age, identifying learning. And as this variation decreases, the mean inclusive share increases, identifying risk-aversion.

Figure 8: Identifying learning (and risk aversion) for clinical trial vs. observational learning. Plots of mean across products (product means removed) of mean utility estimates $\frac{1}{J_a}\sum_{j=1,a=a}^{J_a}(\ln(s_{jt}/s_{0t}) - Q_j)$—and plus and minus one standard deviation of inclusive share across products—by age since EU introduction. Left panel (a) uses only products undergoing clinical trials for US introduction. Right panel (b) uses all other products, where learning is only observational.



Comparing the two panels in Figure 8 shows how we are able to separately estimate the rates of learning in FDA required clinical trials σ_{A^c} and observationally σ_A because we observe all products post market approval in the EU, and a subset of these products are concurrently involved in clinical trials required for eventual FDA approval. For the products in the right panel where learning is only observational or through non-FDA required trials, there is little if any of the narrowing of variance or increase in mean observed for the products in clinical trials. The learning and risk parameters are estimated using the within-product variation, as they are all conditional on the product fixed effects whose parameters provide estimates of the product qualities Q_j.

We use the empirical distribution of the product fixed effects estimated from the EU data to estimate the mean $\overline{Q}$ and variance σ_Q^2 of the distribution of product qualities developed. This amounts to an assumption that all products that a firm might want to introduce to the market are in fact introduced in the EU. This is plausible as the EU has some products with very low market shares and profits, suggesting that the fixed cost of EU trials and introduction

(conditional on having already developed the innovative product) are quite low.

5.2 Results of Demand and Learning Estimation

The parameter estimates from the model are presented in Table 2. The first observation is that there is meaningful underlying variation in product quality that exposes consumers to risk—at $\sigma_Q = 1.23$ the variation in product quality is nearly as large as the match-specific logit standard deviation of $\pi/\sqrt{6} = 1.28$. These estimates imply that, without information revealed through testing, consumers selecting a new product for insertion face a significant probability that the product is significantly worse than the mean product quality.

Table 2: Structural parameter estimates of demand/learning model: mean over all periods and variance of the product quality distribution $F(Q) \sim N(\overline{Q}_t, \sigma_Q^2)$; precision of learning signals from clinical trials $\sigma_{A^c}^2$ and observational σ_A^2; coefficient of risk aversion ρ in doctor choice behavior.

$\overline{Q}$	σ_Q	$1/\sigma_A^2$	$1/\sigma_{A^c}^2$	ρ
-5.63	1.23	0.01	0.71	4.47
(0.023)	(0.002)	(0.001)	(0.024)	(0.116)

$N = 4490$ product-months. Standard errors clustered by month ($N_T = 114$).

The second observation is that the learning rates vary dramatically according to whether the product is in an FDA required clinical trial or not. Interestingly, the parameter estimates indicate that there is virtually no experiential or non-FDA clinical trial market learning occurring. The estimate of $1/\sigma_A^2$ is an economical and statistical zero. By contrast, the precision of clinical trial learning $1/\sigma_{A^c}^2$ is 0.71, corresponding to approximately 18 months of clinical trials for learning to exceed 95 percent completeness.

5.2.1 Comparison of model estimates to external sources

The model estimates provide several opportunities for validation with other data and research. One such comparison we find particularly reassuring is that the implied coefficient of risk aversion is quite sensible and aligns with the estimates of this parameter from the literature. The parameter estimate in Table 2 is not directly interpretable as it is in logit utility units. However, if we convert that estimate into a dollar equivalent by normalizing the total surplus per stenting procedure to \$50,000 (the estimated dollars in quality adjusted life years from the procedure), then the estimated risk aversion parameter is $\rho_{\$^{-1}} = 1.03 \cdot 10^{-4}$.[15] This is within the range of estimates of risk aversion in well-designed studies such as Cohen and Einav (2007).

Another opportunity to validate our model with external data is to compare the price data to the marginal contributions (sometimes also called added values) implied by our estimated

[15] Because we find that price does not influence stent demand, we do not have the standard price coefficient available to scale demand estimates from logit utils to dollars. Instead we take advantage of the fact that like many medical technologies, the procedure of angioplasty with a stent has been subject to numerous studies attempting to value the average quality adjusted life years added by the procedure in dollar terms. We use \$50,000 (published estimates range from \$32,000 to \$80,000) to calibrate the mean total surplus generated per procedure into dollars. Source: Cost Effectiveness Analysis Registry (https://research.tufts-nemc.org).

demand model, given by $AV_j := TS(\mathcal{J}) - TS(\mathcal{J} \setminus \{j\})$, the increase in total surplus from adding each product j to the choice set. Our demand model estimates imply an average added value of \$1609 for DES and \$831 for BMS. These can be compared to average prices in the data of \$1077 for DES and \$558 for BMS in the EU during our sample. Models of negotiated prices such as the Nash Equilibrium of Nash Bargaining used for stents in Grennan (2013, 2014) suggest that price should be equal to marginal cost plus a markup where the supplier receives a fraction of the added value that depends on his bargaining parameter vs. the bargaining parameter of the buyer. Though we do not specify a full supply side model, for the purpose of comparing our demand estimates to prices, we calculate that setting costs equal to the lowest observed prices of \$571 for DES and \$168 for BMS implies bargaining splits where the supplier captures 31-47 percent of the added value, whichs seem plausible and is close to the range estimated in Grennan (2013, 2014) for the US stent market. Thus our demand model estimates, calibrated into dollars using studies that estimate the dollar value of quality adjusted life years associated with angioplasty with stent procedures, yields both added value and risk aversion measures that fall near those in related literature.

5.2.2 Model fit and comparison with fixed effects least-squares estimates

Figure 9: Comparison of estimates from fixed effect and learning models. Left panel (a) plots the estimated distribution of product qualities from the parametric learning model and age fixed effects model. Right panel (b) plots the estimated discount due to uncertainty versus product age for the two models.

