ONLINE APPENDIX

Regulating Innovation with Uncertain Quality: Information, Risk, and Access in Medical Devices

Matthew Grennan and Robert J. Town

Data Appendix

A1. 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 US and EU 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 (and while our independent research found most introduction dates, we were not able to find all). Finally, there is some apparent misreporting in the survey. The following tables illustrate how key sample summary statistics compare across our cleaning steps for the EU and US 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`.

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 EU and US 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 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

Table A1—: Dataset modifications and effect on sample

EU dataset modifications

	Diagnostic procedures	Stents implanted	BMS $\overline{\lvert\mathcal{J}\rvert}$	DES $\overline{\lvert\mathcal{J}\rvert}$	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	Stents implanted	BMS $\overline{\lvert\mathcal{J}\rvert}$	DES $\overline{\lvert\mathcal{J}\rvert}$	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

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 on average about two stents implanted per procedure. The third row drops hospital-months if the number of diagnostic angiographies is less than two times the number of stents implanted in that hospital-month. 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, none of which were frequently used. There are no products for which the US approval dates could not be ascertained.

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.

A2. Clinical trial data

Our collected clinical trial data, and a detailed document on the sources, are available in the online archive and upon request from the authors.

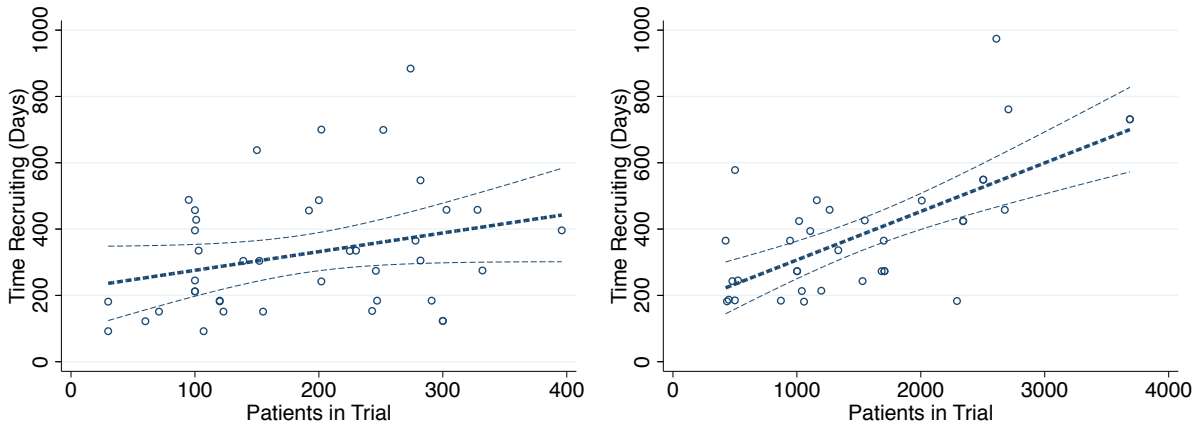
In addition to clarifying the differences between EU and US trial policy and validating our product quality estimates, the trial data make clear the strong relationship between the size of clinical trial in terms of patients and the time spent on the trial via the time it takes to recruit patients. Figure A1 plots the data on patients and length of recruitment in days for smaller and larger trials (broken down to roughly correspond to the scale of trials required for "EU" and "US" approval). One can see from the fitted lines that larger trials take longer. The fit is not perfectly linear, as there are of course idiosyncracies to particular trials, but especially for the larger "US" trials, which tend to be run by professional units within large firms or third party research organizations that do this as their core business, the fit is reasonably tight, implying an average arrival rate of 186 patients per month.

Robustness and Alternative Explanations: Supplemental Figures and Discussion

B1. 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 across products. Figure A2 provides two types of evidence of this variation. First, the figures in the panels provide patterns for a few individual products, illustrating how learning does not always bring good news, and lack of learning brings a volatile mix of good and bad over time. Second, the table below the panels provides summary statistics on the raw changes in usage patterns with age $\ln(s_{jt}/s_{0t}) - \ln(s_{jt+1}/s_{0t+1})$ for products in the EU, undergoing US trials.

The patterns documented previously regarding decreases in volatility and increasing mean usage with age might be worrisome if they were driven by increasing



(a) Small "EU" Trials **(b)** Large "US" Trials

Figure A1. : Relationship between trial size and time.

usage for all product with age that then asymptotes as in a diffusion process. The table on the raw usage changes show this is not the case—there is a large fraction of changes that are "bad news" for products.

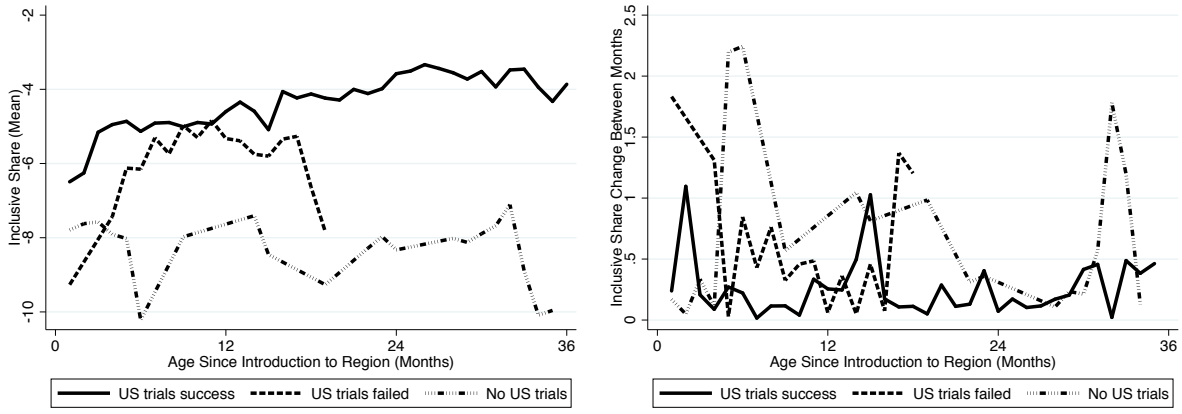
Case Study: CoStar and the Role of Bad News

Here we focus on a clear example of the impact of bad news. A small firm named Conor Medsystems developed a drug-eluting stent with an intuitively appealing new design for drug release that performed well in small early trials (CoStar I (87 patients) and EuroStar I (149 patients)), which were received enthusiastically at conferences in late 2005 through 2006. During this period, pivotal US trials were begun. The stent saw growing market share after receiving a CE mark and being released in the EU in February 2006.[57] In November 2006, Johnson & Johnson was sufficiently optimistic about CoStar to buy Conor for $1.4B. J&J took over CoStar's pre-market notification submission to the FDA. In May 2007 J&J announced the results of a large US trial (CoStar II (1675 patients)), where safety evidence was good but efficacy was disappointing with TLR rates 8% for CoStar versus 4% for its competitor and the control stent, Taxus. Shortly after, J&J announced that it was terminating its FDA mandated clinical trials as the stent was failing to meet its primary endpoints.[58]

The CoStar story demonstrates many of the themes of our analysis. CoStar's usage rose as early trial results were communicated at physician conferences and it underwent US trials. As more information was generated via the clinical trial,

[57] See http://www.ptca.org/pr_conor/20060217.html

[58] See http://www.investor.jnj.com/releasedetail.cfm?releaseid=241182.



(a) $\ln(s_{jt}/s_{0t})$ **(b)** $|\ln(s_{jt}/s_{0t}) - \ln(s_{jt+1}/s_{0t+1})|$

	mean	s.d.	p25	p50	p75	N
$\Delta_t \ln(s_{jt}/s_{0t})\|_{a=1}$	0.24	1.14	-0.16	0.12	0.60	27
$\Delta_t \ln(s_{jt}/s_{0t})\|_{a=12}$	0.17	0.50	-0.04	0.08	0.27	29
$\Delta_t \ln(s_{jt}/s_{0t})\|_{a=24}$	-0.11	0.30	-0.31	-0.06	0.11	32

Figure A2. : 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. Table below summarizes data on raw changes over time for products.

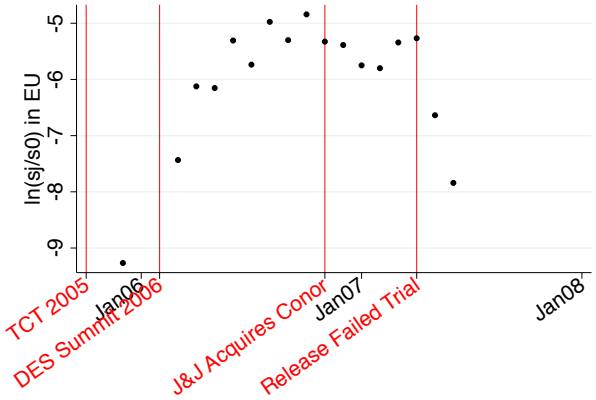


Figure A3. : Evolution of $\ln(s_{jt}/s_{0t})$ for CoStar.

that information is reflected in the inclusive share. Presumably J&J shared this optimism and did not possess differential information, even after due diligence that would have made it privy to the same information as Conor. And when trial results on efficacy were unfavorable, market share dipped and the product was pulled from the market.

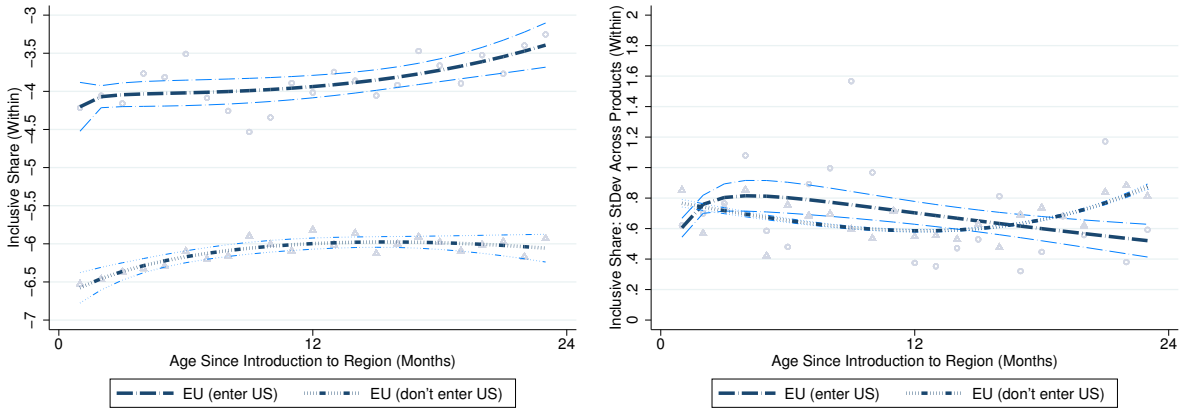
B2. Robustness and Mechanism Tests: Supplemental Figures and Discussion

Placebo Test: PTCA Balloons

One alternative explanation for the findings in Figure 2 would be that the set of manufacturers/products that undergo US trials promote their products differently than other products in the EU, and also differently than for the same products upon US introduction. While we believe the evidence on decreasing variance and on the same products upon US launch make this unlikely, it is not impossible. To further explore this possibility, we perform a placebo test using percutaneous transcatheter coronary artery (PTCA) balloons, which are FDA Class II devices and thus face similar regulatory requirements in both the EU and US. Thus PTCA should not display the differential signs of learning we document for stents if our proposed mechanism is true. The results here show that we do see more total entry in the EU (presumably due to pre-existing complementary sales and distribution assets in the US for some manufacturers); but the differences in amount of entry are smaller than in stents, there is no gap in time of entry on average, and usage patterns with age show no evidence of learning.

