Proliferation and entry deterrence †

Farasat A.S. Bokhari [a] and Weijie Yan [b,c,*]

[a] *School of Economics and Centre for Competition Policy, University of East Anglia, UK*

[b] *Economic and Social Research Institute, Ireland*

[c] *Department of Economics, Trinity College Dublin, Ireland*

Spring 2020

Abstract

Do firms increase product lines to deter entry and if so, when is this strategy successful? We use data from UK pharmaceuticals to examine how incumbents respond to change in the threat of entry. Consistent with entry deterrence motive, we find that originators slow their product launch rate when they find out that entry is very likely, particularly in medium-sized markets. We also find that originators can successfully deter entry via product proliferation in medium-size markets, but product hopping is a more successful strategy in larger markets. Our findings have implications for competition policy and goes beyond pharmaceutical industry.

Key words: Product proliferation, product hopping, entry deterrence, pharmaceuticals, hazard models

JEL Classification: L40, L12, L41, I11, L79

†This paper has benefitted from comments from Imad Ahmed, Ashish Arora, Steve Davies, Paul Dobson, Margaret Kyle, and Franco Marriuzo as well as from participants at the Centre for Competition Policy seminar series (2020) and presentations at EARIE (2019), and IEA (2019), IIOC (2022), CRESSE (2022), BECCLE (2022). The usual caveats apply. This paper is based on a chapter of Weijie Yan's PhD thesis.

*Corresponding Author: Weijie Yan (weijie.yan@esri.ie)

1. INTRODUCTION

Do firms invest in capacity to deter entry and, if so, how can we tell it apart from other unilateral profit maximizing incentives for investment? Further, when are these strategies effective, and what form do these investments take? Firms may invest in capacity to meet an expected increase in future demand, or may invest in product variety to appeal to a larger customer base and to grow. But these can also be pre-emptive actions to fill the product space and prevent entry. Although there is extensive theoretical research on this topic, and when such actions might be credible, empirical evidence is limited. This is because entry deterrence via strategic investments is empirically difficult to isolate from other profit maximizing incentives. Also, the impact of such a strategy is hard to measure due to the lack of an observable counterfactual: did the non-entrant not enter because of a successful deterrence strategy, or because entry was never intended? Our article contributes to the growing empirical literature on strategic deterrence and its effectiveness using the timing of product line extensions in pharmaceutical markets in the UK. It also explains when this strategy works, and at least in the pharmaceutical context, the form of investment. Our results are of interest to industrial organization economists, and to antitrust authorities.

Some of the earlier theoretical literature has argued that deterrence is not rational or on the equilibrium path, and that investments can be delayed till after entry to drive out competition. Later models, such as those building on commitment mechanisms showed that pre-emption can be a subgame perfect outcome (Spence, 1977, Dixit, 1979, 1980). In line with this literature, our results show that originators indeed launch additional variants of their drugs pre-emptively under the threat of entry, and change their launch strategy once entry becomes very likely. Further, this result is mostly present in medium-sized markets where motive for deterrence is strongest. We find that these pre-emptive launches are successful in deterring entry, but in a non-obvious way. Deterrence is successful in medium-sized markets when the originator covers the horizontally differentiated product space with enough patients for each of its product line extensions, and thus making the potential market for any

single presentation thin enough so a generic firm finds entry difficult. However, as we show later, this strategy does not work in small or large markets. An alternative (and rarer) strategy is to shift most of the patients to only the new variants of the drug. In small markets this opens up an opportunity for entry and we find that shifting patients to new drugs is correlated with competitive entry. In larger markets however, shifting patients to the new drug is likely to be accompanied with significant marketing efforts that convince patients and doctors of the higher quality of the new variant i.e., products are vertically differentiated. In these markets entry can be blocked as it is no longer attractive to enter with a generic of the original variant, while the originators' new variants may be protected by intellectual property. Accordingly, we find that entry in large markets is deterred when patients are mostly shifted to the newer variants.

A central issue in identifying entry-deterring investments is to separate the decisions of incumbents under the threat of entry from entry itself, since in the latter case incumbents may be adjusting to the new market structure (key in Goolsbee and Syverson (2008) and Cookson (2017) who study incumbent behavior in response to exogenous changes in potential entry). To this end, pharmaceutical markets present an ideal opportunity for testing deterrence versus accommodation. In most western markets, originators are protected against generic competition for a fixed period of time due to a combination of patent laws and data/market exclusivity rules that prevent generic firms from filing entry applications. While the date of actual entry by a generic firm is potentially endogenous, the date of the end of marketing exclusivity (henceforth 'loss of exclusivity' or LoE) is a pre-determined fixed period of time starting from when the initial market authorization was issued to the originator. We exploit the variation in originator behaviour around this exogenously set date to detect entry deterrence motives.

An often noted point in the literature is that large markets are more likely to attract generic entry than small markets (Scott Morton, 1999, Reiffen and Ward, 2005). This observation is further developed in Ellison and Ellison (2011) and in Dafny (2005), where they note that originators would deviate

from optimal investment only if such an action can lead to profitable deterrence. In their case, this deviation is observed in medium-sized markets: in small markets entry may already be blocked due to the size of the market, while in large markets it may not be feasible to deter, and hence over/under investment would happen only in the medium-sized markets if firms were investing for deterrence motives. We combine this with insights from Goolsbee and Syverson (2008) and Cookson (2017) about change in the threat of entry to compare product launch rate before and after the loss of exclusivity but *before* any entry takes place. Post the loss of exclusivity, generic applications for entry can be filed, and ceteris paribus the threat of entry would increase. Thus we may see firms increasing their effort to deter entry if entry is not yet assured.

By contrast, if a generic application is actually filed, the originators would find out (in the EU, the European Medicines Agency publishes monthly a list of original drugs for which it is reviewing generic applications). In the latter case, entry is imminent and originators know about it, and *if* the motive for investment was deterrence, i.e., they had deviated from their investment trajectory under a blocked entry, they would stop investing any further and return to the optimal path. Thus, the loss of exclusivity marks a sharp point at which firms' actions would change in the presence of deterrence, but not necessarily if they were accommodating entry. To that end, we explore changes in launch rates around the loss of the exclusivity period — and again before any actual entry – by originator type: those that eventually experience entry or not, and by the size of the originators' market.

We find that in medium-sized markets, the launch strategy of the two types of originators deviates from their earlier rates and from each other in a way that is consistent with entry deterrence. Originators that do not experience any eventual entry increase their product launch rate in response to the exogenous increase in threat of entry, while those that eventually experience entry slow down their rate after the end of exclusivity. For the latter group we assume that when a generic entry application is filed, originators know that an

application has been filed, and hence act accordingly. While this is not exactly 'smoking gun' evidence, as we cannot randomize which firms eventually experience entry and which do not, it does provide strong evidence consistent with entry deterrence incentive: one group that never experiences any entry increases its launch rate after the end of exclusivity in response to the increased threat of entry, while the other group that eventually experiences entry, slows its product launch rate after finding out that entry is imminent. Further, while the originators differ in their observable characteristics in the full sample, when we re-do the analysis by market-size (small, medium, large), they become more like 'matched samples' with very similar characteristics for originators with and without entry in small and medium-sized markets. It turn this may alleviate some concerns about comparing originators with and without entry given that it is not a random assignment.