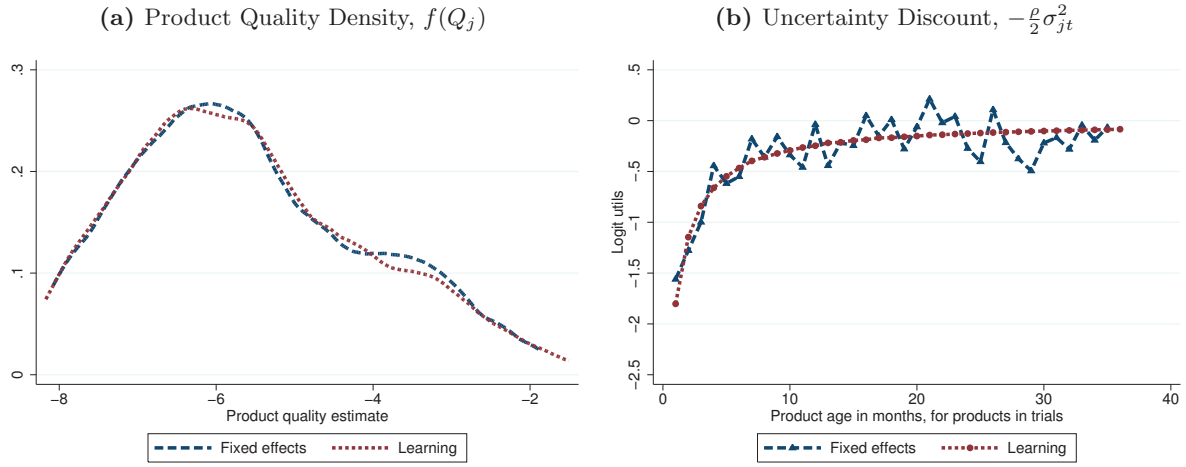


Figure 9 shows the estimated distribution of product qualities Q_j and uncertainty discounts $-\frac{\rho}{2}\sigma_{jt}^2$ for both the learning model and the more flexible model with product and age fixed effects. The message here is that despite its parsimony, the simple learning model fits the data well and its fit is comparable to the much more nonparametric fixed effects model (R^2 of 0.95 vs. 0.98), so our results are not driven by the functional form of the learning model.

5.3 Uncertain Quality and Market Outcomes

Our model implies that imperfect information in conjunction with risk aversion leads to welfare loss. The magnitude of that welfare loss depends upon the mean quality level, the variance in that quality level and the amount of information the consumer possesses. Note that this welfare loss is driven from symmetric yet imperfect information and is not driven by informational asymmetries which is a central concern in much of the quality information literature.

In order to quantify the importance of the imperfect information we examine the welfare and market penetration implications of several different parameters of our model. Table 3 explores the role of uncertainty in the market by using the demand model to calculate total surplus per stent $\frac{TS}{1-s_0}$ (calibrated to be equal to \$50,000 in the observed EU data) and the percent of patients undergoing a diagnostic angiography who choose some stent over the outside good $(1-s_0)$. Here we are positing hypothetical markets where all products have uncertainty in their quality, varying from the unconditional variance of the quality distribution σ_Q^2 (if there was no testing/learning at all), to the estimated uncertainty upon first entering the EU $\sigma^2_{a_{EU}=0}$ (after undergoing EU testing requirements), to varying lengths of US trials $\sigma^2_{T^c}$. The first set of rows show results for the estimated quality distribution in the data (mean $\overline{Q} = -5.63$), and also reports the estimates for the actual EU data, which has quantity-weighted average uncertainty across product months $\overline{\sigma^2_{jt}}|_{EU} = 0.30$, similar to that of $T^c = 2.9$ months of US clinical trials. The subsequent two sets of rows explore the interaction of this uncertainty effect with the level of product quality by shifting the product quality distribution one standard deviation of the logit distribution $\pi/\sqrt{6} = 1.28$ in each direction.

Table 3: The effect of uncertainty and mean quality on surplus per stent and the total number of stenting procedures. Total surplus per stent and share of diagnostic patients receiving stent (vs. alternative non-stent treatment) as a function of mean quality in the market (rows: estimated mean quality and plus or minus one logit standard deviation) and amount of uncertainty (columns: from unconditional product distribution to EU testing to various lengths of US testing).

		$\sigma_Q^2 =$ 1.51	$\sigma^2_{a_{EU}=0} =$ 0.80	$\sigma^2_{T^c=6} =$ 0.18	$\sigma^2_{T^c=12} =$ 0.10	$\sigma^2_{T^c=18} =$ 0.07	$\sigma^2_{T^c=24} =$ 0.05	$\overline{\sigma^2_{jt}}\|_{EU} =$ 0.30
$\overline{Q} = -5.63$:	$\frac{TS}{1-s_0}$ (\$)	44,286	46,241	52,709	54,191	54,829	55,183	50,000
	$1-s_0$ (%)	2	10	31	35	37	38	24
$\overline{Q} = -6.91$:	$\frac{TS}{1-s_0}$ (\$)	43,925	44,487	46,525	47,030	47,251	47,375	
	$1-s_0$ (%)	1	3	11	13	14	15	
$\overline{Q} = -4.35$:	$\frac{TS}{1-s_0}$ (\$)	45,550	51,857	68,709	72,030	73,413	74,170	
	$1-s_0$ (%)	8	29	61	66	68	69	

Table 3 makes several important points. First, the stent market would fail without any learning. This can be seen in the first column of the table in which the percentage of consumers having a stent implanted is close to zero for all values of $\overline{Q}$. Clinical testing provides the necessary information to make this market operate.

Second, increasing the information available to consumers by a modest amount generates significant improvements in welfare. Moving from a world in which there are no clinical trials to one in which there is a FDA required clinical trial of 6 months leads to large increases in the

number of procedures performed and the surplus created per procedure. Increasing the trial length another 6 months generates significantly smaller welfare increases and correspondingly smaller increases in the percentage of patients selecting a stent.

Third, the impact of information is dependent upon the mean quality levels of the stent distribution. In particular, the higher the quality of the average stent, the more valuable is clinical trial information. At first blush, this is somewhat counterintuitive as one make think that if the mean stent is of low quality, avoiding that stent would generate significant improvement in welfare. However, the converse is also true. If the average stent is of high quality, the greater gain to selecting the right stent ex post and thus the more valuable is clinical trial information in our setting.

6 Welfare Implications of Regulatory Policy

With the model and estimated structural parameters, we can examine the impact of different regulatory regimes on welfare. There is a longstanding debate over the appropriate device approval and clearance policy for medical devices and our estimates can shed much needed light on this issue. We examine two different dimensions that could be influenced by regulatory policy: (1) the amount of information T^c that device manufacturers need to generate prior to marketing their products and (2) the precision of observational post-market learning $1/\sigma_A^2$. While the parameter values we explore are within the support of the EU and US data, the role of the model is in predicting the equilibrium responses of firms and consumers at intermediate values that we do not observe.