As another check that our results are indeed capturing learning in the EU from US clinical trials, we perform a "placebo" type analysis by looking at a device where we know such trials are not required. We perform the analysis on PTCA balloons catheters, which are often used to clear a blockage in the artery before the stent is placed. Standard balloons (ones that do not have drug coatings or special cutting capabilities) typically have little, if any, gap between US and EU approval requirements. This is evident in the lag between US and EU introduction of on average two months (here we calculated entry from first observation in the data instead of looking up press releases, and so the confidence interval includes zero when sampling error is taken into account). Despite this lack of lag for those products introduced in both the US and EU, we still observe many balloons introduced only in the EU because they are sold by the same sales force as stents, but are much lower revenue products, so that only a few companies enter the US market for the purpose of selling balloons only. During our ten year sample, 40 manufacturers sell 113 different balloons in the EU and 6 manufacturers sell 40 different balloons in the US. Thus we can execute our same research design on balloons, with the expectation of no differential learning between products that are EU only versus those that enter the US as well.

Figure A4 shows the results of this placebo test, comparing EU data for products that do and do not enter the US as well. The results illustrate the importance



(a) $\text{Mean}_{j|a}\ \ln(s_{jt}/s_{0t})$ (within product) **(b)** $\text{SD}_{j|a}\ \ln(s_{jt}/s_{0t})$

	$x_{a=1}$	$x_{a=24}$	$x_{24}-x_1$	$(x_{24}^{trials}-x_1^{trials})-(x_{24}^{not}-x_1^{not})$
$\text{Mean}_{j\|a}^{EU\|trials}\ \ln(s_{jt}/s_{0t})$	-4.22	-3.25	0.96	0.37
	(0.17)	(0.27)	(0.32)	(0.33)
$\text{Mean}_{j\|a}^{EU\|not}\ \ln(s_{jt}/s_{0t})$	-6.52	-5.93	0.60	
	(0.16)	(0.18)	(0.21)	
$\text{SD}_{j\|a}^{EU\|trials}\ \ln(s_{jt}/s_{0t})$	0.62	0.59	-0.03	0.01
$\text{SD}_{j\|a}^{EU\|not}\ \ln(s_{jt}/s_{0t})$	0.85	0.81	-0.04	

$N = 789$ product-month observations (all in EU). Standard errors clustered by month $N_t = 114$ in parentheses.

Figure A4. : PTCA Balloons—EU only, products that enter US vs. not.

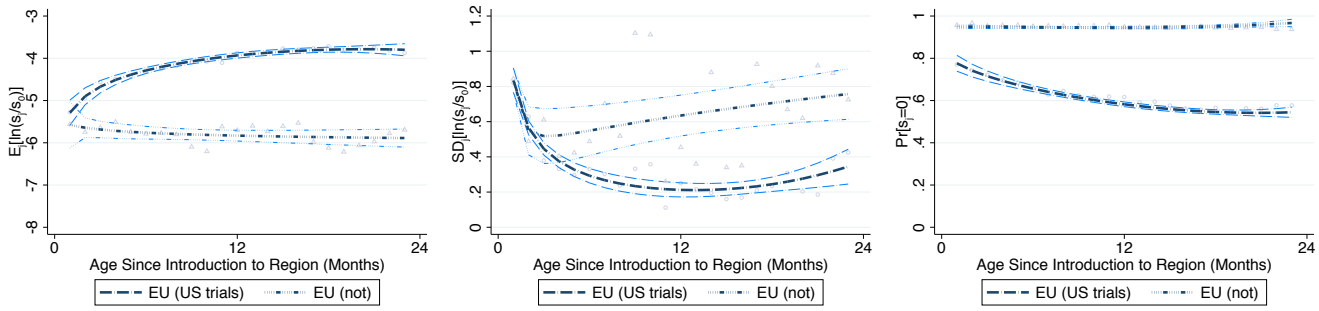
of looking at learning evidence in the volatility along with trends in means as well as the importance of having comparison groups to be able to look at differences-in-differences. There is no evidence of learning in the volatility figure. Mean usage of products in both groups trend up slightly with age, but these trends are statistically identical, suggesting a slight diffusion process that affects all balloons in the EU that is not driven by learning about product quality.

Alternative Explanation: Observational Learning with Different Initial Sample Size

Another potential explanation for the results in Figure 2 is that there is learning in the EU sample undergoing US trials, but this learning is observational (all or in part). The difference between the patterns in the two samples is then plausibly driven by the fact that those stents undergoing US trials enter with higher usage levels, which generate sufficient sample sizes for observational learning to occur, whereas the EU sample not undergoing trials contains too many products that

do not gain enough early traction to enable learning.

We examine this hypothesis by reformulating the same figures and tests for a set of products with overlapping support on initial values of $\frac{1}{J_a}\sum_j \ln(s_{ja}/s_{0a})$ at $a_j = 1$, so they all have similar chances to generate early observational learning. The pattern in Appendix Figure A5 is essentially identical to that in Figure 2, suggesting that our results are not driven by selection on initial quality/usage levels.[59]



(a) $E_{j|a}\left[\ln(s_{jt}/s_{0t})\right]$ **(b)** $SD_{j|a}\left[\ln(s_{jt}/s_{0t})\right]$ **(c)** $Pr_{j|a}\left[s_{jht}=0\right]$

	$\theta_{a=1}$	$\theta_{a=24}$	$\Delta\theta_a$	$\Delta\theta_a^{EU\|_{UStrials}} - \Delta\theta_a^{row}$
$E_{j\|a}^{EU\|_{UStrials}}\ \ln(s_{jt}/s_{0t})$	-5.27	-3.87	1.40***	
	(0.26)	(0.15)	(0.30)	
$E_{j\|a}^{EU\|_{not}}\ \ln(s_{jt}/s_{0t})$	-5.56	-5.70	-0.13	1.54***
	(0.33)	(0.31)	(0.46)	(0.56)
$SD_{j\|a}^{EU\|_{UStrials}}\ \ln(s_{jt}/s_{0t})$	0.82	0.42	-0.40***	
	(0.18)	(0.05)	(0.19)	
$SD_{j\|a}^{EU\|_{not}}\ \ln(s_{jt}/s_{0t})$	0.84	0.72	-0.12	-0.28
	(0.08)	(0.12)	(0.15)	(0.25)
$Pr_{j\|a}^{EU\|_{UStrials}}(s_{jht}=0)$	0.78	0.54	-0.24***	
	(0.04)	(0.02)	(0.05)	
$Pr_{j\|a}^{EU\|_{not}}(s_{jht}=0)$	0.95	0.96	0.01**	-0.25***
	(0.01)	(0.02)	(0.02)	(0.05)

$N^{EU|_{UStrials}} = 197$ (8 products), and $N^{EU|_{not}} = 159$ (10 products) product-month observations.
Standard errors clustered by month $N_t = 114$ in parentheses. $\Delta\theta_a := \theta_{a=24} - \theta_{a=1}$.

Figure A5. : Stent usage patterns after product entry, by region and trial status (subsample matched on $age = 1$ usage)

[59]For this matched sample, selection into US trials must be based on level shifts in expected US profit due to the fact that those products that enter the US all have pre-existing complementary assets for sales and distribution (while those that don't enter do not). This is consistent with the challenges firms such as Biotronik have faced in develop US sales forces. See, "Tipping the Odds for a Maker of Heart Implants," *New York Times*, April 2, 2011.

Alternative Explanation: Asymmetric Information and Signaling

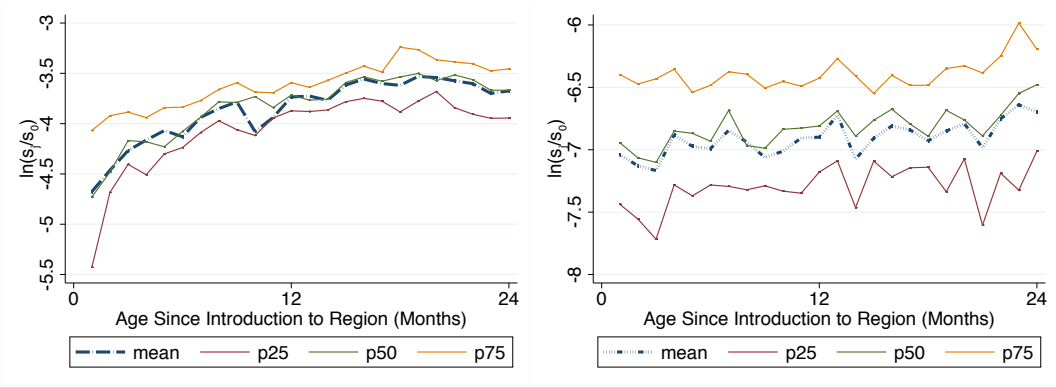
Another potential explanation that could rationalize Figure 2 is manufacturer signaling. Under this hypothesis, after the release of EU trial data, manufacturers retain a sufficiently large degree of private information about expected product quality, and so undertaking costly US trials signals expected product quality to physicians. To produce the observed data patterns, such a signaling model also needs to include some combination of slow signal diffusion across hospitals and/or increasing signal strength as a trial continues. We explore this hypothesis by looking more closely at the shapes of the distribution of $\ln(s_{jt}/s_{0t})$ with age.

Appendix Figure A6 shows the evolution with age of different quantiles of the $\ln(s_{jt}/s_{0t})|_a$ distribution. Under a model where manufacturers and physicians are similarly informed about quality after the release of trials for EU entry, and then learn similarly as data from US trials is released, the distribution of product quality estimates should converge symmetrically to the true product quality distribution. In an asymmetric information setting, consumers do not receive direct information about quality, but instead infer quality must be above some threshold if a manufacturer is willing to continue with costly testing (see Appendix Figure A7 below for more on this intuition). Learning in this way would cause the lower tail of the distribution for product in US trials to become truncated. In the Figure, the 25th and 75th percentiles appear to move symmetrically towards the median as information arrives. Below the figure, we present relevant test statistics. The change in the skewness of the distribution and the change in the ratio of the 75th-50th percentile to the 50th-25th are both insignificant.

Our test of information symmetry in Figure A6 relies upon the intuition that symmetric learning (as we assume in our model) suggests that the inferred distribution of product qualities should tighten from both ends of the distribution as learning occurs (and also shift up if consumers are risk averse). This contrasts with a model where suppliers have private information about their product qualities, where consumer learning should take the form of realizing that manufacturers who engage in costly testing must have product quality exceeding some threshold, which suggests that the inferred distribution of product qualities should tighten from the bottom as learning occurs. Figure A7 illustrates these ideas graphically.