If strategic deterrence is present, it raises the second issue: when are these strategies effective? There are two related but distinct mechanisms by which entry can be deterred via product launches. First, as discussed in Schmalensee (1978) or Smiley (1988), firms can fill the product space via proliferation. However, as pointed out in Judd (1985), this is not credible because the incumbent can always withdraw a product post entry, and since the competitor knows that, product proliferation would not be a deterrent without a commitment mechanism. In pharmaceuticals, the commitment to not withdraw its products comes via physician detailing efforts (i.e., advertising and sales calls which require significant sunk costs), where a significant portion of the existing patients are moved to the newer formulation or dosage before entry by a competitor. If both the original and the new variants of the drug are covered by the originator such that there are enough patients using both type of drugs, then the product space is covered, and there may not be any room for entry. More relevantly however, since the entry is typically for a generic version of an existing variant and not a new formulation, the horizontally differentiated products by the the incumbent split the patient base across formulations so that any one variant where a generic may try to enter may be thin enough not to be able to recover entry costs.

Second, an alternative mechanism is via 'product hopping', where entry can be blocked by switching almost all the patients to the newer formulation, again via physician detailing, and sometimes in conjunction with withdrawing the initial formulation from the market prior to entry. This is the case of vertical differentiation where the manufacturer's newer version is supposedly better for all patients and not just some of them. This strategy was highlighted in the European Commission's pharmaceutical sector inquiry as well as in several antitrust cases in the US (see EC, 2009, Carrier and Shadowen, 2016). To understand how this works, note that in the EU, data exclusivity granted to the original formulation applies simultaneously to any additional strengths or formulations, and hence any product line extensions launched later do not receive exclusivity extensions by the drug approval agencies. By contrast, in the US, a new formulation can receive three years of exclusivity. Nonetheless, as pointed out in Kyle (2016), new formulations in the EU may still be protected against generic entry via secondary patents on the product line extensions. Thus, if through physician detailing efforts patients can be switched to newer formulations or dosages, and via secondary patents generic entry into new product line extensions can be blocked, then originators can deter entry even in the original formulation due to lack of patients. Such detailing efforts are likely in large markets, but less likely in smaller markets. In the latter case if the originators switch to the newer version, it may be because it is costly to maintain multiple product lines and the newer variant is more profitable. Thus product hopping could lead to entry in smaller markets and to deterrence in larger markets.

Thus we also estimate hazard models at the market level to test if launching additional variants deters entry. Given our earlier finding that originators with eventual entry launch more drugs than those without entry, it is hardly surprising that in any such models the count of products is positively correlated with probability of entry. We confirmed this to be the case in our estimates as well, and that higher (lagged) sales attract more entry. More importantly however, we focus on the role of shifting patients to the newer drugs. We find that if the originator launches additional variants, and *if* the relative market share of the originator's drugs are evenly spread across its various formulations

to fill the product space, then it reduces the risk of entry significantly in medium-size markets. This strategy does not deter entry in large markets. The alternative strategy of product hopping, where most of the patients are shifted to the newer formulations, seems to be effective in deterring entry in large markets, and attracts entry in small markets, though the evidence is somewhat weaker and product hopping is rare.

In summary then, we find that some incumbents slow their product launch rate after the end of exclusivity while others increase it, and that these changes take place before any actual entry. These are firms that eventually experienced entry or not respectively, and are in medium-sized markets. If this change in launch rate is driven purely by demand side fundamentals, then we would expect easier/faster generic entry after the loss of exclusivity as well, but it is precisely in the medium-sized markets that we find that entry is less likely if (i) the incumbent has launched more products and (ii) has covered all variants with significant share of patients. By contrast, product hopping is rare but creates opportunity for entry in small markets, and deters entry in large markets, possibly due to differences in detailing efforts across these two markets.

Related Literature. Our paper is related to two streams of literature, the first in empirical industrial organization that focuses on entry deterrence strategies such as capacity, product proliferation, advertising, fighting brands, licensing, and pricing, and builds on the insights from theoretical models (see Wilson, 1992, for a review of the theory). In the earlier empirical literature, Lieberman (1987) did not find any evidence of investment in capacity to deter entry, but Weiman and Levin (1994) documented in a case study that Southern Bell Telephone dramatically expanded capacity in long-distance and toll lines as pre-emptive investment. Particularly, they note that the company expanded its pole miles from 2,000 to 8,600 and their toll wire coverage from 5,000 to over 55,000 miles, and that the timing of this investment was strategic. More recently, Dafny (2005) used a monotonicity test to document evidence of investing in capacity to deter entry in an invasive cardiac procedure by US hospitals. Goolsbee and Syverson (2008) analyzed incumbent reactions when

Southwest airlines became present at both ends of a route, but before flying the specific route itself. They found strong evidence that incumbents lower their fares when faced with the threat of entry, but the reasons for such pre-emptive actions for deterrence versus accommodation have mixed evidence. Seamans (2012) found that incumbent cable TV providers responded to the threat of entry by municipal entrants (but not private entrants) by upgrading their cable systems, but conditional on upgrade, they were less likely to offer services that ran on those upgrades when compared to incumbents not facing a similar threat. Additionally, they reported that the strategy appeared to have worked: of the 400 cases where an incumbent faced potential entry by a public utility and upgraded their system, only nine actually experienced entry. Finally Cookson (2017) provided evidence from the American casino industry and documents both the investment in capacity to deter entry as well as the effectiveness of these investments that reduced the likelihood of entry.

Our paper is also related to a second more specialized literature in pharmaceuticals that considers similar strategies as above (related to capacity, prices, advertising etc.), but within the context of intellectual property and market exclusivity rules or other country specific regulations. For instance, Caves et al. (1991) analyzed price and advertising expenditures for 30 drugs that experienced patent expiration in the US between 1976 and 1987, and found no evidence of limit pricing. They also noted that while reduction in advertising expenditures typically starts two years before patent expiration, it is because innovators expect lower returns from advertising once the generic entry takes place rather than to deter entry. Reduction in advertising two years prior to loss of patent is also noted in other studies, including most recently in Castanheira et al. (2019). Similarly, Scott Morton (2000) found that pre-patent expiration advertising by the originator did not deter entry by generics. On the other hand, Bergman and Rudholm (2003) found evidence of limit pricing in the Swedish market, where once a branded firm lowered its drug price, it committed to that price because regulations in the country prevented it from raising them easily post generic entry.

In line with the focus of our paper on product-line extensions, Huskamp et al. (2008) maintained that new formulations allowed pharmaceutical firms in the US to extend market exclusivity, and instead investigated how new formulations affected a branded firm's advertising strategy. They found that promotions are shifted away from the original formulation and towards the newer formulation well before generic entry takes place. Ellison and Ellison (2011) developed and used the non-monotonicity test and provided some (weak) evidence from the US showing that additional products were launched more in medium-sized markets, as would be predicted under a deterrence hypothesis. By contrast, Danzon and Furukawa (2011) looked at the effectiveness of launching additional formulations pre-patent expiration and found that it lowered the probability of generic entry in the US, but not in other countries, including in the UK. While we find a similar result for the number of formulations (like in their case, positive but not significant for the UK), we find that the probability of entry is lowered if originators can successfully shift patients to the newer formulation in medium or large markets.