As discussed in Section 3.3, a fully specified supply model for this market is an undertaking deserving of its own separate paper. Instead of solving the fully dynamic equilibrium which would yield point estimates, we use our model to place upper and lower bounds on implications from the full equilibrium model. First, we compute outcomes in the case where there are no direct fixed costs of longer clinical trials, so all firms enter in equilibrium, and the only role of increasing trial length on market structure is to delay access to the newest technologies (in addition to increasing learning). This represents an upper bound on the total surplus generated under any clinical trial requirement. Next, we compute outcomes assuming that the cost of trials is \$1.6M per month, but with firms' entry decisions based on realized EU profits (Makower et. al. (2010) survey that reports the average pivotal trial required by the FDA to cost 1.6 million dollars per month). Because this doesn't allow expected market shares and prices to increase as fixed costs increase and the market becomes more concentrated, this represents a lower bound on the total surplus generated under any clinical trial requirement.

6.1 Optimal Premarket Clinical Testing

The first exercise we perform is examining the optimal regulatory standard for clinical trial length. This addresses a fundamental question facing any industry where new products are developed with uncertain quality and safety: How much testing is enough? Answering this

question requires understanding the consequences of alternative testing requirements. One way to summarize these consequences is to plot the expected surplus generated versus length of trial required.

Figure 10: Optimal Regulation: Plot of total surplus per patient (measured in the percent change from EU benchmark) versus length of (US sized) clinical trial required (in addition to EU requirements). The two cases plotted provide bounds for the full dynamic equilibrium: The case with zero direct costs of trials provides an upper bound where all products enter and delay is the only cost of longer trials. The case with trial costs but where entry decisions are based on realized EU profits provides a lower bound where firms do not take increased market power into account as entry costs rise. The black vertical lines indicate the identified set of optimal trial lengths. 95 percent confidence intervals, clustered by month, provided by dotted lines.

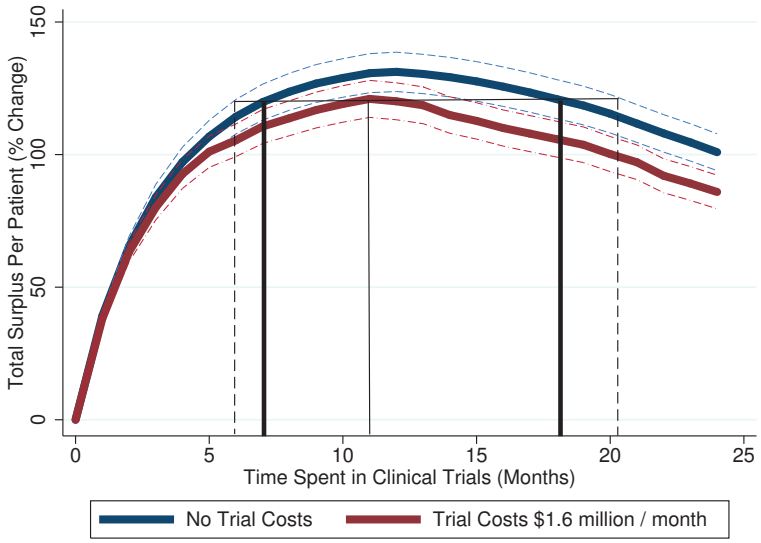


Figure 10 does just this, plotting expected total surplus per patient treated in the market versus the required length of time spent in clinical testing (relative to the current EU required clinical testing). More specifically, Figure 10 plots our estimates of the upper and lower bounds with the corresponding confidence intervals, and uses these to place bounds on the optimal regulatory policy. The results suggest that the optimal tradeoff of access vs. risk is reached between $T_c^* = [7, 19]$ months of premarket clinical testing. An interesting feature of the estimated total surplus as a function of time in premarket clinical testing is that it is relatively flat for a wide range of trial lengths near the optimum.

Outside of the flat range, however, surplus drops rapidly with zero month trials resulting in a 110 percent drop in surplus relative to the optimal. At first this seems to suggest that the EU could make extremely large welfare gains by increasing its standards—until one realizes that the EU is able to free-ride off of the information being generated in trials for US entry. In effect, the EU is getting free post-approval learning which makes it's market look closer to one with US trials of approximately 2.9 months from a risk perspective. That is, conditional on the US regulatory policy, the gain to the EU from increasing its standards is more modest,

but still substantial at 20 percent or more.

Recall that in our data the mean lag between US and EU entry is 10 months for all products and 17 months for DES. That is, we find that the current FDA policy falls within our confidence interval for the optimal policy conditional on the rate of observational learning. This result speaks directly to the current policy debate over the FDA medical device approval pathway and supports the FDA argument that reductions in their standards for device approval will reduce consumer welfare. Though of course we are comparing a different time and product market, our results stand in contrast to the Peltzman (1973) influential analysis of the 1962 Kefauver-Harris Amendments to the Food, Drug, and Cosmetic Act which required proof of efficacy and made the testing procedures required to prove that efficacy subject FDA oversight. He concludes that the Kefauver-Harris Amendments led to a significant decrease in welfare. Peltzman's analysis, however, does not speak to the optimal informational requirements pharmaceutical manufacturers should face when introducing a new molecular agent. To the best of our knowledge, the our analysis is the first that provides an estimate of the optimal policy on the amount of information creation.

6.2 Alternative Policy: Shorter Trials with Increased Post-market Learning

We estimated the post market approval observational learning rate is zero for the set of products in our data. There are several potential reasons for the lack of post-market approval learning. For some products, observational learning from real world use may make it difficult to infer product quality (not having the randomization into treatment and control as in a clinical trial). For other products, though—and likely for those in our sample—the problem is simply a lack of systematic data collection and sharing of information.

One frequently proposed change to FDA regulatory policy is to relax requirements on premarket clinical trials but increase requirements on post-market surveillance, including data collection, analysis and reporting. This policy has a direct connection to our model in the sense that it's intention is to increase the rate of post-market approval observational learning—in the language of our model, this means increasing the precision $1/\sigma_A^2$ of the signals that arrive outside of FDA required clinical trials. We analyze this policy by taking the estimated model, varying σ_A, and calculating the corresponding optimal trial length $T_c^*(\sigma_A)$ and total surplus generated $TS(\sigma_A, T_c^*(\sigma_A))$. Figure 11 displays the results, again using our bounds to generate a partially identified set of predictions.