The left panel (a) plots two distributions directly from our EU data for stents undergoing US trials: (Pre-learning) plots the density of $\ln(s_{jt}/s_{0t})|_{a=1}$; and (Post-learning) plots the density of $\ln(s_{jt}/s_{0t})|_{a=12}$. As one would expect from Figure A6, the distribution shifts up and tightens symmetrically after 12 months in US clinical trials.

The right panel (b) plots the same pre-learning distribution, and displays the expected post-learning distribution from applying a truncated learning rule $\ln(s_{jt}/s_{0t})|_{a=1,\ln(s_{jt}/s_{0t})>-6}$. The plot illustrates the type of distribution we might expect if there were learning with asymmetric information. This is clearly different from the symmetric model and from our data, which is why our test in Figure A6 fails to reject the null hypothesis of symmetric learning.

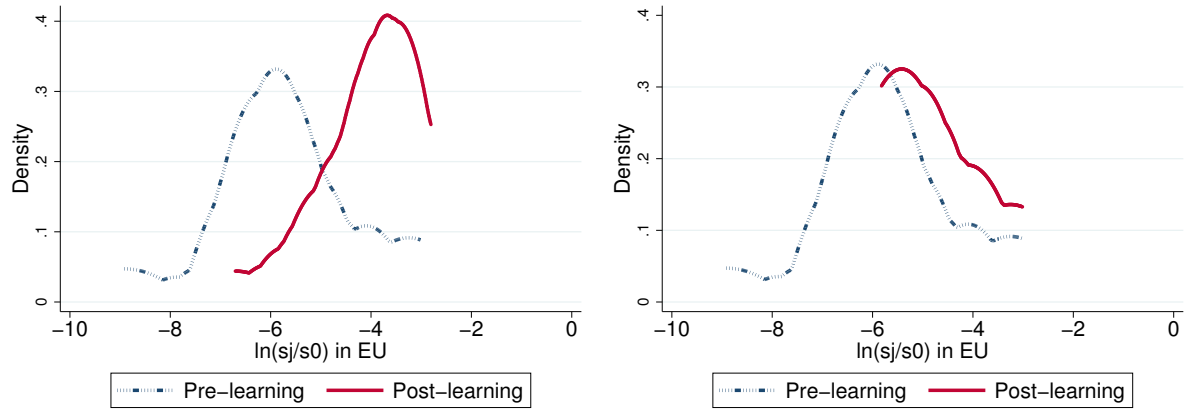


(a) EU, in trials **(b)** EU, not

	EU, in trials			EU, not		
	$\theta_{a=1}$	$\theta_{a=24}$	$\Delta\theta_a$	$\theta_{a=1}$	$\theta_{a=24}$	$\Delta\theta_a$
$\left(\frac{\mu-p50}{\sigma}\right)_{j\|a}$ $\ln(s_{jt}/s_{0t})$	0.06	-0.08	-0.14	-0.09	-0.11	-0.01
$\left(\frac{p75-p50}{p75-p25}\right)_{j\|a}$ $\ln(s_{jt}/s_{0t})$	0.49	0.41	-0.08	0.53	0.42	-0.10

$N = 383$ product-months (in EU; US trials). Standard errors clustered by month $N_t = 114$ in parentheses.

Figure A6. : Symmetry of changes in quality distributions

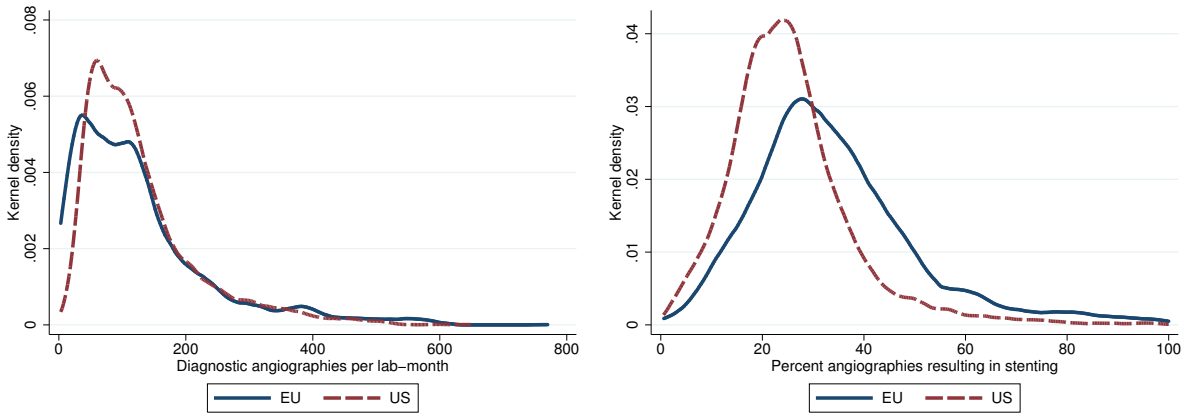


(a) Symmetric Info (Our Estimates) **(b)** Asymmetric Info (Hypothetical)

Figure A7. : Learning effects on inferred product quality distributions under symmetric and asymmetric information mechanisms.

B3. EU vs. US: Other Differences Driving Entry and Diffusion Patterns?

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.[60] This modest differential seems unlikely to account for the stark differences of entry rates between the two regions.



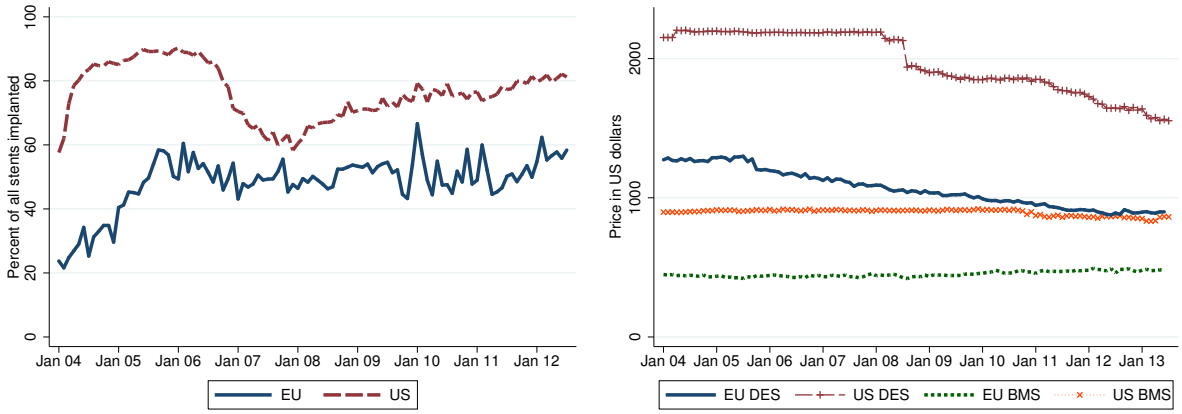
(a) Diagnostic Angiographies **(b)** Percent Resulting in Stenting

Figure A8. : 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.

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

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

EU performing a larger number of angiographies or having a higher rate of stenting conditional on the angiography rate. Figure A8 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.



(a) DES Share of All Stents **(b)** Average Price Paid

Figure A9. : Comparison of usage and price patterns EU vs. US.

Figure A9 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 A9 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 are driven by differences in regulation and not underlying demand.

Theory Appendix

This appendix provides formulas and proofs to supplement the results provided in the body of the paper.

C1. Nested Logit Demand Formulas

Choice probabilities are given by:

$$\text{(C1)}\quad cp_{jht} = Pr[U_{ijht} > U_{ikht}, \forall k \in \mathcal{J}_t] = \frac{\exp\left(\frac{\delta_{jht}}{1-\lambda^{g_j}}\right)}{\sum_{k\in\mathcal{J}_t^{g_j}} \exp\left(\frac{\delta_{jht}}{1-\lambda^{g_j}}\right)} \frac{\left(\sum_{k\in\mathcal{J}_t^{g_j}} \exp\left(\frac{\delta_{jht}}{1-\lambda^{g_j}}\right)\right)^{1-\lambda^{g_j}}}{1+\sum_{\mathcal{J}_t^{g_j}\subset\mathcal{J}_t}\left(\sum_{k\in\mathcal{J}_t^{g_j}} \exp\left(\frac{\delta_{jht}}{1-\lambda^{g_j}}\right)\right)^{1-\lambda^{g_j}}}$$

where $\delta_{jht} := Q_{jht} - \frac{\rho}{2}\sigma_{jt}^2 - \theta^p p_{jht} + \xi_{jh}$ is the mean ex-ante expected utility across patients given beliefs regarding the mean stent performance characteristics and the variance of those beliefs. The corresponding elasticity of choice probabilities with respect to own price is given by

$$\text{(C2)}\qquad \eta_{jht} := \frac{\partial q_{jht}}{\partial p_{jht}}\frac{p_{jht}}{q_{jht}} = -\theta^p\left(\frac{1-\lambda^{g_j} cp_{jht|g} - (1-\lambda^{g_j})cp_{jht}}{1-\lambda^{g_j}}\right) p_{jht}.$$

The ex-ante expected consumer surplus (relative to the outside option) as a function of information and choice set is

$$\text{(C3)}\quad CS_{ht}(\mathcal{J}_t, \mathcal{I}_{ht}) = \theta^{scale} \ln\left(1 + \sum_{\mathcal{J}_t^{g_j}\subset\mathcal{J}_t}\left(\sum_{j\in\mathcal{J}_t^{g_j}} \exp\left(\frac{\delta_{jht}}{1-\lambda^{g_j}}\right)\right)^{1-\lambda^{g_j}}\right).$$

where θ^{scale} is set to make fully informed the average treatment on the treated vs. non-stent alternatives for DES introduced to the US equal to \$5000, as motivated by the clinical literature discussed in the body of the paper.

$$\text{(C4)}\qquad 5000 = \theta^{scale}\frac{1}{|\mathcal{J}_{DES,US}|}\sum_{j\in\mathcal{J}_{DES,US}} \frac{\ln\left(1+\exp\left(\delta_{jht}\right)\right)}{\left(\frac{1+\exp\left(\delta_{jht}\right)}{\exp\left(\delta_{jht}\right)}\right)}$$

C2. Regulator's Total Surplus Tradeoff: Illustrative Case

The general total surplus function is complicated by the entry policies of firms, tracking observational learning for firms that entered the market at different times, and potential distortions due to heterogeneity in marginal costs and price markups. To clearly see the core tradeoff between uncertainty and access in the

model, it is helpful to consider a simple case of a simple logit demand model with testing and entry costless, no observational learning, homogenous marginal costs (normalized to zero for convenience), and no distortions in usage due to price. In this case, the regulator's tradeoff simplifies as follows:

$$\begin{aligned} TS_t(T^c+1) - TS_t(T^c) &= \ln\left(\frac{\sum_{j\in\mathcal{J}_t(T^c+1)} e^{Q_{jt}-\frac{\rho}{2}\sigma^2_{jt}(T^c+1)}}{\sum_{j\in\mathcal{J}_t(T^c)} e^{Q_{jt}-\frac{\rho}{2}\sigma^2_{jt}(T^c)}}\right) - \chi\,|\mathcal{J}^e_t(T^c+1)\setminus\mathcal{J}^e_t(T^c)| \\ \text{(C5)} &= \ln\left(\frac{\sum_{j\in\mathcal{J}_t(T^c+1)} e^{Q_{jt}-\frac{\rho}{2}\sigma^2_{jt}(T^c+1)}}{\sum_{j\in\mathcal{J}_t(T^c)} e^{Q_{jt}-\frac{\rho}{2}\sigma^2_{jt}(T^c)}}\right) \\ \text{(C6)} &= \underbrace{\frac{\rho}{2}\left(\sigma^2(T^c)-\sigma^2(T^c+1)\right)}_{\text{gain from decreased risk}} - \underbrace{\ln\left(\frac{\sum_{j\in\mathcal{J}_{t+1}(T^c+1)} e^{Q_{jt}}}{\sum_{j\in\mathcal{J}_t(T^c+1)} e^{Q_{jt}}}\right)}_{\text{gain from tech change/entry}} \end{aligned}$$

where (C5) follows from no fixed costs, and (C6) follows from no observational learning and recognizing $\chi = 0 \Rightarrow \mathcal{J}_t(T^c) = \mathcal{J}_{t+1}(T^c+1)$.