Product-line extensions in pharmaceuticals often rely on secondary patents, which may be weaker and draw challenges from potential generic entrants. In the context of market exclusivity rules in the US, Grabowski and Kyle (2007) reported that generic firms are increasingly engaging in a 'prospecting' approach, i.e., where even a small probability of a win in a patent litigation can draw many generic challenges (particularly for drugs that have large sales) and shorten the effective market exclusivity period for the branded firms. Hemphill and Sampat (2012) also reported a similar increase in patent challenges. However, they also found that weaker and later expiring patents i.e., those associated with product line extensions via firms' 'evergreening' strategy, drew disproportionately more challenges and in fact maintained the historical effective exclusivity period for new molecular entities.

Finally, a closely related literature considers pre-emption by a branded firm into the generic segment either via the launch of an in-house 'pseudo' or branded generic, or via a licensing agreement with a third party to launch an 'authorized generic'. Hollis (2003) and Hollis and Liang (2007) argued that

authorized generics diminish incentives for independent entry, particularly in small markets (in the US, the Hatch-Waxman Act rewards the first successful generic with a 180-day market exclusivity over other generics). Consistent with that view, the Federal Trade Commission estimated that authorized generics can reduce first generics' revenues by 40-52% during the exclusivity period issued to the generic and by 53-62% in the following 30 months (FTC, 2011). Reiffen and Ward (2007) used calibration for the US market to estimate that anticipated entry of a branded generic crowds out 1.7-2.4 independent generic entries regardless of market size. However, they argued that deterring motivation is likely present in small and medium-sized markets, as it also helps maintain higher prices (manage cannibalization) in the branded segment, while in large markets, their motivation is to capture generic profits rather than deterrence itself. Berndt et al. (2007) claimed that despite the increasing rate of authorized generics, the rate of challenges under the provisions of the act is also high, and there is no evidence on the entry deterrence effect of authorized generics. Finally, Appelt (2015) provided more direct evidence from the German pharmaceutical market that authorized generics have no significant impact on entry of independent generic drugs.

2. Background and Hypotheses

In this section, we first summarize the relevant information related to market entry and the exclusivity period and then develop the hypotheses that we will test. To bring a new drug to a European national market, a firm requires market authorization (MA) from either a national authority, such as the Medicines and Healthcare Products Regulatory Agency in the UK or, as of 1995, from the European Medicines Agency. This process starts with the firm filing for a new drug application in case of a new molecular entity, or an abridged application for a generic drug. In the former case, MA is granted after establishing safety and efficacy via three phase clinical trials that take several years to complete, while in the latter case, the applicant references the safety and efficacy data of the originator's drug, and aims to establish therapeutic and bioequivalence to it.

Since a patent life is 20 years from the date of filing, and significant time is lost in drug development, the EU provides two routes that allow innovators to extend the exclusive marketing of their products. The first, available since 1993, is the Supplementary Protection Certificate (SPC) which allows originators to extend the patent for up to five years after the expiration of the original patent, or fifteen years from the first market authorization date in the EU, whichever is less. Second, there is an explicit data exclusivity period which was introduced in 1984 at the EU level. Prior to that, drug approval was at the national level and with varying rules, during which a generic firm was not allowed to reference the originator's data. In the European community, data exclusivity extended either to six years or ten years from the start of MA, depending on the member state (UK had ten years) and started from the date of first market authorization registered anywhere in the community.

Further, the data exclusivity protects original novel substances, for instance the molecule in the original drug, while subsequent improvements such as new therapeutic indications, dosage strengths, or formulations are not granted any additional protection. Nonetheless, these product line extensions may be protected via secondary patents. As of 2005, a new '8+2(+1)' exclusivity period applies, introduced at the EU level which provides unified rules of exclusivity across all member states – eight years of data exclusivity during which a generic cannot file for an abridged application, plus two additional years of market exclusivity, i.e., the generic may file the abridged application, but not market the drug, and a final one additional year of market exclusivity for new indication(s) if they constitute a significant clinical benefit. In general then, a generic entry application can be filed after the original 10 years of drug market exclusivity ends, but application or entry may be delayed further if there are other active patents still protecting the drug. Further details on entry and market exclusivity period are given in Appendix A.

Presentation proliferation can affect entry via two different potential mechanisms. In the first case, the originator could create horizontal differentiation from the original product. For instance, a new strength, or a different formulation such as a capsule may be better suited for some patients than others.

This would allow closer matching of the product characteristics with the preferences of the patient. On the one hand, this can also expand the market to newer patients for whom the original formulation was not ideal, and attract entry. However, if the new variant is protected via secondary patents (but not necessarily via market or data exclusivity as discussed above), then this form of market expansion is not likely to encourage entry. On the other hand, some of the existing patients would leave the original presentation and move to the newer variant as it is closer to their medical needs. This would make the market for the original presentation thinner, making entry less attractive as there are non-trivial entry costs.

Further, as mentioned earlier, presentation proliferation is likely to be increasing in the size of the market as it may be profitable to do so for reasons other than entry deterrence. Thus, in line with the insights from Ellison and Ellison (2011), a firm would deviate from its optimal investment strategy of additional presentations typically only in medium-sized markets. Information that its original drug is being reviewed by the EMA for a generic application would make the originator aware that entry is imminent. Accordingly, the firm would return to an optimal path of presentation proliferation. Such information will be received by incumbents who eventually see entry. Those who do not see any entry, may or may not receive any such information (the ambiguity, in this case, is if the generic application was not successful). Thus our first two hypotheses are as follows.

H1: In medium-sized markets, incumbents that experience entry would slow down their launch rate after the loss of exclusivity (LoE) and before any entry.

H2: By contrast, in the same markets, the effect on the proliferation rate of those that do not experience entry is ambiguous. They may continue with the same rate as before LoE or may increase it.

Continuing from the above discussion, we expect the proliferation strategy to be more successful if the incumbent does not withdraw from the original presentation as that would create an opening for a new entrant, albeit with a smaller market. Thus, our next hypothesis is as given below.

H3: In medium size markets, conditional on presentation proliferation, probability of entry is reduced if incumbent maintains patients more evenly across presentations.

The second mechanism for deterring entry via proliferation is when the originators create vertical differentiation between variants – or at least a vast majority of the patients consider the new presentation to be a superior product. An example is Shire pharmaceuticals which introduced a mixed amphetamine salt Adderall XR in the US in 2001. This was a once-a-day variant of their original drug Adderall and was introduced one year before generic entry for their popular attention deficit hyperactivity disorder (ADHD) drug. In 2001, the sales value of all ADHD drugs was over a billion dollars and the mixed amphetamine salts segment, provided only by Shire at the time, had over 31% of the ADHD market. Also, ADHD was considered a disorder that primarily, but not exclusively, affected school-aged children. Since providing additional dosages during school hours is difficult, particularly since there is also some stigma associated with mental health problems, a once-a-day drug provided a significant benefit for school-aged patients by eliminating the need for administering the drug during school. By contrast, the original variant may be administered multiple times and can be better suited for non-school-aged children or adults (via greater ability to inhibit the reuptake and/or promote the additional release of neurotransmitters in the brain). One year after generic entry in the mixed amphetamine salt market, Shire retained 70% of the market (by value) via its XR version and only 8.45% from the original non-XR version while the remaining 22% of this segment went to the two entering generics (Bokhari and Fournier, 2013).