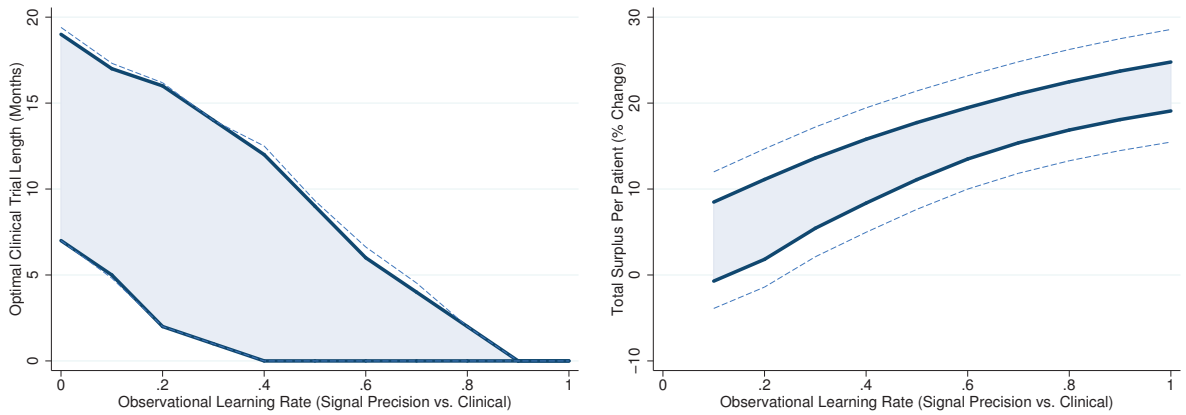
When observational learning is as fast as FDA clinical trial learning, there is no reason to run clinical trials at all. Total surplus is highest—24 percent higher than with no observational learning—because there is no tradeoff to be made between access and learning. The value of this increase is very large. Using baseline estimates of utilization and a value of $50,000 per treatment yields an estimate of $7.6 billion per year in increased welfare from this increase in post-market learning in the US.[16]

Not surprisingly, as the precision of observational learning decreases (relative to clinical

[16] In 2009, over 640,000 stent procedures were performed in the US (Auerbach, 2012).

Figure 11: The Value of Post-Market Surveillance: Plots of optimal trial length (left panel (a)) and total surplus (right panel (b)) as observational learning precision $1/\sigma_A^2$ varies from zero to the clinical trial precision $1/\sigma_{A^c}^2$. 95 percent confidence intervals, clustered by month, provided by dotted lines.

(a) Optimal Trial Length, $T_c^*(\sigma_A)$ **(b)** Total Surplus, $TS(\sigma_A, T_c^*(\sigma_A))$



trial learning), it becomes optimal to require longer clinical trial periods prior to market access in order to take advantage of the faster learning rate of clinical trials. The lesson from this policy experiment is that there is merit to the argument that a requiring shorter trials with post-approval testing would improve consumer welfare. The gains from this policy critically depend on the rate and cost of learning via post-market surveillance.

6.3 Discussion and Implications

As discussed in the Introduction, there is an important literature measuring the optimal regulatory policies across settings as diverse as pharmaceuticals (Filson 2012; Budish, Roin, and Williams 2015), liquor distribution (Seim and Waldfogel 2013; Miravete, Seim, and Thurk 2014), and water management (Timmins 2002). In general, the estimates show that the regulator's behavior departs from the socially optimal policy by anywhere from 10 to 50 percent. Three of our main results add to this literature: First, to the best of our knowledge, our work is the first to find a regulator (the FDA) whose policies may in fact be maximizing total surplus. Second, our results are consistent with the previous literature in that we find an EU regulatory regime that is suboptimal. The EU could meaningfully increase welfare by increasing the informational criteria required to receive market access. Third, we show that adding another dimension of regulatory policies that improve market learning and reduce pre-market clinical requirements could dramatically increase social welfare.

It is important to note that our analysis holds the technology fixed and abstracts away from the feedback effects of regulatory approval regime to firms incentive to invest in new products. However, an important takeaway from our analysis is that the value of a technological innovation to the marketplace depends to a large extent on the regulatory regime's informational

requirements for product testing. In fact, our estimates of the value of information for medical technologies is large and comparable in magnitude to the estimates of the value of innovation itself. The welfare implications of medical technology innovation has long played a central role in health economics, but there is much less of an emphasis in the literature on the value of information regarding these new technologies. Murphy and Topel (2006) show that medical technology innovation has lead to massive improvements in welfare over the twentieth century. They find gains on the order of $1.2 million per representative American in 2000. More recently, Budish, Roin, and Williams (2015) estimate that increasing effective patent lengths for cancer drugs would yield an $89 billion increase in welfare for patients diagnosed in 2003 by inducing more investment in treating cancers where the effective patent life on new drugs is relatively short. Coronary stents treat a narrower set of conditions than cancer drugs; but scaled for market size, our finding that increasing post-market learning rates can increase welfare $7.6 billion per year suggest that the role of information can be comparable to the role of new technology innovation in affecting welfare.

Thus a broader takeaway from our research is that the innovation process should be considered holistically from idea to consumer—the value of innovations can be significantly enhanced or diminished by the information regulators require technology firms to produce and disseminate. In the case we study, the availability of new medical technologies with uncertain quality can only achieve their welfare potential if firms undertake the necessary studies to document the product's clinical performance. For coronary stents, the market would shut down without at least some initial testing as required by the EU, and benefits even more from the further testing required by the US. Thus it seems that there are important complementarities between the value of new medical technologies and the regulatory approval product regime, and our work provides one of the first quantifications of that value.

7 Conclusion

The tradeoff between access and risk in regulating the market entry of new products is important in a variety of industries, and in particular in medical devices, where it is an active topic of policy debate in almost every country in the world. In this paper we develop a model with products introduced when quality is still uncertain, learning over time, and regulator (and manufacturer) decisions regarding product testing and market entry. We show that the empirical predictions of the model are borne out in market share data in the US and EU medical device markets and are consistent with the beliefs that the US regulatory environment is more restrictive than the EU. We then estimate the structural parameters of the model for use in welfare analysis of policy analyses affecting: (1) the length of clinical trials required before market entry and (2) observational learning after market entry.

For the set of devices on which we have data, we estimate that the US is close to the optimal policy, but the EU is too lax (despite free-riding off of information generated by US trials). We also estimate that if it is possible to achieve post-market learning rates close enough to those we observe from clinical trials at a comparable cost, then embracing recent calls for more active

"post-market surveillance" could further increase total surplus by as much as 24 percent.

Of course, our analysis is limited in the set of devices for which detailed market data is available, and extrapolating to policy for all devices should be done with care. The theoretical model we develop provides some guidance for how this extrapolation should depend on the uncertainty in quality of new product introductions, the rate of technological improvement, the learning rate in clinical trials, and the observational learning rate for any type of device being considered. Because the model is quite general and flexible, and the type of data we use is available for many markets, we hope that we have provided a starting point for analysis of regulation and market structure in other industries where new product development and testing play an important role.

We also hope that we have provided a building block that, in future research, could be used to provide a more complete picture of how regulation affects market structure, innovation, and ultimately welfare. While estimating the welfare effects of the access/risk tradeoff for an exogenously given set of innovations is an important step towards better understanding this phenomenon, a more complete understanding would allow for the regulatory regime to effect the types of innovations firms develop for the market and vice-versa. A more dynamic analysis of this type would require a significant extension to the theory, and would also require detailed data on innovative activities of the firms in a market.