C3. Learning and Risk Aversion Predictions for Shares

The data exploration section explores several patterns of distribution of market shares, conditional on age. This Section provides further justification for the relationship between these patterns and learning / risk aversion.

Prediction 1 (Learning): If initial product quality is uncertain $\sigma_Q > 0$, then learning $1/\sigma^2_A > 0$ implies that expected volatility in product-specific quality estimates (demeaned by true product quality) converge by decreasing with age to zero $Var_{j|a}(\delta^a_j - \bar{\delta}_j) \searrow^{a\to\infty} 0$.

Proof: It is clear from the model setup and Bayes' rule that nonzero precision of the learning process $1/\sigma^2_A > 0$ (in and/or out of trials, so here we suppress that subscript) implies convergence of quality estimates to the true quality $Q^a_j \longrightarrow^{a\to\infty} Q^*_j$ and the convergence of uncertainty about that estimate to zero $\sigma^a_j \searrow^{a\to\infty} 0$ for any product j. Our further claim is that evidence for this learning will be found by looking at measures of volatility

of the mean utilities across products:

$$\begin{aligned}
\lim_{a\to\infty} Var_{j|a}(\delta_j^a - \bar{\delta}_j) &= \lim_{a\to\infty} Var_{j|a}(Q_j^a - Q_j^*) - 0 \\
&= \lim_{a\to\infty} Var_{j|a}\left(\frac{\sigma_{EU}^2}{a\sigma_{EU}^2+\sigma_A^2}\sum_{\tau=1}^{a}\left(Q_j^*+\epsilon_j^\tau\right) + \frac{\sigma_A^2}{a\sigma_{EU}^2+\sigma_A^2}Q_j^1 - Q_j^*\right) \\
&= \lim_{a\to\infty} E_{j|a}\left[\frac{\sigma_{EU}^2}{a\sigma_{EU}^2+\sigma_A^2}\sum_{\tau=1}^{a}\epsilon_j^\tau + \frac{\sigma_A^2}{a\sigma_{EU}^2+\sigma_A^2}(Q_j^1 - Q_j^*)\right]^2 \\
&= \lim_{a\to\infty}\left(\frac{\sigma_{EU}^2}{a\sigma_{EU}^2+\sigma_A^2}\right)^2 a\sigma_A^2 + \left(\frac{\sigma_A^2}{a\sigma_{EU}^2+\sigma_A^2}\right)^2 {\sigma_j^2}^1 \\
&\searrow 0 \qquad \text{(C7)}
\end{aligned}$$

where the first equality follows from the mean utility specification used in the body of the paper (and relies on linear separability of the quality term). The second substitutes for Q_j^a using Bayes' rule. The third from distributing the Q_j^* term and the fact that the term inside brackets has expectation zero. The fourth from distributing the square and taking expectations. And monotonically decreasing convergence of that quantity to zero is clear. *Q.E.D.*

Prediction 2 (Risk Aversion): If consumers (doctors making decisions on behalf of their patients) are risk averse ($\rho > 0$), then expected product usage, conditional on age, will increase strictly as learning occurs ($E_j[\delta_j^a] \nearrow^{a\to\infty} \overline{Q}$).

Proof: This again follows from the basic structure of the learning model. Consider the quantity of interest:

$$\begin{aligned}
\lim_{a\to\infty} E_j[\delta_j^a] &= \lim_{a\to\infty} E_j[Q_j^a] - \frac{\rho}{2}{\sigma_j^2}^a \\
&= E_j[Q_j^*] - 0 \\
&= \bar{Q} \qquad \text{(C8)}
\end{aligned}$$

where the first line follows from the mean utility specification. The second line follows from the convergence of $\{Q_j^a\}$ and $\{\sigma_j^{2a}\}$. And ${\sigma_j^2}^a \searrow^{a\to\infty} 0$ and $\rho > 0$ guarantee the convergence is increasing and strictly so. *Q.E.D.*

Demand/Learning Estimation: Supplementary Details

D1. 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) Construct an initial estimator for σ_Q using the empirical equivalent from the Q_j^* from the estimator using age by trial status fixed effects instead of the learning model.

2) Guess initial values for learning precisions $(\sigma_{EU}, \sigma_A, \sigma_{A^c})$ and hospital heterogeneity $(\sigma_H^{des}, \sigma_H^{bms}, \gamma_H)$.

3) Compute the full vector of σ_{jt}^2 implied by σ_Q^2, the learning precision parameters, and which products are in trials when.

4) Least squares then gives an estimator for the linear parameters $(\rho, Q_j^*, \theta^p, \lambda^g)$.

5) Repeat 2-4 until minimize the GMM objective function.

6) Recompute σ_Q using the empirical equivalent from the Q_j^* from this stage.

7) Repeat 2-6 until σ_Q converges.

D2. Robustness and Alternative Structural Demand Models

Table A2 displays results for several robustness checks on our demand/learning model specification. The first two columns use age fixed effects, interacted with a dummy variable indicating whether the product is in US clinical trials, to provide a less parametric way to capture how demand changes over time with age and trial status. The first column (NL) estimates a simple nested logit model, shutting down any variation in preferences across hospitals. The second column (NLQW) estimates the Quan and Williams (2018) model. The results show how across hospital heterogeneity is important for fitting the data as the criterion function reaches a lower minimum with this added flexibility. As expected, this acts as a selection correction for the product-hospital-months with zero shares, which shifts the product fixed effect estimates.

The third column (NLQWNN) adds the structure of the Normal-Normal learning model in place of the age and trial status fixed effects. There are two primary differences: (1) the learning model parameterization forces learning to be smooth over time (vs. the nonparametric fixed effects); and (2) the learning model uses the rational expectations assumption to link the product fixed effect estimates Q_j^* to how demand evolves with age and trial status. Under rational expectations, the fixed effect estimates must be consistent with the prior distributions $F^{UStrials}(Q)$ and $F^{not}(Q)$, and the precision parameters in the learning model ($1/\sigma_{EU}^2$, $1/\sigma_{A^c}^2$, $1/\sigma_A^2$) link the prior to how the variance and levels moments of product usage evolve with age and trial status.

Figure A10 plots the age fixed effects in NLQW and uncertainty discounts $-\frac{\rho}{2}\sigma_{jt}^2$ in NLQWNN versus age. The left panel shows the products in US trials; the right

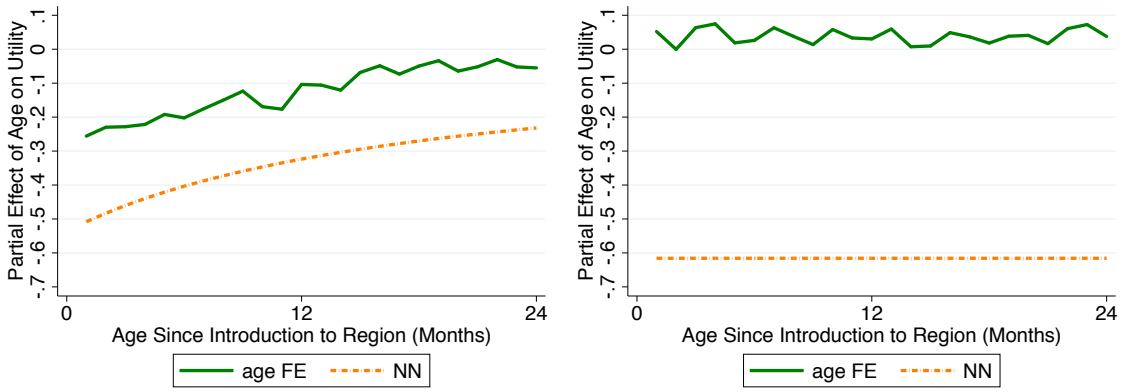
Table A2—: Estimates of physician preference and learning model parameters

	NL	NLQW	NLQWNN	$\sigma_A(q_{jt-1})$	H Lags
θ^p (utils/\$1000)	0.21	0.20	0.10	0.10	0.11
	(0.03)	(0.04)	(0.04)	(0.04)	(0.05)
λ^{des}	0.88	0.84	0.81	0.81	0.81
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
λ^{bms}	0.91	0.88	0.82	0.82	0.81
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
σ_H^{des}		0.14	0.19	0.19	0.19
		(0.03)	(0.04)	(0.04)	(0.04)
σ_H^{bms}		0.07	0.18	0.18	0.18
		(0.01)	(0.02)	(0.02)	(0.03)
$\rho \cdot \theta^p$ (1/\$1000)			3.36	3.29	3.66
			(1.70)	(2.30)	(1.99)
$1/\sigma_{EU}^2$			18.82	18.75	18.52
			(2.16)	(2.09)	(2.74)
$1/\sigma_{A^c}^2$			1.73	1.70	1.88
			(0.51)	(0.60)	(0.70)
$1/\sigma_A^2$			0.00	0.00	0.00
			(0.00)	(0.15)	(0.13)
γ_H					
β^q ($\frac{1}{\sigma_A^2}$/100)				0.00	
				(0.00)	
μ^{lag} (months)					0.00
					(0.10)
$\overline{Q}_j$	-2.06	-2.58	-2.37	-2.37	-2.41
	(0.05)	(0.06)	(0.10)	(0.35)	(0.31)
$\sigma_Q\vert_{UStrials}$	0.27	0.30	0.26	0.26	0.27
	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)
$\sigma_Q\vert_{not}$	0.30	0.36	0.34	0.34	0.34
	(0.02)	(0.03)	(0.01)	(0.01)	(0.02)
$age \times UStrials$ FE	Y	Y	N	N	N
$\min(GMM_{criterion})$	101.47	15.47	15.53	15.54	15.54
$RMSE(\xi_{jt})$	0.413	0.401	0.281	0.279	0.297

Estimates for demand model $\ln(s_{jt}/s_{0t}) = \lambda^{g_j}\ln(s_{j|_g t}) - \theta^p p_{jht} + Q_j^* - \frac{\rho}{2}\sigma_{jt}^2 + \xi_{jt}$ with separate nests for DES and BMS. $N_{JHT} = 407,191$ product-hospital-months and $N_{JT} = 4,888$ product-months. Standard errors in parentheses and are clustered by month ($N_T = 114$).

panel products not in US trials. The patterns show that: (1) With regards to the smooth parameterization of the learning model, the fit is still quite close to the pattern of the age fixed effects, so the parametric form imposes very little on the data (this can be seen in the figures and also in the $\min(GMM_{criterion})$ (fitting the aggregate usage, aggregate volatility, and hospital moments) and $RMSE(\xi_{jt})$ (of aggregate usage moments only) in the table being close for the two models). (2) The rational expectations assumption allows the model to extract much more information from the data – the prior is now linked to the fixed effects, and so we

can infer the amount of learning from EU trials/approval from the gap between that and the variation in usage patterns at $age = 1$. Pinning uncertainty to these two points then allows us to infer the amount of uncertainty remaining as learning does or doesn't occur, and it then requires the product quality estimates Q_j^* to adjust for this. Finally, the NN learning model separates learning/uncertainty and risk aversion parameters – these structural parameters have a clear interpretation that allows for validation of the results, and they allow estimation of counterfactuals where the nature of uncertainty/learning might change due to regulatory changes.