While this did not prevent entry in this example, it highlights the case when introducing a vertically differentiated product can prevent entry in the original variant if all or nearly all patients are moved to the newer presentation. However, such a strategy is likely to be expensive and relatively rare where a new variant is superior and allows product hopping. Thus, we have the following additional hypothesis.

H4: Product hopping can deter entry in large markets.

3. Data and Descriptive Statistics

We use the 1996:Q3-2016:Q3 British Pharmaceutical Index (BPI) data series by Intercontinental Marketing Services (IMS) which provides national level sales for all drugs sold in the UK but disaggregated by individual items at the pack level. The BPI contains information in terms of total shipments by nominal sales value and various measures of quantity from wholesalers to retail pharmacies and dispensing doctors, but does not include direct sales from manufacturers to hospitals, or to non-pharmacy stores (e.g. grocery stores). Drugs are identified by manufacturer (except for generics), product name, which is either a brand name or its international non-proprietary name in the case of generics, main/active molecule(s), and strength and package size e.g. 20mg 28pills. In our data, the identity of a generic drug's manufacturer is typically not known but other information about the drug is known. For each item, the data lists its associated four-digit anatomical therapeutic chemical code (ATC4) and a three-digit code for formulation (NFC3), which tells us what the drug is used for and its route of administration, and whether it is a tablet, a capsule, an extended release version, an ointment or some other formulation. The data also includes information on whether a drug is branded or generic, and the month and year a pack was first launched in the UK. We use the UK launch date as a proxy of actual market authorization, and forty quarters from then as the end of exclusivity.

We combine the information on ATC4 with the molecule name to identify an originator as the manufacturer with the first launch date on any individual drug within the ATC4-molecule combination. Thus our originator is the first firm to provide a molecule within a ATC4 therapeutic class. Accordingly, our unit of analysis is the originator over time, or equivalently, molecule in a therapeutic class over time. Line extensions by the originator are all other drugs in the same ATC4-molecule combination that differ either just by formulation, i.e. tablet, capsule, liquid, etc. (given by the NFC3 code), or by pack variety, i.e. different dosage or pack size, and with a later launch date. Entry by a competitor is identified in a similar manner, i.e., when a drug is introduced by another manufacturer which is in the same ATC4-molecule combination

(in that respect an entrant could be a generic or a branded competitor with a 'me-too' drug which has the same ATC4-molecule but perhaps a different formulation).

TABLE 1. Risk Sets And Entry Events

Originator's class and formulation		Loss of Exclusivity Period			
		1996 - 2016		2001 - 2011	
		At risk	Entry	At risk	Entry
A	Alimentary t.& metabolism	32	9	21	8
B	Blood + b.forming organs	26	3	15	1
C	Cardiovascular system	51	23	24	15
D	Dermatologicals	23	4	15	2
G	G.u.system & sex hormones	28	11	20	8
H	Systemic hormones	10	0	6	0
J	Systemic anti-infectives	65	14	33	6
L	Antineoplast+immunomodul	47	14	35	13
M	Musculo-skeletal system	27	10	14	5
N	Nervous system	70	42	48	28
P	Parasitology	4	0	3	0
R	Respiratory system	22	5	14	4
S	Sensory organs	25	2	15	2
Total		430	137	263	92
Solid	Tablets, capsules, extend release, etc.	194	91	121	62
Liquid	Liquids & aerosols	40	6	22	3
Injection	Ampules, vials, pre-filled syringes, etc.	115	23	67	14
Ointment	Ointments, creams, gels & sols	26	2	18	2
Other	All others & multiple formulations	55	15	35	11
Total		430	137	263	92

Notes. The data contains 181 ATC4 classes (132 ATC3, 64 ATC2, and 13 ATC1 classes) and 78 values for NFC3 formulation codes. The latter are collapsed into simplified formulation classifications. See data appendix Table A-1 for details.

For our analysis we constructed two primary data sets. The first data set was constructed to observe sales and product launches by originators before and after the expiration of their market exclusivity period. Since our data series is for 1996-2016, and market exclusivity lasts for ten years since the initial launch, we selected those originators that would have lost exclusivity between 2001-2011 i.e., working backward, their UK launch dates were between 1991-2001. This window gives us at least five years of observations if exclusivity ended as early as 2001, and at least five years after the end of exclusivity if it ended as late as 2011. This resulted in a final data set of 263 originators consisting

of 58 distinct firms (as some firms are originators in multiple classes), and of these 263 original drugs, 92 experienced entry by a competitor in our data. The second set consists of all originators with the loss of exclusivity anytime between 1996 and 2016. This larger data is used in hazard models to estimate the probability of entry by a competitor and consists of 430 originators as 70 distinct firms at risk of competitive entry. Of these, 137 actually experienced entry. Table 1 gives a summary of originators by first-digit therapeutic classes as well as by (simplified) formulations of the original drug. Notably however, entry can in fact also happen before the 10th year, as was the case for 30 originators in the larger data set used in the hazard models. The reasons could be a prior launch elsewhere in the EU or after patent litigation. Further details about selection criteria and data cleaning are given in Appendix A.2.

There is significant variation in entry by therapy classes and formulations. The nervous system class has the highest number of entries by originators as well as by competitors where, of the 70 originators that entered in this class, 42 experienced entry by a competitor. This is followed by molecules treating anti-infectives for systemic use with 65 originators and 14 competitor entrants. Others such as parasitology draw very few originators and competitors. Similarly, among formulations, ointments draw fewest entries while solid form drugs, e.g. tablets, capsules etc. have the highest entry rates.

Table 2 provides summary statistics of all the variables related to the 430 originators and used in the analysis (for the smaller sample of 263 originators, descriptive statistics are very similar and are given in Table A-2 in the appendix). For each originator, we count product line extensions with two main measures. The first is D1, which provides for any given originator a count of the total number of drugs in the same ATC4-molecule class that differ by their formulation code (the NFC3 code can take upto 78 unique values). For instance, if the original drug was launched as a regular tablet, and at some point the originator launches an extend release tablet or a capsule then that would be counted as an additional formulation. However any variation by dosage or pack size differences are ignored, and do not increment the count. Our second measure is D2, which further allows for variation by dosage or

TABLE 2. Originator's characteristics: full sample (430 Originators)

Variable	Description	Mean	Std. Dev.			Min	Max
			overall	between	within		
$D1$	Count based on formulations	1.35	0.67	0.58	0.30	1	5
$D2$	Count based on pack variations	3.23	3.40	3.18	1.33	1	37
$S1$	1/HHI from shares of D1	1.12	0.29	0.23	0.15	1	3.6
$S2$	Share of D1 launched after 5 years	0.05	0.20	0.16	0.10	0	1
Sales (log)	Sales by originator	10.91	4.31	4.05	2.05	0	18.3
†Monopoly	Originator monopolist in other classes	0.92	0.27	0.25	0.13	0	1
†Nearby	Other monopolists in ATC3 class	0.82	0.39	0.34	0.19	0	1
†Chronic	Chronic disease drug	0.71	0.46			0	1
†SPC	Originator enters after 1993	0.74	0.44			0	1
†1Form	Single original formulation	0.94	0.23			0	1
†Solid	Tablets, capsules, extend release, etc.	0.43	0.50			0	1
†Liquid	Liquids & aerosols	0.10	0.30			0	1
†Injection	Ampules, vials, pre-filled syringes, etc.	0.27	0.45			0	1
†Ointment	Ointments, creams, gels & sols	0.07	0.25			0	1
†Other	All others & multiple formulations	0.12	0.33			0	1