References

Ackerberg, D. A. (2003). Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination. *International Economic Review*, 44(3):1007–1040.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *RAND Journal of Economics*, 25(2):242–262.

Budish, E., Roin, B., and Williams, H. (2015). Do fixed patent terms distort innovation? evidence from cancer clinical trials. *American Economic Review*, forthcoming.

Ching, A. T. (2010). A dynamic oligopoly structural model for the prescription drug market after patent expiration. *International Economic Review*, 51(4):1175–1207.

Cohen, A. and Einav, L. (2007). Estimating risk preferences from deductible choice. *The American Economic Review*, pages 745–788.

Crawford, G. S. and Shum, M. (2005). Uncertainty and learning in pharmaceutical demand. *Econometrica*, 73(4):1137–1173.

Dranove, D. and Jin, G. (2010). Quality disclosure and certification: Theory and practice. *Journal of Economic Literature*, 48(4):935–63.

Erdem, T. and Keane, M. P. (1996). Decision-Making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets. *Marketing Science*, 15(1):1–20.

Filson, D. (2012). A markov-perfect equilibrium model of the impacts of price controls on the performance of the pharmaceutical industry. *RAND Journal of Economics*, 43(1):110–138.

Grennan, M. (2013). Price discrimination and bargaining: Empirical evidence from medical devices. *American Economic Review*, 103(1):145–177.

Grennan, M. (2014). Bargaining ability and competitive advantage: Empirical evidence from medical devices. forthcoming in Management Science.

Lee, R. S. and Fong, K. (2013). Markov-perfect network formation an applied framework for bilateral oligopoly and bargaining in buyer-seller networks. *working paper*.

Makower, J., Meer, A., and Denend, L. (2010). FDA impact of us medical technology innovation: A survey of over 200 medical technology companies. Stanford University Report.

Miravete, E. J., Seim, K., and Thurk, J. (2014). Complexity, Efficiency, and Fairness in Multi-Product Monopoly Pricing. Technical report, Working paper.

Murphy, K. and Topel, R. (2006). The value of health and longevity. *Journal of Political Economy*, 114(3):871–904.

Peltzman, S. (1973). An evaluation of consumer protection legislation: The 1962 drug amendments. *Journal of Political Economy*, 81(5):1049–1091.

Roberts, J. H. and Urban, G. L. (1988). Modeling multiattribute utility, risk, and belief dynamics for new consumer durable brand choice. *Management Science*, 34(2):167–185.

Seim, K. and Waldfogel, J. (2013). Public monopoly and economic efficiency: Evidence from the pennsylvania liquor control board's entry decisions. *American Economic Review*, 103(2):831–62.

Timmins, C. (2002). Measuring the Dynamic Efficiency Costs of Regulators' Preferences: Municipal Water Utilities in the Arid West. *Econometrica*, 70(2):603–629.

Appendices

A Theory Appendix

A.1 Supply Bounds on Total Surplus

As discussed in Section 3.3 of the paper, developing a full supply model of entry/exit, contracting, and pricing would add frontier modeling efforts in both the bargaining and dynamic oligopoly literatures that would distract from the current focus on information and the tradeoff between access and risk in regulating new products when quality of the innovations are uncertain. Instead, we develop bounds on the total surplus $TS(T^c)$ generated under any regulatory policy required pre-market clinical trials of length T^c.

The key to constructing these bounds is that we only consider counterfactuals where the entry costs are greater than those required to obtain EU approval, and we assume that EU approval costs are sufficiently low that all potential entrants find it profitable to enter under the current EU regime.

Proposition UB (Upper Bound on $TS(T^c)$): The total surplus generated in equilibrium when there are no direct fixed costs of longer clinical trials ($\phi^e(T^c) = 0, \forall T^c \geq 0$) provides an upper bound for the surplus generated in equilibrium when entry costs are increasing in trial length ($\phi^e(T^c + 1) > \phi^e(T^c), \forall T^c \geq 0; \phi^e(0) = 0$).

In this case in any period t we have:

$$\begin{aligned}
TS_t^{UB}(T^c) &:= TS_t(0, T^c) \\
&= \ln\left(\sum_{j \in \mathcal{J}_t(0,T^c)} e^{Q_{jt}(T^c) - \frac{\rho}{2}\sigma_{jt}^2(T^c)}\right) \\
&\geq \ln\left(\sum_{j \in \mathcal{J}_t(\phi^e(T^c),T^c)} e^{Q_{jt}(T^c) - \frac{\rho}{2}\sigma_{jt}^2(T^c)}\right) \qquad (12) \\
&\geq \ln\left(\sum_{j \in \mathcal{J}_t(\phi^e(T^c),T^c)} e^{Q_{jt}(T^c) - \frac{\rho}{2}\sigma_{jt}^2(T^c)}\right) - \phi^e(T^c)|\mathcal{J}_t \setminus \mathcal{J}_{t-1}| \qquad (13) \\
&= TS_t(\phi^e(T^c), T^c).
\end{aligned}$$

The inequality in (13) holds for any nonnegative fixed cost of trials. The inequality in (12) obtains from the fact that the choice set under nonnegative fixed costs is a weak subset of the choice set under no fixed costs $\mathcal{J}_t(\phi^e(T^c), T^c) \subseteq \mathcal{J}_t(0, T^c)$ (and the fact that in our model the utility parameters are not a function of the choice set itself). That this will hold weakly is certain due to the assumption that all potential entrants enter in the observed EU regime (which has cost zero and also $T^c = 0$), and so in expectation all of these same products would enter in the upper bound case. The subset will be strict in

the case that $\phi^e(T^c)$ is large enough so that some product j finds it unprofitable to enter. *Q.E.D.*

Proposition LB (Lower Bound on $TS(T^c)$): The total surplus generated in equilibrium when entry costs are increasing in trial length ($\phi^e(T^c+1) > \phi^e(T^c), \forall T^c \geq 0; \phi^e(0) = 0$) is bounded from below by the total surplus generated in equilibrium with these same fixed costs, but where firms follow a naive entry strategy that assumes all other firms will enter in equilibrium.