(a) In US Trials **(b)** Not in US Trials

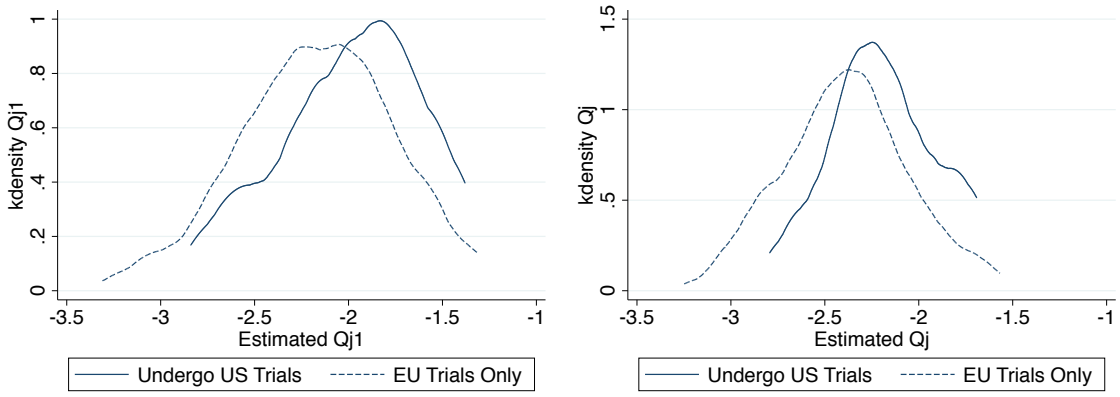
Figure A10. : Comparison of estimates from fixed effect and learning models. Plots the estimated discount due to uncertainty versus product age in Normal-Normal learning model $-\frac{\rho}{2}\sigma_{jt}^2$ vs. model with age and $age \times UStrials$ fixed effects.

Returning to Table A2, the final two columns show the results for two extensions of our model. Column 4 allows observational learning in the market place to be a function of past demand, $\sigma_A(q_{jt-1})$, in order to check that the lack of market learning that we estimate is not being driven by low usage levels. This is similar in spirit to our "overlapping $Q_{j,age=1}$" test in the reduced form section, and like that test, we do not find any evidence of observational learning for products not in US trials, even for those that are used in large quantities.

The final column, $HLags$, reports the estimates for a model that allows different hospitals to learn with different random lags from each information shock. The goal is to allow for the patterns in the data to potentially be generated by an information diffusion process that is not intrinsically linked to information generation (in which case we might be conflating this diffusion process with the trial information generation process, which is the process the regulator controls most directly). Specifically, each hospital receives each signal with a delay of a

number of months drawn from a Poisson(μ^{lag}) distribution, and we estimated μ^{lag} using simulated GMM. Similar to the hospital signal correlation parameter in our preferred model, this lag parameter is identified by the difference in the aggregate vs. hospital level patterns of usage. The estimate does not suggest that learning lag heterogeneity across hospitals explains the patterns in the data.

Estimated product quality distributions



(a) $f(Q_j^1|UStrials)$ **(b)** $f(Q_j^*|UStrials)$

Figure A11. : Estimated Quality Distributions. Density plots of the product fixed effect parameters $\{Q_j^*\}$ and market expectations upon EU entry $\{Q_j^1\}$.

One advantage of the GMM algorithm vs. ML (besides the ability to use instruments, which is of course important) is that it allows a nonparametric distribution of product quality estimates. Figure A11 plots the distribution of Q_j^1 (left panel (a)) and Q_j^* (right panel (b)). The results help to validate several of the maintained assumptions. The distributions are not perfectly normal, but appear to be symmetric and reasonably approximated by normals (especially since the tails are inherently difficult to estimate). Also, it does seem that the $UStrials$ distribution may indeed be best thought of as a different distribution with a slightly higher mean and smaller variance. But the distribution does not appear to have a different shape in a way that would make the two groups difficult to compare or that would suggest an asymmetric information signaling equilibrium.

Counterfactuals: Supplementary Details

E1. Partial equilibrium effect of risk: dependence on quality relative to outside option

Table A3 replicates Table 2 in the paper body, allowing the mean qualities to vary by shifting the entire quality distribution by plus or minus a standard deviation of the logit horizontal error term. As referred to in the paper, the effects of decreasing risk become more dramatic as mean quality increases relative to the outside option.

Table A3—: The effect of uncertainty on number of stenting procedures, surplus per stent implanted, and expected ex post loss.

		$\sigma_Q =$ 0.312	$\sigma_{a_{EU}=0} =$ 0.185	$\sigma_{T^c=6} =$ 0.160	$\sigma_{T^c=12} =$ 0.143	$\sigma_{T^c=18} =$ 0.131	$\sigma_{T^c=24} =$ 0.121	$\sigma_{T^c=30} =$ 0.113
Baseline $\overline{Q_j^*}$	$1-s_0$ (%)	12.5	24.0	26.4	27.9	29.0	29.7	30.3
		(2.5)	(1.4)	(1.3)	(1.3)	(1.3)	(1.4)	(1.4)
	$\frac{TS}{1-s_0}$ ($)	5776	6103	6184	6238	6276	6304	6327
		(176)	(167)	(167)	(168)	(169)	(170)	(171)
	$E[Q_j^* - Q_{jt}\|j^*]$ ($)	-1096	-560	-429	-348	-292	-252	-221
		(127)	(23)	(37)	(41)	(41)	(39)	(37)
$\overline{Q_j^*} + \sqrt{\pi/6}$	$1-s_0$ (%)	33.8	52.9	56.1	58.0	59.2	60.1	60.8
		(5.0)	(1.8)	(1.6)	(1.6)	(1.6)	(1.6)	(1.6)
	$\frac{TS}{1-s_0}$ ($)	6525	7458	7663	7795	7887	7955	8007
		(301)	(233)	(230)	(232)	(234)	(236)	(238)
	$E[Q_j^* - Q_{jt}\|j^*]$ ($)	-1083	-554	-425	-344	-289	-249	-219
		(127)	(23)	(36)	(40)	(40)	(39)	(37)
$\overline{Q_j^*} - \sqrt{\pi/6}$	$1-s_0$ (%)	3.9	8.1	9.1	9.8	10.2	10.6	10.9
		(0.9)	(0.6)	(0.6)	(0.6)	(0.6)	(0.6)	(0.6)
	$\frac{TS}{1-s_0}$ ($)	5533	5611	5634	5651	5663	5672	5679
		(137)	(138)	(138)	(139)	(139)	(139)	(139)
	$E[Q_j^* - Q_{jt}\|j^*]$ ($)	-1102	-563	-432	-350	-294	-254	-223
		(127)	(23)	(37)	(41)	(41)	(40)	(38)

E2. Algorithms for computing equilibrium counterfactuals

The Matlab code for the counterfactuals is available in the electronic archive *code4RegulatingInnovation.zip*. This appendix outlines the main steps of the algorithms.

For each potential $T^c = 0, 1, ..., 30$ (and given estimated or comparative static parameter values) we calculate the M and L cases as follows:

M case:

1) Simulate $r = 1, ..., 20$ draws from the posterior signal distribution for each observation, given $\sigma_{jt}(T^c)$ and qualities at EU entry Q_j^1.

2) Restrict sample to products that would be active in each month $\mathcal{J}^M(T^c) = \mathcal{J}_{t+T^c}$.

3) Given $\mathcal{J}^M(T^c)$, use demand/learning models to compute equilibrium quantities and surplus measures for each r.

4) Estimate expected surplus measures using mean of r realizations.

L case:

1) Given M case results and $\chi T^c = 1.6E6 * T^c$, restrict sample to products that would enter, under the naive assumption that firms assume other products enter as if $\chi = 0$ (i.e. single agent entry, assuming competing against $\mathcal{J}^M(T^c) \setminus j$), such that

$$E\left[\sum_{t=\underline{t}_j+T^c}^{\bar{t}_j+T^c}\left(\sum_h q_{jht}(p_{jht}-mc_j)\right)^{1-.01(t-\underline{t}_j-T^c)} |\mathcal{I}^1, \chi_{-j}=0\right] > \chi_j T^c$$

, with the expectation computed as the mean across the r simulations of the profits for product j competing against $\mathcal{J}^M(T^c) \setminus j$ (i.e. exactly the expected product profits from the M case). The set of products that would enter is $\mathcal{J}^L(T^c)$.

2) Given $\mathcal{J}^L(T^c)$, use demand/learning models to compute equilibrium quantities and surplus measures for each r.

3) Estimate expected surplus measures using mean of r realizations.

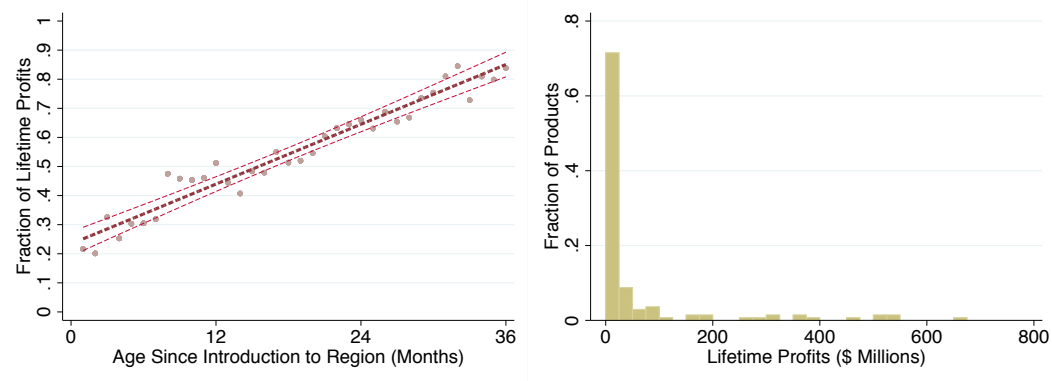
In the Appendix where we also model pricing, equilibrium prices are also computed simultaneously with quantities. Standard errors for counterfactuals are calculated using a nonparametric delete-10 jackknife, blocked at the month level.