Notes. Summary statistics from unbalanced panel of 430 originators over 80 quarters with 21,670 observations. For time invariant variables, there is no *within* standard deviation and overall standard deviation is the same as *between*. For the smaller sample with 263 originators, see Table A-2 in the appendix.†1/0 Dummy variable, 1 if true.

pack size as well. Thus D2 $\geq$ D1. Based on formulations, on average, an originator has 1.35 drugs in their portfolio with a standard deviation of 0.67 and a maximum of 5 drugs. Most of the variation is cross-sectional as indicated by *between* standard deviation, which is 0.58, while the *within* standard deviation is 0.30 ('within' is due to variation over time for a given originator). Based on the second measure, originators launch 3.23 presentations with a standard deviation of 3.40 and a max of 37.

Since we are also interested in measuring the effect of product proliferation and product hopping on entry, i.e., if an originator has not just launched an

extension of the ordinal drug, but has successfully moved patients to the newer drugs fully or spread them evenly into all variants, we use relative market share of an originator's products to compute two additional variables. The first is the inverse of the Herfindahl-Hirschman index constructed from market share s_{ij} (by value) over the set of $\mathcal{N}_j$ formulations sold by the j-th originator, i.e., $S1_j = 1/\sum_{i\in\mathcal{N}_j}(s_{ij}^2)$. This measure can be thought of as a count of formulations, but only if the formulation has a significant share of an originator's portfolio. A second measure is the total share of all formulations launched by an originator after five years of their original launch, $S2_j = \sum_{i\in\mathcal{N}5_j}(s_{ij})$, where $\mathcal{N}5_j$ is the subset of new formulations launched by the originator after five years of initial entry. Note that both measures are inclusive of shares of all pack variations (dosage or strength variation) within a formulation. The mean value of S1 is 1.12, indicating that not all formulations retain significant market share, while the mean share of drugs launched after five years is only .05 of the originator's portfolio, but sometimes going up to 1.0. Nonetheless, variance in this measure is less than the S1 variable, both overall as well as within and between. For instance if we define an indicator variable as $I(S2 > 0.5)$, then only 19 out of 430 originators, or 4.42%, engage in product hopping.

In the analysis that follows, we also use several additional variables about the originator or their initial drug. The variable (log) sales is the sum of sales from all drugs by the originator within the ATC4-molecule class and is recorded for each period (converted to constant 2015 value using the consumer price index for the UK). The mean value is 10.91 with almost twice as much variation between originators than over time (within). While not shown in this table, we also used the time invariant value of sum of sales over the two years prior to the loss of market exclusivity (or two years prior to entry in the handful of cases where entry occurred before the loss of exclusivity) to classify the originators as belonging to small, medium and large markets. The mean and median of log of sum of sales over these two years is 13.19 and 13.88 respectively, and we used the 33rd and 66th percentile values of the distribution, 12.29 and 15.43 respectively, to classify the originators into the three equal sized groups. Other variables include whether the originator entered before or after 1993 (when SPC came into effect), type of originator's original formulation, whether the

originator entered with a single formulation (23 originators entered with more than one formulation), codes of therapeutic class (they are used at two digit level in most of the regression analysis), whether the original drug is for a chronic disease or not, if the originator is a monopolist in any other class, and finally whether there are other monopolists in the same ATC3 class as the reference drug. Summary statistics of these variables are also given in Table 2.

4. RESULTS

Product launches by the originator. To test whether originators launch additional products to deter entry, we checked if the product launch rate changes before and after the loss of exclusivity (LoE), and *prior to any entry* by a competitor. Figure 1 plots the value of counts of products over time (left for D1 and right for D2) and the vertical line marks the LoE period. For either graph, the black dotted line in the middle shows the count of products for the sample of 263 originators described earlier. As can be seen, there is a small change in slope before and after LoE, which is more obvious for D2 than D1.

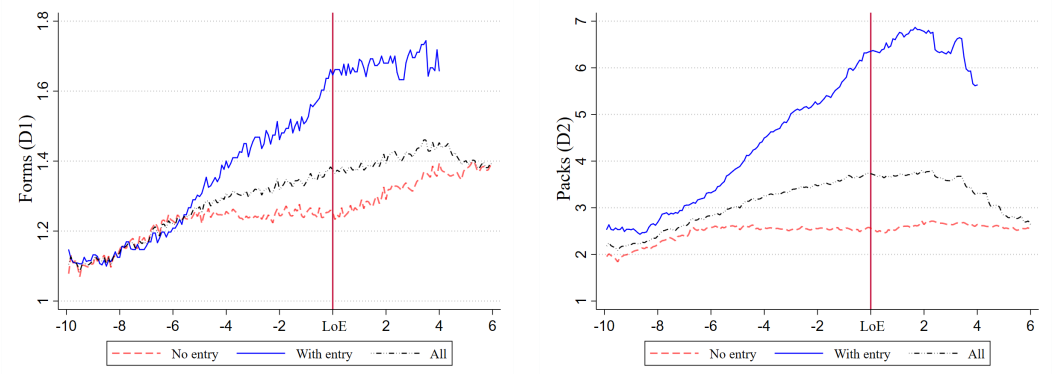

FIGURE 1. Count of products by originators

Next, we looked at the counts over time by sub-samples as discussed earlier: originators that eventually experienced entry versus those that did not experience competitive entry (solid/blue line and dashed/red line respectively). The mean value of counts of products for the group with entries (but before entry) is higher than for the mean value for the group without any entries. The mean values of D1 are 1.51 and 1.12 for the two subsamples (with and without entry

respectively), and similarly those of D2 are 5.45 and 2.24 in the subsamples (these stats are for the cross-section two years before LoE but the numbers are similar for the panel). There are other differences in these two subsamples as well. For instance, the mean of log sales is 14.04 for the group that experienced entry, and 9.13 for those that did not, confirming that in our data too, entry is more likely in larger markets as often observed in the literature (additional statistics by entry status are given in Table A-3 in the appendix).

As the graph indicates, not only is the overall count for the two groups different, the launch rate, as measured by the slope is higher for the originators that experience eventual entry than those that do not. This is particularly obvious before the LoE. However, there is a discernible *change* in the slope for the originators that experienced entry after the LoE compared to those that do not experience any entry. For D1 (the graph on the left), the slope decreases after LoE for originators who experienced entry, and increases for those without entry. For D2 (the graph on the right) there does not appear to be a change in slope after LoE for originators without entry, but it decreases for originators with entry.