In this case in any period t we have:

$$\begin{aligned} TS_t^{LB}(T^c) &:= TS_t((\phi_j^e(T^c), 0_{-j}), T^c) \\ &= \ln\left(\sum_{j\in\mathcal{J}_t((\phi_j^e(T^c),0_{-j}),T^c)} e^{Q_{jt}(T^c)-\frac{\rho}{2}\sigma_{jt}^2(T^c)}\right) - \phi^e(T^c)|\mathcal{J}_t(j,-j)\setminus\mathcal{J}_{t-1}(j,-j)| \\ &\leq \ln\left(\sum_{j\in\mathcal{J}_t(\phi^e(T^c),T^c)} e^{Q_{jt}(T^c)-\frac{\rho}{2}\sigma_{jt}^2(T^c)}\right) - \phi^e(T^c)|\mathcal{J}_t\setminus\mathcal{J}_{t-1}| \qquad (14) \\ &= TS_t(\phi^e(T^c), T^c). \end{aligned}$$

The inequality in (14) obtains from the fact that the choice set under the full equilibrium is a subset of the choice set under the naive equilibrium and that there is never over entry in either equilibrium (the social benefit of firm entry is always greater than the fixed cost of entry). Under the assumption that all potential products entered in the zero cost case, potential profits (after fixed costs) will always be weakly higher in the full equilibrium, increasing the probability of entry. Under the additional assumption that there is never over entry (which would be the case with any pricing model where a product cannot capture more than it's marginal contribution), the surplus gain is always greater than the fixed cost of entry. *Q.E.D.*

These upper and lower bound scenarios are equivalent to each other and to the full equilibrium at $T^c = 0$. Both will become further from the true equilibrium as the costs of entry increase.

A.2 Total Surplus with No Fixed Costs or Observational Learning

$$\begin{aligned} TS_t(T^c+1) - TS_t(T^c) &= \ln\left(\frac{\sum_{j\in\mathcal{J}_t(T^c+1)} e^{Q_{jt}-\frac{\rho}{2}\sigma_{jt}^2(T^c+1)}}{\sum_{j\in\mathcal{J}_t(T^c)} e^{Q_{jt}-\frac{\rho}{2}\sigma_{jt}^2(T^c)}}\right) - \phi^e\left|\mathcal{J}_t^e(T^c+1)\setminus\mathcal{J}_t^e(T^c)\right| \\ &= \ln\left(\frac{\sum_{j\in\mathcal{J}_t(T^c+1)} e^{Q_{jt}-\frac{\rho}{2}\sigma_{jt}^2(T^c+1)}}{\sum_{j\in\mathcal{J}_t(T^c)} e^{Q_{jt}-\frac{\rho}{2}\sigma_{jt}^2(T^c)}}\right) \qquad (15) \\ &= \frac{\rho}{2}\left(\sigma^2(T^c) - \sigma^2(T^c+1)\right) - \ln\left(\frac{\sum_{j\in\mathcal{J}_t(T^c)} e^{Q_{jt}}}{\sum_{j\in\mathcal{J}_t(T^c+1)} e^{Q_{jt}}}\right) \qquad (16) \end{aligned}$$

where (15) follows from no fixed costs, and (16) follows from no observational learning. Then averaging over any period of time $t = 1, ..., T$ and recognizing $\phi^e = 0 \Rightarrow \mathcal{J}_t(T^c) = \mathcal{J}_{t+1}(T^c + 1)$ so that the log sum term is telescoping yields (9).

B Data Appendix

B.1 Dataset construction

The dataset used in this paper is from Millennium Research Group's Marketrack survey of catheter labs, the source that major device manufacturers subscribe to for detailed market research. The goal of the survey is to provide an accurate picture of market shares and prices of medical devices. For our purposes, the key variables in the data are the price paid and quantity used for each stent in each hospital in each month. In addition, the hospitals report monthly totals for different procedures performed, such as diagnostic angiographies. The data span January 2004 through June 2013 and cover the U.S. and E.U. markets.

There are three main challenges in constructing a usable dataset from the raw survey data. First, the survey was not as concerned with collecting price data as it was with collecting quantity data. Second, the survey measures stent usage rather than availability and our data go back only to 2004, so it is not always possible to infer regulatory approval dates from the data. Finally, there is some misreporting in the survey. The following tables illustrate how key sample summary statistics compare across the cleaning steps for the E.U. and U.S. datasets. These steps are summarized below; full detail can be found in the Stata code used to execute them, `cleaning-eu-data-3-sample.do` and `cleaning-us-data-3-sample.do`.

EU dataset modifications

	Diagnostic procedures	No. of stents implanted	No. of BMS products	No. of DES products	Average stent age	Stent-hospital-months	Hospital-months	Hospitals
Raw data	151	108	3.8	3.3	54.3	88,144	15,064	542
Rm. suspect q	161	98	3.3	2.8	54.5	61,098	13,477	540
Rm. if q¿2*diagnostics	152	107	3.8	3.3	54.3	86,672	14,812	537
Rm. suspect diagnostics	151	108	3.8	3.3	54.4	87,349	14,933	542
Rm. outlier p	148	106	3.8	3.3	54.4	81,646	14,149	532
Rm. unknown entry	150	108	3.8	3.3	54.0	87,516	14,995	541
Final sample	160	95	3.2	2.8	54.6	54,771	12,313	524

US dataset modifications

	Diagnostic procedures	No. of stents implanted	No. of BMS products	No. of DES products	Average stent age	Stent-hospital-months	Hospital-months	Hospitals
Raw data	137	76	2.2	2.5	36.8	68,603	17,183	526
Rm. suspect q	147	68	1.9	2.1	37.8	44,218	14,631	509
Rm. if q¿2*diagnostics	138	76	2.2	2.5	36.7	67,783	16,982	517
Rm. suspect diagnostics	138	76	2.2	2.5	36.8	67,857	16,997	526
Rm. outlier p	136	75	2.2	2.5	37.1	66,293	16,720	525
Final sample	147	67	1.8	2.1	38.0	41,779	13,900	478

The table rows record the sample means for key summary statistics across various cleaning steps. The summary statistics are means of quantities calculated at the hospital-month level. The means reported are of the total number of stents implanted; the total number of diagnostic angiographies; the number of different bare-metal stents (BMS) used; the number of different

drug-eluting stents (DES) used; and the weighted average age, in months, of the stents used. The table also shows the total number of stent-hospital-month observations, number of hospital-month observations, and number of hospitals in each sample.

The table rows correspond to different samples. The first row of each table summarizes the raw E.U. and U.S. survey data. The second row drops hospital-months with suspect total quantities. The criteria for dropping are threefold. First, we drop hospital-months for which the total quantity of stents changes by (weakly) more than 50% relative to the previous month in which the hospital appears in the data. Second, for "low-quantity" hospitals with mean monthly stent quantities below 15, we drop hospital-months with usage strictly greater than 1.5 standard deviations from the hospital's mean. For "high-quantity" hospitals with mean monthly stent quantities (weakly) greater than 15, we drop hospital-months with usage strictly greater than 3.0 standard deviations from the hospital's mean. Third, for hospital-months with flagged quantity changes that were accompanied by a 30% or greater change in diagnostic angiography procedures, the hospital-months were undropped. Diagnostic angiography procedures are performed prior to coronary stent implantation, so large changes in monthly stent quantities should be accompanied be similarly large changes in angiographies.