E3. Distribution of Profits Over Product Lifetime and Across Products

The counterfactual L case 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.

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.

In our counterfactuals, we hold take time from entry until exit as exogenously given, and we use this same extrapolation to calculate expected lifetime profits



(a) Cumulative Profits Over the Product Lifecycle

(b) Distribution of Lifetime Profits over Products

	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 ($1000s)	1,347	2,119	41	262	4,067	68
Distribution of lifetime profits (extrapolated where necessary):						
Lifetime Profit ($M)	72.6	141.5	1.3	10.5	304.3	109

Figure A12. : Distribution of Profits Over Age and Across Products.

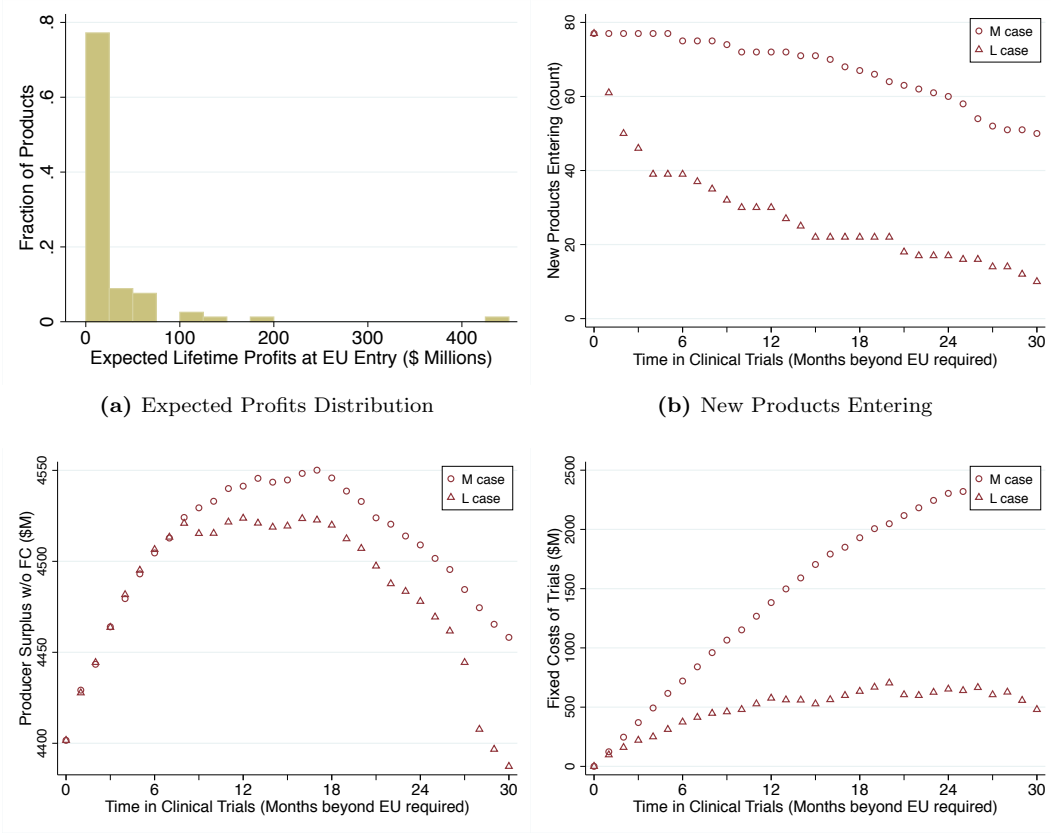
for any product that has not exited by the end (or entered by the beginning) of the sample period in the counterfactual. Our counterfactual estimates are robust to a variety of approaches to this extrapolation. This is in part because the extrapolation is typically for the beginning or end of lifetime tail of product profits, so that lifetime profit projections are not very sensitive to the method we choose. Further, the products that are marginal in the sense that they exit as entry costs increase, are also marginal in their contribution to consumer surplus.

Also notable here is the distribution of estimated lifetime profits in the last row of the table at the bottom of Figure A12, which makes it clear that many products with quite low profitability enter the EU. This supports our assumption that the products in the EU market represent a reasonable approximation to the set of products developed that firms might consider testing and bringing to market.

E4. Additional Results for $E[\pi_j(T^c)]$,$\mathcal{J}(T^c)$, $PS(T^c)$, $FC(T^c)$

Each of our counterfactuals fully characterizes outcomes (under our M and L cases), but we only show results for CS and TS in the body of the paper. Figure A13 shows: (a) the distribution of expected profits across products given their

information sets upon EU entry and beliefs about other product entry, (b) the number of new products entering, (c) variable producer surplus, and (d) fixed cost expenditures on trials.



(a) Expected Profits Distribution

(b) New Products Entering

(c) Firm Variable Profits

(d) Fixed Costs of Trials

Figure A13. : Additional Results for $E[\pi_j(T^c)]$,$\mathcal{J}(T^c)$, $PS(T^c)$, $FC(T^c)$ Computed for our baseline scenario of a change in policy in January 2004 (no spillovers).

Several highlights from these results are discussed in the body of the paper. The full results here provide a more complete picture of how delay of entry from longer trials (subfigure (b), M case) affects the choice set of products available for our period of study. Additionally, they show how fixed costs of additional trials (subfigure (d)) plus the distribution across products of expected profits (subfigure (a)) can further combine to limit the set of products available (subfigure (b), L case). The bulk of this decreased entry due to larger fixed costs of testing comes from the 76 percent of products whose expected lifetime profits are below $25M.

E5. Robustness of Counterfactual Estimates to Modeling Assumptions

The paper makes a number of modeling assumptions to facilitate the counterfactual computations. Here we explore robustness of the results to the extent to which the government/regulator discounts future surplus, and to the assumptions on marginal costs and prices. Table A4 shows the results.

The assumed amount of discounting has rather large effects on the amount of surplus generated by optimal testing. The reason for this is that it takes time for many products used in the market to be products that have benefited from further testing, whereas the costs of lack of access to new products are more uniformly distributed in time. Thus more heavily discounting the future reduces the benefit from further testing. However, this effect is primarily borne out in the level of surplus gain from optimal testing, with the amount of increased testing under optimal policy changing little with the discounting applied.

Table A4—: Robustness: Discounting, Marginal Costs, and Pricing. Top row repeats our baseline results (discounted at one percent nominal per month, $mc_j = .50 * \min(p|g_j)$, and prices held fixed). Subsequent rows modify these assumptions as indicated in the first column.

State of Market at Policy Change	$\Delta CS(T^{c*})$ (%)		T^{c*}_{CS} (months)		$\Delta TS(T^{c*})$ (%)		T^{c*}_{TS} (months)	
	(L)	(M)	(L)	(M)	(L)	(M)	(L)	(M)
Jan 2004	5.2	6.7	16	17	3.9	6.3	16	17
	(1.8)	(1.9)	(4)	(2)	(1.6)	(1.9)	(4)	(2)
Jan 2004, Govt. Discount 7% real	5.8	7.3	18	18	4.5	6.9	16	17
	(2.1)	(2.2)	(4)	(3)	(2.0)	(2.1)	(4)	(3)
Jan 2004, Govt. Discount 3% real	6.6	8.2	19	19	5.2	7.7	16	19
	(2.1)	(2.2)	(4)	(3)	(2.0)	(2.1)	(4)	(3)
Jan 2004, No Govt. Discount	7.7	9.5	19	19	6.2	8.9	16	19
	(2.1)	(2.2)	(4)	(3)	(2.0)	(2.1)	(4)	(3)
Jan 2004, $mc_j = .25 * \min(p\|g_j)$	5.4	6.7	13	17	3.9	6.3	13	17
	(1.8)	(1.9)	(3)	(2)	(1.6)	(1.9)	(4)	(2)
Jan 2004, $mc_j = .75 * \min(p\|g_j)$	5.2	6.7	16	17	3.8	6.3	13	17
	(1.8)	(1.9)	(4)	(2)	(1.6)	(1.9)	(4)	(2)
Jan 2004, Pricing Modeled	5.2	6.6	13	17	4.0	6.7	13	18
	(1.7)	(1.9)	(3)	(2)	(1.6)	(1.8)	(4)	(3)

$N_{JHT} = 407,191$ product-hospital-months and $N_{JT} = 4,888$ product-months. Standard errors, clustered by month ($N_T = 114$) using a delete-10 block jacknife, in parentheses.

Turning to marginal costs and prices, our primary specification holds prices fixed in the counterfactual and assumes marginal costs at half the minimum price observed within each product group ($g \in \{bms, des\}$) for the purpose of computing firm expected and realized profits. The bottom rows of Table A4 show that results change very little, if at all, when we change marginal costs to be either one quarter or three quarters the group minimum, or when we instead model pric-

ing and use that model to estimate marginal costs and compute new equilibrium prices in the counterfactuals.

Pricing: model, identification, estimation, and results

Here we describe the pricing model and estimation for the robustness check where we model pricing explicitly.

In the EU, device pricing practices vary somewhat across countries and hospitals, but are typically negotiated between manufacturers and either the hospital or some regional body responsible for procurement for a set of hospitals. For this Appendix, we model this process using a static Nash Equilibrium of Nash Bargaining models for each period, following the theory developed in Horn and Wolinsky (1988) and Collard-Wexler et al. (2019) and recent empirical work by Crawford and Yurukoglu (2012), Grennan (2013), and Gowrisankaran et al. (2014). These approaches assume that prices maximize the bilateral Nash product

(E1)
$$\max_{p_{j\mathcal{H}t}} \left(\sum_{h \in \mathcal{H}} \pi_{j\mathcal{H}}(\mathcal{J}_t, I_{ht}, p_{j\mathcal{H}t}) \right)^{b_{jt}(\mathcal{H})} \left(\sum_{h \in \mathcal{H}} CS_{ht}(\mathcal{J}_t, I_{ht}, p_{j\mathcal{H}t}) - CS_{ht}(\mathcal{J}_t \setminus \{j\}, I_{ht}, p_{j\mathcal{H}t}) \right)^{b_{\mathcal{H}t}(j)},$$

for each $j \in \mathcal{J}_t$ in each market (group of hospitals in bargaining unit $\mathcal{H}$ in each month t), taking other prices in the market $\{p_{k\mathcal{H}t}\}_{k \in \mathcal{J}_t}$ as given. Here $\pi_{j\mathcal{H}} := q_{jht}(p_{j\mathcal{H}t} - mc_j)$ are manufacturer variable profits at marginal cost, mc_j, for each device. CS_{ht} is the hospital level consumer surplus. The parameters $b_{jt}(\mathcal{H})$ and $b_{\mathcal{H}t}(j)$ are the Nash bargaining weights, determining the extent to which equilibrium prices weight manufacturer profit (minus its outside option of not producing the stent) versus hospital surplus (minus its outside option of optimal choice for each patient from a choice set that excludes j).[61]

With the demand and learning model parameter estimates in hand, we turn to estimating the parameters from the bargaining model between hospitals and device manufacturers. To take the bargaining model to the data, we follow Grennan (2013) and rewrite (E1) as

$$\text{(E2)} \quad p_{j\mathcal{H}t} = mc_j + \frac{b_{jt}(\mathcal{H})}{b_{jt}(\mathcal{H}) + b_{\mathcal{H}t}(j)} \sum_{h \in \mathcal{H}} \frac{M_{ht}}{\sum_{h \in \mathcal{H}} M_{ht}} AV_{jht}(\mathcal{J}_t, I_{ht}, p_{j\mathcal{H}t}, mc_j; \theta^D) \ ,$$

where the added value of product j to hospital h is defined as

(E3)
$$AV_{jht} := \left(1 + \frac{\partial q_{jht}}{\partial p_{jht}} \frac{p_{j\mathcal{H}t} - mc_j}{q_{jht}}\right) \frac{CS_{ht}(\mathcal{J}_t, I_{ht}, p_{j\mathcal{H}t}) - CS_{ht}(\mathcal{J}_t \setminus \{j\}, I_{ht}, p_{j\mathcal{H}t})}{q_{jht}} + p_{j\mathcal{H}t} - mc_j$$

[61] Assuming constant returns to scale in distribution and manufacturing on the margin at $\sum_{h \in \mathcal{H}} q_{jht}$. We also follow previous work in maintaining the Nash-like assumption that other prices remain the same in the case of disagreement, which is consistent with "passive beliefs" in the theory literature that provides noncooperative foundations for this concept.