Whether an originator eventually sees entry or not is certainly not exogenous, nor is the date on which entry takes place (though the loss of exclusivity date is exogenous). However, in the regression analysis that follows where we test whether the slopes indeed change, we rely on two key factors for identification. First, that the date of LoE is predetermined, and hence the change in threat level of entry is exogenous for all originators. Thus originators could respond to this change in threat by potentially launching more products. Second, since the potential competitors can file for market authorization/entry only after the LoE, and it takes time before competitors can obtain the required authorization and actually enter, originators actually find out whether entry is imminent or not. These originators may then subsequently change their behaviour based on this information. While we do not observe the exact date on which this subset of originators becomes aware of imminent entry, on average they would find out before entry and after LoE, and it is during this

window that we should see them reduce launches if indeed deterrence was the motive.

To test whether there is a change in launch strategy, we estimated a reduced form equation for the total number of drugs by all 263 originators as a function of time,

$$D_{jt} = \beta_0 + \beta_1 T_{jt} + \beta_2 LoE_{jt} + \beta_3 T_{jt} LoE_{jt} + \beta_4 \ln(\text{Sales})_{j,t-1} + X_{j,t-1}\gamma + \epsilon_{jt}. \quad (1)$$

In the equation above, D is either D1 or D2, LoE is a 1/0 indicator variable equal to one after the LoE event, T is time to LoE, negative before and positive afterwards. The variable $\ln(\text{Sales})_{j,t-1}$ is one period lagged value of sales, and $X_{j,t-1}$ is a vector of other variables listed in Table 2, plus dummy variables for ATC2. Some of these variables are time invariant, but those that are not enter with lagged values. Our interest is in whether product launch rate changes after LoE, and hence in the coefficient β_3 which measures if rate is different before versus after LoE. We estimated the equation given above for the full sample via pooled OLS as well as by various subsamples of interest. Later we also describe results from random effects estimation as well as from Poisson distribution for count of products.

Table 3 shows selected regression coefficients for different sub-samples, along with robust and clustered standard errors, where clustering is at the originator level. The full set of regression coefficients are given in the appendix in Table B-1 (in some of the subsamples to follow the number of originators becomes too few and it does not make sense to always use clustered standard errors, hence we provide two types of standard errors, see Cameron et al. (2008)). Columns (1)-(3) correspond to when the dependent variable is D1, and (4)-(6) when it is D2. Initially, in sample A, we used all the observations for the 263 originators that reached the LoE within the 2001-2011 period (results are in columns (1) and (4) for D1 and D2 respectively). The average increase in number of products over time is .006 and .039 per quarter for D1 and D2, and both are statistically significant. Further, the interaction terms are negative and significant (-.005 and -.056, respectively) indicating a decrease in launch rate after the LoE for the overall sample of originators.

TABLE 3. Product launch rate

	D1 and D2 by samples A,B,C						D1 and D2 by subsamples of C			
	D1			D2			D1		D2	
	(1) A	(2) B	(3) C	(4) A	(5) B	(6) C	C(7) With	C(8) W/out	C(9) With	C(10) W/out
T	0.006	0.006	0.003	0.039	0.039	0.028	0.009	0.001	0.088	-0.005
	(0.001)^a	(0.001)^a	(0.001)^b	(0.003)^a	(0.003)^a	(0.008)^a	(0.003)^a	(0.001)	(0.013)^a	(0.008)
	[0.001]^a	[0.002]^a	[0.002]^c	[0.007]^a	[0.007]^a	[0.010]^a	[0.003]^a	[0.002]	[0.018]^a	[0.007]
LoE	0.017	-0.001	0.018	-0.030	-0.001	0.182	-0.004	0.021	0.315	0.070
	(0.018)	(0.019)	(0.025)	(0.103)	(0.104)	(0.139)	(0.052)	(0.022)	(0.228)	(0.136)
	[0.026]	[0.026]	[0.018]	[0.134]	[0.126]	[0.091]^b	[0.038]	[0.018]	[0.193]	[0.079]
LoE×T	-0.005	-0.005	-0.001	-0.056	-0.054	-0.053	-0.008	0.003	-0.098	0.001
	(0.001)^a	(0.001)^a	(0.002)	(0.004)^a	(0.004)^a	(0.011)^a	(0.005)^c	(0.002)^c	(0.020)^a	(0.012)
	[0.002]^c	[0.002]^c	[0.003]	[0.012]^a	[0.012]^a	[0.015]^a	[0.007]	[0.003]	[0.029]^a	[0.012]
Obs	13,559	12,560	8,052	13,559	12,560	8,052	2,325	5,727	2,325	5,727
R^2	0.41	0.421	0.425	0.384	0.396	0.423	0.551	0.533	0.666	0.383
Originators	263	212	212	263	212	212	66	146	66	146

	Ho: LoE×T equal across samples†	
	C7 v C8	C9 v C10
Chi2(1)	2.80	10.54
p-value	.094	.001

Robust standard errors in parentheses followed by clustered standard errors in brackets. Superscripts a, b, c indicate significance at 1%, 5% and 10% levels, respectively. All regressions include other controls and ATC2 dummies. See Table B-1 for full set of coefficients. Sample (A) is all initial 263 originators, (B) is restricted to 212 originators with sales observed before/after LoE, and (C) is the same as (B) but with observations restricted to within 5 years of the LoE.
†Test of equality of the interaction term across samples restricted to the originator with and without eventual entry.

Some originators in our sample do not have observations both before and after the LoE. For instance, we may observe sales for a given original drug in our data only after 2006, even though the drug entered the UK market in 1995 and reached its LoE period in 2005. Thus, we restricted the sample further to 212 originators for whom we have observations both before and after the LoE, and re-estimated Equation 1. The results for sample B are summarized in columns (2) and (4) for D1 and D2. The coefficients of interest do not change by much either in magnitude or in significance levels.

We imposed one final restriction on the combined sample (sample C), where we required that all observations for an originator be within five years of the LoE. This is so that observations that are too far before or after the LoE do not contribute to the measurement of change in slopes, as there may be other factors unrelated to LoE that can affect product launches as well. Doing so does not reduce the number of originators in the sample any further, only the periods over which they are followed which leads to a drop in observations from 12,560 to 8,052. The results from this sample are given in Columns (3) and (6). The magnitudes of the coefficients decrease slightly, but the overall pattern remains as before. The interaction terms remains negative and significant (significant for D2 but not for D1 anymore), indicating that there is an overall slowdown in product launches after the LoE.

To further investigate which originators changed their product launch, we created two sub-samples of C based on whether the originators eventually experienced entry or not (66 and 146 originators respectively). We need to be careful in interpretation as assignment into the two groups is not a random allocation. Nonetheless, grouping by ex post outcome of entry or no entry is still a useful exercise to learn how these firms differ, if at all, in their ex ante actions. The results are given in columns (7) and (8) for D1, and in columns (9) and (10) for D2. Focusing again on the interaction terms, we can see that for both measures D1 and D2, there is a decrease in product launch during the window from LoE to actual entry for originators that experience eventual entry (see columns 7 and 9). However, those that do not experience entry (columns 8 and 10) either continue with the same rate of product launch

as before LoE, or slightly increase their product launch, though the evidence for the latter statement is weaker. We further tested whether the interaction terms in the two subsamples (with/without) differ from each other. The null that the slopes are equal across the subsamples is rejected both for D1 (p-value = .094) and for D2 (p-value = .001), thus indicating that the launch strategies differ for these two groups of originators post LoE.