The third and fourth rows of the table drop hospital-months with suspect diagnostic angiography counts. Diagnostic angiographies should be bounded below by some multiple of the number of stents used; in our data and anecdotally according to clinicians, there are at least two diagnostic angiography procedures per stent implant. The third row drops hospital-months if their total quantity of stents exceeds twice the number of diagnostic angiographies in that hospital-month. The angiography count itself could be suspect. The fourth row drops hospital-months if the number of diagnostic angiographies is more than 2 standard deviations away from the hospital's mean and if the ratio of angiographies to stents was 2 standard deviations from the hospital's mean.

The fifth row of the table drops hospital-months with problematic prices. We drop hospital-months with outlier prices based on a regression of log-price on the hospital's number of BMS products and number of DES products used that month, in addition to a hospital fixed-effect. Hospital-months with products whose regression residuals were more than 2 standard deviations from the mean of all residuals were dropped.

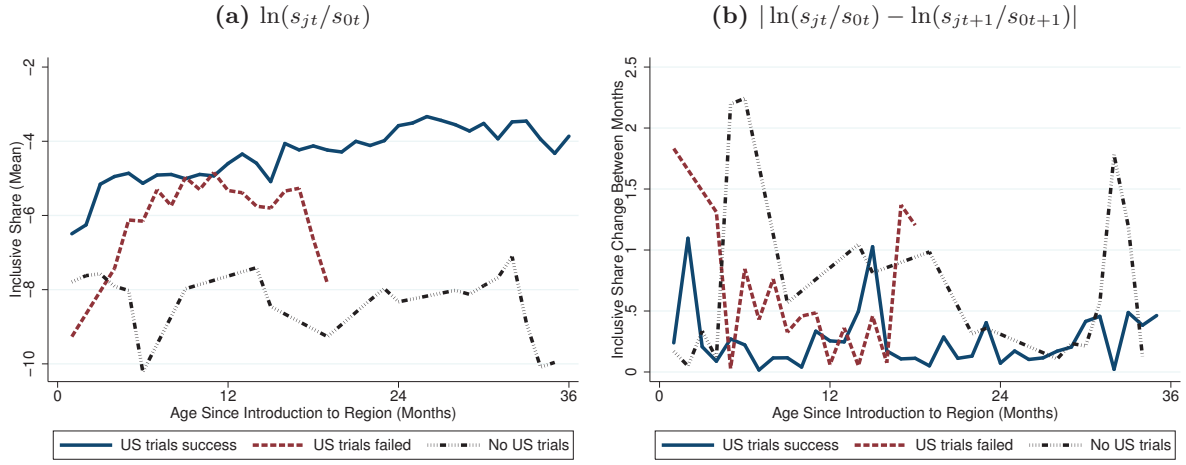
The sixth and penultimate row of the E.U. table drops hospital-months with positive quantities for stents for which E.U. regulatory approval dates are not known. Since the age of the product is an important component of our analysis, the products for which an entry date could be pinned down with reasonable certainty must be removed from the analysis. This drop affects only a few products. There are no products for which the U.S. approval dates could not be ascertained, so this row is missing from the U.S. table.

The final row in each table reports summary statistics for the final sample, which drops all observations that meet one or more of the dropping criteria described above.

B.2 Evidence of learning from individual products

Averaging across products conditional on age provides patterns in the data that have direct relation to expected patterns in our model. However, these averages cloud heterogeneity in learning across products. A look at the patterns for a few individual products illustrate how learning does not always bring good news and lack of learning brings a volatile mix of good and bad over time.

Figure 12: Learning patterns for selected individual products. Three representative products that receive good and bad news from trials or not much (useful) news at all. Left panel (a) plots mean utility estimate for each product $\ln(s_{ja}/s_{0a})$ by age since introduction into the EU. Right panel (b) plots absolute differences $|\ln(s_{ja}/s_{0a}) - \ln(s_{ja+1}/s_{0a+1})|$ by age, which should be larger with more uncertainty, and converge toward zero with learning.



B.3 Distribution of Profits Over Product Lifetime and Across Products

The counterfactual lower bounds with fixed costs of entry require calculation of expected lifetime profits under the assumption that all firms who enter in the EU do enter in equilibrium. This number can be directly acquired from the EU data for the 41 of 109 products that both enter and exit the market during our sample period. However, for the other 68 products whose lifetimes are truncated at the beginning or end, we need to extrapolate.

Table 4: Product age and profitability.

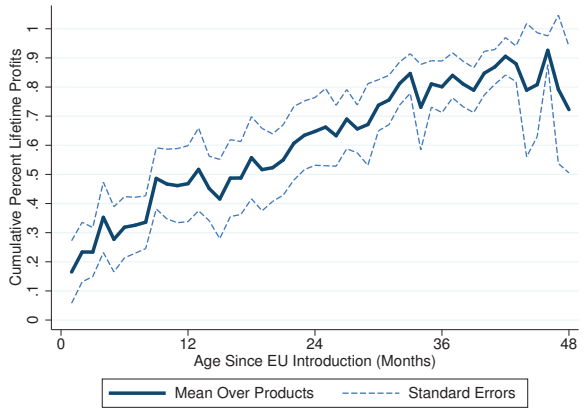
	mean	s.d.	10ptile	median	90ptile	N
Products with full lifetime during sample period:						
Months in sample	21.5	19.8	5	15	47	41
Profit per month ($1000s)	179	612	18	70	211	41
Products with censored lifetime:						
Months in sample	53.1	34.9	10	46	100	68
Profit per month ($1000s)	1,347	2,119	41	262	4,067	68

We perform this extrapolation by estimating the percent of cumulative lifetime profits the

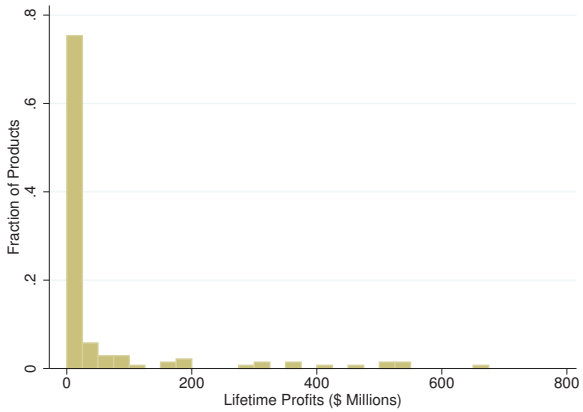
average product has earned at each age. We then use this percent to extrapolate the missing profits, for whatever age at which the truncation occurred. We do this unconditionally on any covariates besides age. Our counterfactual estimates are robust to a variety of approaches to this extrapolation. This is in part because the extrapolation is typically for the beginning or end of lifetime tail of product profits, so that lifetime profit projections are not very sensitive to the method we choose. And further the products that are marginal in our counterfactuals, in the sense that they exit as entry costs increase, are also marginal in the computation of consumer surplus in that which enter (and even to some extent how many enter) does not greatly affect total welfare.