To estimate the structural parameters, the bargaining weights and the marginal costs, we use our utility model estimates. We calculate the substitution patterns by simulating η_{jht}, the elasticities across hospitals as defined in Equation (C2), over the distribution of hospital level unobservables $f_H(\xi_{jh}; \sigma_H^{g_j})$ (suppressing dependence on hospital-specific learning for simplicity since $1/\sigma_A^2$ and γ_H are both estimated to be zero). Similarly, we use the consumer surplus equation derived from the utility function in Equation (C3) to compute the buyer surplus portion of the added value.

Because of physicians may be imperfect agents for patients and/or hospitals, estimated physician price sensitivity measures may not reflect the correct scaling for measuring hospital and/or consumer surplus. For this reason, we deviate slightly from the standard demand estimation approach in that we add a scaling coefficient to relate consumer surplus derived from consumer utility maximization to estimates of the dollar value of quality adjusted life years obtained in clinical studies. We normalize the total surplus per stenting procedure to \$5,000, which is the approximate median of the estimated dollars in quality adjusted life years from the procedure relative to a coronary artery bypass graph surgery, a more invasive alternative to receiving angioplasty and a stent.[62] This alternative scaling is only for translating welfare measures into dollars—we continue to use the estimated θ^p in quantity and elasticity calculations, as revealed preference indicates this is the level of price sensitivity that best fits the demand patterns in the data.

In addition to the standard set of issues that the bargaining literature has identified in estimating marginal costs and bargaining parameters, we face two additional challenges. First, the challenge in estimating demand at the hospital level means that our demand estimates only provide the distribution of added values across hospitals, not the hospital-specific added values. Second, because we only observe a sample of hospitals, we do not have added value measures for all hospitals in a group $\mathcal{H}$ in cases where our hospitals may negotiate as part of a group.

We address both of these supply estimation challenges by aggregating our estimation strategy across hospitals to the product-month level. Otherwise, we follow Grennan (2013, 2014): We assume the econometric error enters relative bargaining weights multiplicatively:

$$mc_j := \mu_g^C \quad ; \quad \frac{b_{jt}(\mathcal{H})}{b_{\mathcal{H}t}(j)} := \frac{\beta_j}{\beta_{\mathcal{H}}} \nu_{j\mathcal{H}t} \tag{E4}$$

where μ_g^C allows marginal cost to vary across BMS/DES, and β_j are product-specific bargaining parameters to be estimated. Substituting into Equation (E2), rearranging and taking logs to obtain a linear equation in the unobservable, and

[62] Among studies reported in the Cost Effectiveness Analysis Registry (https://research.tufts-nemc.org). We could also scale into dollars using the standard approach of the inverse of the price coefficient $\frac{1}{\theta^p} = 10,482$, which would approximately double all related consumer welfare estimates.

aggregating over hospitals gives the equations at the product-month level that we take to the data:

$$\text{(E5)} \qquad \sum_h \frac{M_{ht}}{M_t} \ln\left(\frac{p_{jht}-\gamma_j}{AV_{jht}}\right) = \ln(\beta_j) + \underbrace{\sum_h \frac{M_{ht}}{M_t}(-\ln(\beta_h)+\ln(\nu_{jht}))}_{\tilde{\nu}_{jt}}$$

The parameters (β_j, γ_j) can then be estimated by a GMM algorithm, assuming $E_{jt}\left[\tilde{\nu}_{jt}|Z_{jt}\right] = 0$ for a set of instruments including product-specific constants and $\frac{\partial \tilde{\nu}_{jt}}{\partial \gamma_j}$.

Table A5 presents the parameter estimates from the bargaining model. As we estimate the parameters at the product level, we present the means and standard deviations of those estimates. The first two variables are the elasticity and average value parameters that come from the demand model that feed into the bargaining model. The price elasticity is somewhat higher from DES stents and the typic BMS stent adds $1155 of value while the average DES stent adds significantly more value at $1424.

Table A5—: Structural parameter estimates for pricing model

	η^p_{jht}		AV_{jht} ($)		mc_j ($)		$\frac{b_{jt}(\mathcal{H})}{b_{jt}(\mathcal{H})+b_{\mathcal{H}t}(j)}$	
	mean	sd	mean	sd	mean	sd	mean	sd
BMS	-0.25	0.11	1155	118	87	-	0.41	0.12
	(0.06)	(0.06)	(172)	(41)	(124)		(0.06)	(0.04)
DES	-0.42	0.14	1424	224	361	-	0.60	0.14
	(0.32)	(0.11)	(312)	(60)	(117)		(0.08)	(0.04)

$N_{JHT} = 407,191$ product-hospital-months and $N_{JT} = 4,888$ product-months. Standard errors clustered by month ($N_T = 114$) using a delete-10 block jacknife in parentheses.

The next two sets of variables are parameter estimates from the bargaining model. The marginal cost estimates align with expectations and prior literature. BMS cost an average of $87 to produce while DES are more costly at $361. Finally, the last two columns present the estimates of the relative bargaining weights, $\frac{b_{jt}(\mathcal{H})}{b_{jt}(\mathcal{H})+b_{\mathcal{H}t}(j)}$. The results imply that for BMS, hospitals retain the majority of the surplus (manufacturers obtain 41 percent on average) from the implantation of the device, with a modest amount of variance across products. However, for the newer DES technology, on average, the manufacturers receive the majority of the surplus (60 percent).[63]

[63]The DES bargaining parameter is nearly double that in Grennan (2013) in the US 2004-07 subsample, but this corresponds closely to the magnitude of our alternative scaling of consumer surplus, and may also be related to lower reimbursements to hospitals and the different competitive environment for DES in the EU relative to the US.

E6. Additional Comparative Static Counterfactual Results

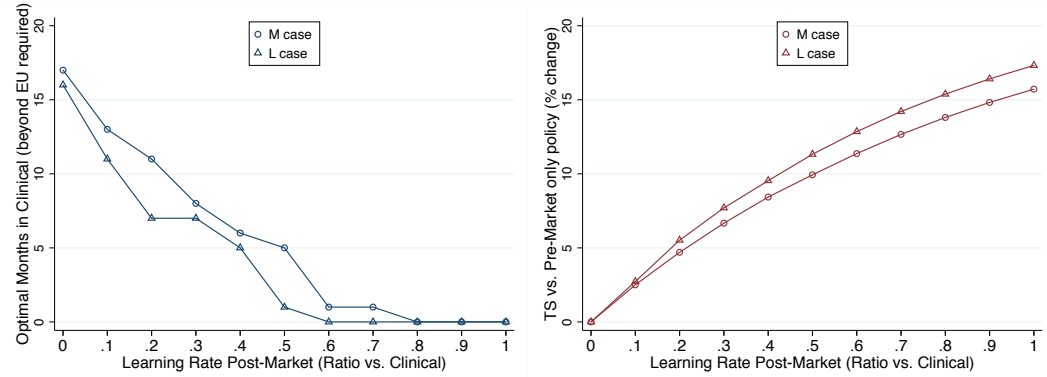
Table A6 provides results for a more complete set of comparative static parameter values than those shown in the body of the paper in Table 3. The key takeaways were discussed in the body of the paper. We supply the full set of numerical values here for the interested reader. We find these useful in considering the robustness of our estimated results for coronary stents 2004-13, and also for thinking about how these results might extrapolate to other product categories with different primitives. Perhaps the most non-obvious and important qualitative result is the nonlinearity of optimal testing *length* with respect to the precision of information generated by testing. As mentioned in the body of the paper, this thought experiment is tightly linked to pushes for validating more intermediate/surrogate endpoints (e.g. for stents this would be measuring 6 month "loss" in the target vessel diameter instead of revascularization or mortality endpoints) that could allow trials to generate more information with smaller sample sizes.

Table A6—: Sensitivity of Optimal Regulation to Key Parameters: Full results for all parameter values we have explored.