Matched samples and market size. As noted before, there are observable differences in originators that experience entry versus those that do not (see Table A-3) and the assignment is non-random. In part these could be driving our results. To account for that, as well as to investigate if there is heterogeneity in results by market size, we re-estimated the last sample (sample C) by further subsamples based on market size being small, medium or large as previously described. One advantage of this approach is that it generates subsamples that are far better matched in terms of the observable differences between originators with and without entry, at lest for the small and medium sized markets. See Table A-4 which reports covariates by entry status for the medium sized markets (there is a similar good match in small markets but less so in the large markets).

Thus we estimated the reduced form regression for each combination of market size and subsamples based on with and without entry for both D1 and D2 (a total of $3 \times 2 \times 2 = 12$ regressions). Selected coefficients are shown in Table 4. The top panel is for D1 and the bottom panel is for D2 (full set of coefficients are in the online appendix in Table B-2 and Table B-3). The main result that stands out is that in the medium-sized markets, the interaction coefficients are negative and positive respectively for originators with and without entry, and that the coefficients are significantly different from each other. In small markets, the coefficients are neither significantly different from zero nor from each other, where as in the large markets, the same is true for D1. However for D2 we have the anomalous result that the interaction terms are negative and significant, but not different from each other (but even this is not true per the clustered standard errors, which indicate that the slopes are not statistically different from zero). Further, as noted above, the sample is less well matched in

TABLE 4. Product launch rate by market size

	Small		Medium		Large	
	With	W/out	With	W/out	With	W/out
D1	(1)	(2)	(3)	(4)	(5)	(6)
T	0.000	-0.002	0.013	-0.002	0.002	0.008
	(0.003)	(0.001)	(0.003)^a	(0.002)	(0.003)	(0.003)^a
	[0.001]	[0.002]	[0.010]	[0.003]	[0.004]	[0.008]
LoE	-0.057	0.021	-0.122	0.035	0.057	-0.005
	(0.038)	(0.020)	(0.045)^a	(0.026)	(0.055)	(0.038)
	[0.060]	[0.024]	[0.065]^c	[0.026]	[0.049]	[0.052]
LoE×T	0.001	0.001	-0.019	0.008	0.002	-0.001
	(0.003)	(0.002)	(0.004)^a	(0.002)^a	(0.005)	(0.004)
	[0.001]	[0.003]	[0.015]	[0.004]^c	[0.007]	[0.011]
†$\chi^2(1)$	0.001		3.082		0.062	
p-value	0.9723		0.0792		0.8036	
D2	(1)	(2)	(3)	(4)	(5)	(6)
T	0.020	-0.011	0.032	-0.021	0.060	0.030
	(0.010)	(0.004)^a	(0.006)^a	(0.008)^a	(0.014)^a	(0.007)^a
	[0.020]	[0.008]	[0.015]	[0.012]^c	[0.023]^b	[0.016]^c
LoE	0.020	0.033	-0.019	0.092	0.577	0.137
	(0.156)	(0.056)	(0.117)	(0.101)	(0.235)^b	(0.121)
	[0.198]	[0.062]	[0.179]	[0.093]	[0.239]^b	[0.231]
LoE×T	-0.018	0.002	-0.052	0.028	-0.049	-0.043
	(0.012)	(0.005)	(0.009)^a	(0.010)^a	(0.022)^b	(0.011)^a
	[0.026]	[0.009]	[0.032]	[0.015]^c	[0.042]	[0.037]
†$\chi^2(1)$	0.627		5.323		0.014	
p-value	0.4283		0.0211		0.9051	
Observations	159	2,272	550	2,463	1,616	992
Originators	5	60	15	61	46	25

Robust standard errors in parentheses followed by cluster standard errors in brackets. Superscripts a, b, c indicate significance at 1%, 5% and 10% levels, respectively. Full set of coefficients are in the online appendix in Table B-2 and Table B-3.

†Test of equality of the interaction term across samples restricted to originator with and without eventual entry.

large markets. Thus, our results indicate that firms with and without eventual entry differ in their launch strategies post LoE and before actual entry, and the differences are most pronounced in the medium-sized markets.

Robustness. We repeated the analysis using linear random effects model. Table B-4 and Table B-5 in the appendix provide results analogous to those reported in Table 3 and Table 4 and show very similar estimates and significance levels. We also estimated the model using truncated Poisson distribution (our dependent variable starts with minimum value of one). While this has the advantage that it is a count based model, coefficients on interactions terms are not always straightforward to interpret in non-linear models (see Ai and Norton, 2003, Athey and Imbens, 2006, Puhani, 2012) and hence we focused on the linear version in the main analysis first. Nonetheless, results equivalent to those of Table 3 for count models are shown in Table B-6 and lead to similar conclusions. Finally, we also estimated the truncated Poisson model by market size. Because of lack of variation in D1, this model does not converge but does so for D2. Results are summarized in Table B-7. They too are consistent with the main analysis.

Monotonicity test. We also implemented a test for strategic deterrence based on a monotonicity argument proposed in Ellison and Ellison (2011, 2000 WP paper) and in Dafny (2005). As explained there, originators would launch more products in larger markets, and hence there should be a monotone increasing relation between product launches and size of the market. However, if there were an entry deterrence motive as well, then the originators would launch more products in medium-sized markets relative to when there is no deterrence motive. The test and its results are described in greater detail in Appendix B.2, but generally they do not support strategic deterrence story for our data. We rule those out because the relationship between number of products and market size is very convex in our data, a condition which invalidates the monotonicity test.

Probability of Entry. Next we estimated hazard rate models to assess the impact of product line extensions on the probability of entry. One way to

proceed would have been to model the entry decision by a potential entrant, where the firm enters the market when the future discounted profit from entry becomes greater than zero, and the hazard is modeled as a function of the entrant's characteristics interacted with the market characteristics (see Reinganum, 1989, Bokhari, 2009). However, because our interest is not whether a generic manufacturer with given characteristics enters a specific market, but rather whether an incumbent moving patients to the newer formulations reduces the probability of entry in their market by any generic manufacturer. To that end, we used hazard models to assess the impact of how well these additional drugs have diffused in the patient population on the probability of entry by a competitor in the therapy-molecule class of the originator. Thus, let $\lambda_j(\tau)$ be a continuous time hazard that incumbent j experiences entry at time τ and is given by the proportional form $\lambda_j(\tau) = \lambda_0(\tau)exp(Z_j(\tau)'\beta)$ where $\lambda_0(\tau)$ is the baseline hazard, and $Z_j(\tau)$ is a vector of time varying covariates of the originator. We can generate a discrete time hazard from this by grouping time τ along the quarterly intervals $[0, \tau_1), [\tau_1, \tau_2), \ldots, [\tau_{t-1}, \tau_t), \ldots, [\tau_l, \infty)$. Then λ_{jt}, the probability that originator j experiences entry in quarter t conditional on no entry until the previous quarter, is given by

$$\begin{aligned} \lambda_{jt} &= \Pr[\tau_{t-1} \leq T_j < \tau_t | T_j \geq \tau_{t-1}] \\ &= 1 - exp\{-exp(Z'_{jt}\beta + \alpha_t)\}. \end{aligned} \tag{2}$$