Figure 13: Distribution of Profits Over Time and Across Products.

(a) Cumulative Profits Over the Product Lifecycle



(b) Distribution of Lifetime Profits over Products



C Estimation and Counterfactual Algorithms

C.1 Demand/learning estimation algorithm

The estimation approach is to construct a generalized method of moments estimator that matches the observed market shares in the data (and knowledge of which products are in clinical trials when) to the demand and learning model. The Matlab code for this estimator is available in the electronic archive *code4RegulatingInnovation.zip*. This appendix outlines the main steps of the algorithm.

1. Compute mean utilities $\delta_{jt} = \ln(s_{jt}/s_{0t})$ for all product-months.
2. Construct an initial estimator for uncertainty immediately after EU testing $\sigma^2_{ja=1}$ using the empirical equivalent from the distribution of δ_{jt}.
3. Guess initial values for learning precisions $\sigma_{\mathbf{A}} := (\sigma_A, \sigma_{A^c})$.

4. Compute the full vector of σ^2_{jt} implied by $\sigma^2_{ja=1}$, the learning precision parameters, and which products are in trials when.

5. Least squares then gives an estimator for ρ and the product qualities Q_j as a function of the learning parameters, where $[Q_j; \rho](\sigma_{\mathbf{A}}) = (X'X)^{-1}X'\delta$ with $X = [1_j, -\frac{1}{2}\sigma^2_{jt}]$. (Here Q_j represents the vector of coefficients on product dummy variables, and 1_j the matrix of product dummy variables.)

6. We need to make sure that the distribution of $Q_{jt} := \delta_{jt} + \frac{\rho}{2}\sigma^2_{jt}$ is consistent with the $\sigma^2_{ja=1}$ by recomputing $\sigma^2_{ja=1}(Q_{jt})$ and repeating 4-6 until $\sigma^2_{ja=1}$ converges.

7. Compute the residuals $\xi_{jt} = \delta_{jt} - Q_j + \frac{\rho}{2}\sigma^2_{jt}$.

8. Evaluate GMM objective function based on $E[\xi' Z] = 0$ where $Z = \left[\frac{1}{a_{jt}} \;\; \frac{1}{a^2_{jt}}\right]$.

9. Repeat 4-8 until we find the value of $\sigma_{\mathbf{A}}$ that minimizes the GMM objective function.

C.2 Optimal regulation counterfactual algorithm

The advantage of the upper and lower bounds we have defined on total surplus is that they can be calculated using only the data and demand/learning model estimates. For each potential $T^c = 0, 1, ..., 24$ we calculate the upper and lower bounds as follows:

Upper Bound

1. Given T^c, restrict sample to products that would be active in each month.

2. Use demand/learning model to compute total surplus over the sample period.

Lower Bound

1. Given T^c, restrict sample to products that would be active in each month.

2. Given $\phi^e = T^c \times 1.6E6$, restrict sample to products that would enter, under the naive assumption that firms assume other products enter as if $\phi^e = 0$.

3. Use demand/learning model to compute total surplus over the sample period, remembering to subtract fixed costs when products enter.

C.3 Observational learning counterfactual algorithm

Because we only specify bounds on total surplus for any trial length T^c, we obtain only bounds on the optimal trial length under any parameter values $[T^c_{LB}(\sigma_{\mathbf{A}}), T^c_{UB}(\sigma_{\mathbf{A}}),]$. Thus for each potential value of observational learning precision $1/\sigma^2_A = 0, 1/10\sigma^2_{A^c}, 2/10\sigma^2_{A^c}, ..., 1/\sigma^2_{A^c}$ we calculate the bounds on optimal trial time and surplus generated by these trial times as follows:

1. Given $1/\sigma_A^2$, calculate the upper and lower bounds on surplus generated for $T^c = 0, 1, ..., 24$ as done previously for the zero observational learning case.

2. $T^c_{LB}(\sigma_{\mathbf{A}})$ will be the maximum T^c such that the upper bound total surplus is less than the maximum of the lower bound total surplus (among the T^c below that at which the lower bound surplus is maximized).

3. $T^c_{UB}(\sigma_{\mathbf{A}})$ will be the minimum T^c such that the upper bound total surplus is less than the maximum of the lower bound total surplus (among the T^c above that at which the lower bound surplus is maximized).

4. The tightest bounds on surplus created in this case are simply the max of the upper bound surplus and max of the lower bounds surplus.

D Robustness Checks

D.1 Demand regressions with price

Price is neither a statistically nor an economically significant determinant of demand. The table below reports coefficients on price from a regression of product-month–level stent utilization on price, product fixed effects, and product age (months since market entry) fixed effects. The coefficient on price is reported in the first row of the table. It is consistently small (relative to the fixed coefficients, not shown) and insignificant.

The first column reports OLS results. Column 2 reports IV results, with price instrumented by one-month price lags at the product level. Column 4 reports IV results, with price instrumented by the total number of other stents on the market that month and the number of other stents in the product's DES/BMS category. Finally, column 6 reports IV results with both sets of instruments. All IV specifications pass the weak instrument test. The specification using number of competitors as instruments yields a price coefficient point estimate that is almost identical to the economically small estimate in Grennan (2013), but price does not reach statistical significance in any of the specifications using this data.

Table 5: Fixed effects demand models with price

	OLS	IV1	FS1	IV2	FS2	IV3	FS3
Price (1,000s)	0.01	-0.01		-0.28		-0.01	
	(0.09)	(0.18)		(0.98)		(0.18)	
Instruments:							
Lagged price (1,000s)			0.45***				0.45***
			(0.03)				(0.03)
No. competing stents					-0.0021		-0.0005
					(0.0015)		(0.0014)
No. DES/BMS					0.0074***		0.0038*
					(0.0019)		(0.0017)
Observations	4490	4437	4437	4490	4490	4437	4437

Standard errors in parentheses

FS columns report first-stage results for the preceding IV columns.

IV1: lagged price. IV2: number of competing stents. IV3: lagged price and number of competing stents.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$