State of Market at Policy Change	$\Delta CS(T^{c*})$ (%)		T^{c*}_{CS} (months)		$\Delta TS(T^{c*})$ (%)		T^{c*}_{TS} (months)	
	(L)	(M)	(L)	(M)	(L)	(M)	(L)	(M)
Jan 2004	5.2	6.7	16	17	3.9	6.3	16	17
	(1.8)	(1.9)	(4)	(2)	(1.6)	(1.9)	(4)	(2)
Jan 2004, FC * .1	6.6	6.7	17	17	5.9	6.3	17	17
	(1.9)	(1.9)	(3)	(2)	(1.8)	(1.9)	(3)	(2)
Jan 2004, FC * .2	6.5	6.7	17	17	5.7	6.3	17	17
	(1.9)	(1.9)	(3)	(2)	(1.8)	(1.9)	(3)	(2)
Jan 2004, FC * .5	6.2	6.7	14	17	5.0	6.3	13	17
	(1.9)	(1.9)	(3)	(2)	(1.8)	(1.9)	(3)	(2)
Jan 2004, FC * 2	3.9	6.7	9	17	2.4	6.3	7	17
	(1.6)	(1.9)	(4)	(2)	(1.3)	(1.9)	(6)	(2)
Jan 2004, FC * 5	1.7	6.7	7	17	0.0	6.3	0	17
	(1.1)	(1.9)	(3)	(2)	(0.4)	(1.9)	(2)	(2)
Jan 2004, FC * 10	0.0	6.7	0	17	0.0	6.3	0	17
	(0.3)	(1.9)	(1)	(2)	(0.0)	(1.9)	(0)	(2)
Jan 2004, σ_Q * .5	0.1	0.2	1	5	0.0	0.1	0	1
	(0.3)	(0.5)	(2)	(3)	(0.1)	(0.5)	(1)	(4)
Jan 2004, σ_Q * .75	2.2	3.1	8	13	1.3	2.9	6	13
	(1.3)	(1.5)	(4)	(3)	(1.0)	(1.4)	(4)	(3)
Jan 2004, σ_Q * .9	4.0	5.3	11	17	2.7	5.0	11	17
	(1.6)	(1.8)	(4)	(3)	(1.4)	(1.7)	(4)	(3)
Jan 2004, σ_Q * 1.1	6.4	7.9	18	18	5.0	7.5	16	17
	(1.9)	(2.1)	(3)	(3)	(1.8)	(2.0)	(4)	(2)
Jan 2004, σ_Q * 1.33	8.6	10.2	19	19	6.9	9.6	17	19
	(2.1)	(2.2)	(3)	(3)	(2.0)	(2.1)	(4)	(3)
Jan 2004, σ_Q * 2	12.1	13.7	19	19	10.2	13.1	18	19
	(2.3)	(2.4)	(3)	(3)	(2.2)	(2.3)	(4)	(3)
Jan 2004, ρ * .5	0.8	1.2	6	6	0.2	1.1	3	6
	(0.5)	(0.7)	(2)	(2)	(0.3)	(0.6)	(2)	(2)
Jan 2004, ρ * .75	2.7	3.6	9	13	1.8	3.3	6	13
	(1.2)	(1.3)	(3)	(3)	(1.0)	(1.3)	(3)	(3)
Jan 2004, ρ * 1.33	9.6	11.2	19	19	7.7	10.6	18	19
	(2.6)	(2.7)	(3)	(3)	(2.4)	(2.6)	(5)	(3)
Jan 2004, ρ * 2	16.3	20.0	20	26	14.1	19.1	20	26
	(4.1)	(4.3)	(3)	(2)	(3.9)	(4.1)	(3)	(3)
Jan 2004, $1/\sigma^2_{A^c}$ * .2	0.0	0.1	0	1	0.0	0.1	0	1
	(0.1)	(0.5)	(2)	(4)	(0.0)	(0.4)	(0)	(4)
Jan 2004, $1/\sigma^2_{A^c}$ * .5	1.5	2.6	7	13	0.7	2.4	6	13
	(1.2)	(1.5)	(4)	(4)	(0.9)	(1.4)	(5)	(4)
Jan 2004, $1/\sigma^2_{A^c}$ * .75	3.5	4.9	16	17	2.2	4.6	8	17
	(1.6)	(1.8)	(4)	(3)	(1.4)	(1.7)	(5)	(3)
Jan 2004, $1/\sigma^2_{A^c}$ * 1.33	6.9	8.4	13	17	5.5	8.0	16	17
	(1.9)	(2.1)	(3)	(2)	(1.8)	(2.0)	(3)	(2)
Jan 2004, $1/\sigma^2_{A^c}$ * 2	9.6	10.7	13	17	7.9	10.2	13	17
	(2.1)	(2.2)	(3)	(3)	(1.9)	(2.1)	(3)	(3)
Jan 2004, $1/\sigma^2_{A^c}$ * 5	14.2	15.3	10	12	12.6	14.5	9	11
	(2.3)	(2.3)	(2)	(2)	(2.1)	(2.2)	(2)	(2)

$N_{JHT} = 407,191$ product-hospital-months and $N_{JT} = 4,888$ product-months. Standard errors, clustered by month ($N_T = 114$) using a delete-10 block jacknife, in parentheses.

E7. Post-Market Surveillance and Consumer Surplus



(a) Optimal Trial Length, $T^{c*}_{CS}(\sigma_A)$

(b) Consumer Surplus, $CS(\sigma_A, T^{c*}_{CS}(\sigma_A))$

Figure A14. : The Value of Post-Market Surveillance (Consumer Surplus): Plots of optimal trial length (left panel (a)) and total surplus (right panel (b)) as observational learning precision $1/\sigma^2_A$ varies from zero to the clinical trial precision $1/\sigma^2_{Ac}$.

As noted in the body of the paper, the CS metric generates tighter bounds and greater returns to optimal pre-market policy. The CS metric is of special interest in the post-market surveillance case because it is derived from only the risk-access tradeoff, not the fixed costs savings from less trials. As a result, optimal pre-market trial length decreases less quickly with post-market learning under the CS criterion.

*

REFERENCES

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

Aguiar, L. and Waldfogel, J. (2018). Quality predictability and the welfare benefits from new products: Evidence from the digitization of recorded music. *Journal of Political Economy*, 126(2):492–524.

Auerbach, D., Maeda, J., and Steiner, C. (2012). Hospital stays with cardiac stents, 2009 (hcup statistical brief 128). Technical report, Agency for Healthcare Research and Quality.

Berry, S. and Waldfogel, J. (1999). Free entry and social inefficiency in radio broadcasting. *RAND Journal of Economics*, 30(3):397–420.

Brown, Z. Y. (2017). An empirical model of price transparency and markups in health care. Working Paper, University of Michigan.

Budish, E., Roin, B., and Williams, H. (2015). Do firms underinvest in long-term research? evidence from cancer clinical trials. *American Economic Review*, 105(7):2044–2085.

Cardell, N. S. (1997). Variance components structures for the extreme-value and logistic distributions with applications to models of heterogeneity. *Econometric Theory*, 13:185–213.

Cardon, J. H. and Hendel, I. (2001). Asymmetric information in health insurance: Evidence from the national medical expenditure survey. *The RAND Journal of Economics*, 32(3):408–427.

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

Ching, A. T., Erdem, T., and Keane, M. P. (2013). Learning models: An assessment of progress, challenges, and new developments. *Marketing Science*, 32(6):913–938.

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

Collard-Wexler, A., Gowrisankaran, G., and Lee, R. S. (2019). Bargaining in bilateral oligopoly: An alternating offers representation of the" nash-in-nash" solution. *Journal of Political Economy*, 127(1):163–195.

Crawford, G. and Yurukoglu, A. (2012). The welfare effects of bundling in multichannel television. *American Economic Review*, 102(2).

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

Donahoe, G. and King, G. (2014). Estimates of medical device spending in the united states. Technical report, Advanced Medical Technology Association.

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.

Gandhi, A., Lu, Z., and Shi, X. (2013). Estimating demand for differentiated products with error in market shares.

Gowrisankaran, G., Nevo, A., and Town, R. J. (2014). Mergers when prices are negotiated: Evidence from the hospital industry. *The American Economic Review*, 105(1):172–203.

Grabowski, H. and Wang, Y. R. (2008). Do faster food and drug adminstration drug reviews adversely affect patient safety? an analysis of the 1992 prescription drug user fee act. *Journal of Law and Economics*, 51(2):377–406.

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. *Management Science*, 60(12):3011–3025.

Grennan, M. and Swanson, A. (2018). Diagnosing price dispersion: Demand, bargaining, and search in hospital-supplier contracting. Working paper.

Grennan, M. and Town, R. J. (2018). Regulating innovation with uncertain quality: Information, risk, and access in medical devices. NBER working paper 20981.

Hamilton, B. H., Jungheim, E., McManus, B., and Pantano, J. (2018). Healthcare access, costs, and treatment dynamics: Evidence from in vitro fertilization. *American Economic Review*, 108(12):3725–77.

Handel, B. R. (2013). Adverse selection and inertia in health insurance markets: When nudging hurts. *American Economic Review*, 103(7):2643–82.

Harrington, S. E. (2012). Cost of capital for pharmaceutical, biotechnology, and medical device firms. In *The Oxford Handbook of the Economics of the Biopharmaceutical Industry*. Oxford.

Ho, K. and Lee, R. S. (2019). Equilibrium provider networks: bargaining and exclusion in health care markets. *American Economic Review*. forthcoming.

Horn, H. and Wolinsky, A. (1988). Bilateral monopolies and incentives for merger. *RAND Journal of Economics*, 19:408–419.

Ifrach, B. and Weintraub, G. (2017). A framework for dynamic oligopoly in concentrated industries. *The Review of Economic Studies*, 84(3):1106–1150.

Kalyanasundaram, A. and Karlheinz, P. (2014). Comparison of revascularization procedures in coronary artery disease. *Medscape*. http://emedicine.medscape.com/article/164682.

Kamenica, E. and Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, 101(6):2590–2615.

Kaplan, A. V. and Stern, A. D. (2018). The central and unacknowledged role of the us food and drug administration in the design and execution of medical device pivotal trials. *JAMA Cardiology*, 3(1):5–6.

Kyle, M. (2007). Pharmaceutical price controls and entry strategies. *Review of Economics and Statistics*, 89(1):88–99.

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*.

Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., Abraham, J., Adair, T., Aggarwal, R., Ahn, S. Y., et al. (2012). Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the global burden of disease study 2010. *The Lancet*, 380(9859):2095–2128.

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.

MedTech Europe, S. (2013). The european medical technology industry - in figures 2013. Technical report, MedTech Europe.

Miravete, E. J., Seim, K., and Thurk, J. (2018). Market power and the laffer curve. *Econometrica*, 86:1651–87.

Miravete, E. J., Seim, K., and Thurk, J. (2019). One markup to rule them all: Taxation by liquor pricing regulation. forthcoming.

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

Nistor, C. and Tucker, C. E. (2015). Certification intermediaries: Evidence from the medical device industry. working paper.

Olson, M. K. (2008). The risk we bear: The effects of review speed and industry user fees on new drug safety. *Journal of Health Economics*, 27:175–200.

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

Petrin, A. (2002). Quantifying the benefits of new products: The case of the minivan. *Journal of Political Economy*, 110(4):705–729.

Philipson, T., Berndt, E. R., Gottschalk, A. H., and Sun, E. (2008). Cost-benefit analysis of the FDA: The case of the prescription drug user fee acts. *Journal of Public Economics*, 92(5-6):1306–1325.

Quan, T. and Williams, K. (2018). Product variety, across market demand heterogeneity, and the value of online retail. *RAND Journal of Economics*, 49(4):877–913.

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.

Sampat, B. (2018). A survey of empirical evidence on patents and innovation. NBER working paper 25383.

Schreyögg, J., Stargardt, T., Tiemann, O., and Busse, R. (2006). Methods to determine reimbursement rates for diagnosis related groups (DRG): a comparison of nine european countries. *Health Care Management Science*, 9(3):215–223.

Scott Morton, F. and Kyle, M. (2012). Markets for pharmaceutical products. In Pauly, M., McGuire, T., and Barros, P., editors, *Handbook of Health Economics*, volume 2, chapter 12. Elsevier.

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.

Sorenson, C. and Kanavos, P. (2011). Medical technology procurement in europe: a cross-country comparison of current practice and policy. *Health Policy*, 100(1):43–50.

Spence, M. A. (1975). Monopoly, quality, and regulation. *Bell Journal of Economics*, 6(2):417–429.

Stern, A. D. (2017). Innovation under regulatory uncertainty: evidence from medical technology. *Journal of Public Economics*, 145:181–200.

Stern, A. D., Kramer, D. B., Ouellet, M., and Kesselheim, A. S. (2017). Review times and adverse events for cardiovascular devices. *Nature*, 1(0013):0013.

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

Train, K. (2015). Welfare calculations in discrete choice models when anticipated and experienced attributes differ: A guide with examples. *Journal of Choice Modelling*, 16:15–22.

Van Norman, G. A. (2016). Drugs and devices: comparison of european and us approval processes. *JACC: Basic to Translational Science*, 1(5):399–412.