In the equation above, α_t is the natural log of the baseline hazard within an interval $[\tau_{t-1}, \tau_t)$ and is given by $\ln \int_{\tau_{t-1}}^{\tau_t} \lambda_0(s)ds$ (see Cameron and Trivedi, 2005). The vector Z consists of variables D, S, X and their interactions, where D is one of the variables in $\{D1, D2\}$, and S is a measure of the extent to which patients use these additional drugs, i.e., one of the variables in $\{S1, S2\}$. The variable X includes size of the market (dummy variables M and L for medium and large) as well as other product or originator characteristics listed in Table 2. Specifically,

$$Z'_{jt}\beta = \beta_1 D_{jt} + \beta_2 S_{jt} + \beta_3 M_j + \beta_4 L_j + \beta_5 S_{jt} M_j + \beta_6 S_{jt} L_j + \beta_7 \ln(\text{Sales})_{j,t-1} + X_{j,t-1}. \tag{3}$$

Figure 2 illustrates the survival probability over time as a monopolist and is grouped by market size. Recall that we defined small, medium or large

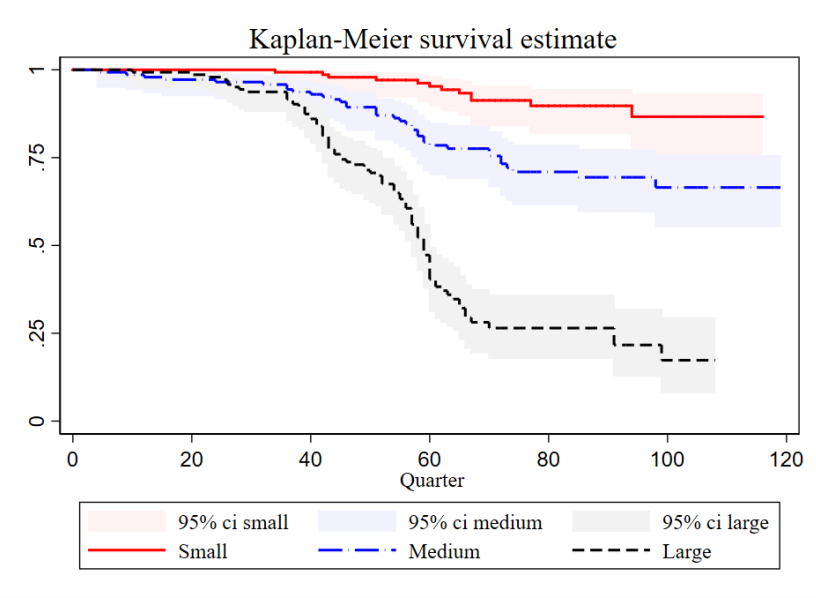

FIGURE 2. Survival by market size.

market based on sales value over the two years prior to the LoE (there were 30 cases where originators experienced entry prior to LoE, and for those cases we defined small, medium or large market based on sales value two years before entry). The Kaplan-Meier curves show that entry probability differs by market size. Unsurprisingly, probability of entry is lowest when market size is small, as an entrant can expect lower profits post entry in these markets. Entry probability is higher when market size is large, and entries are more concentrated between the 10th and 15th year since the launch of the original drug. Entry probability for medium-sized market is located between the other two market sizes.

We already know from descriptive statistics and from Figure 1 that firms that experience entry launch more products than those that do not. Consequently, a hazard model as described above with $D1$ or $D2$ on the right hand side will capture this positive correlation, but cannot be interpreted as causal, nor does it shed any extra light over what we already know. Instead, our primary interest is in the impact of $S1$ and $S2$ and their interaction with categorical variables for market size (medium and large) on the probability of entry after controlling for $D1$ (or $D2$) and other variables. The hazard model given in

(2) is estimated under four different specifications which differ by variables controlled for in the model, the observations used, or the time period used to compute the value of $D1, D2, S1, S2$ and ln(sales).

Selected regression coefficients and clustered standard errors for the variables of interest are shown in Table 5, and the coefficients of other variables are available upon request. Columns (1) to (4) show estimates when $S = S1$, and columns (5) to (8) provide estimates when $S = S2$ for each of the four specifications respectively. Note that the hazard models for $S2$ have fewer observations since this variable measures the relative share of originators drugs that were introduced after five years, and hence observations for the first five years are omitted. Columns (1) and (5) are estimated on the baseline sample of 430 originators, include the duration dummies, and the variables listed in the table, but do not control for other product/originator characteristics listed in Table 2, nor do they include any of the ATC2 dummy variables. By contrast, columns (2) and (6) include all these additional variables as controls. An alternative set of regressions that control for $D2$ instead of $D1$ give similar results, and are available in the appendix in Table B-9.

Columns (1,2,5, and 6). Starting with these four columns (1,2,5 and 6), the probability of entry increases in log sales in all four cases as expected. The dummy variables for size of the market, medium or large, are either positive and significant, or if negative (as in column 5) then it is not significant. The omitted base category is small markets. The coefficient for the variable S can be interpreted as value of $S1$ or $S2$ in small markets. This coefficient is positive and significant in all four columns except in column (2) when it is not significant for $S1$ when additional controls are included in the specification. The positive coefficient implies that probability of entry increases in small markets when originators have more drugs that are well diffused among the patient population.

Importantly however, the interaction terms with the dummy for medium and large size markets are negative and significant in all but one case (the negative coefficient on the interaction between $S1$ and 'Large' is not significant in column (2) when other controls are included in the hazard model). The negative

TABLE 5. Discrete time hazard models for S1 & S2 (controlling for D1)

	S=S1				S=S2			
	(1-A)	(2-B)	(3-C)	(4-D)	(5-A)	(6-B)	(7-C)	(8-D)
S	1.337^{b}	0.731	1.332	1.840	2.379^{a}	2.692^{a}	2.833^{a}	3.259^{a}
	(0.671)	(0.874)	(1.077)	(1.836)	(0.826)	(0.818)	(1.017)	(1.043)
Medium	3.253^{a}	4.479^{a}	5.163^{a}	6.299^{b}	-0.060	0.506	0.286	0.348
	(1.156)	(1.436)	(1.757)	(2.451)	(0.361)	(0.545)	(0.659)	(0.766)
Large	2.168^{b}	2.242^{c}	2.378	3.294	0.488	1.602^{b}	1.080	1.011
	(0.944)	(1.268)	(1.591)	(2.280)	(0.532)	(0.780)	(1.029)	(1.145)
Medium× S	-3.159^{a}	-3.931^{a}	-4.790^{a}	-5.873^{a}	-3.187^{b}	-3.587^{a}	-3.675^{b}	-3.888^{b}
	(0.990)	(1.230)	(1.476)	(2.067)	(1.277)	(1.271)	(1.442)	(1.717)
Large× S	-1.816^{b}	-1.366	-1.798	-2.942^{c}	-3.457^{a}	-3.658^{a}	-4.029^{a}	-6.029^{a}
	(0.749)	(1.008)	(1.171)	(1.785)	(1.069)	(1.271)	(1.429)	(1.476)
Sales (log)	0.274^{a}	0.244^{b}	0.359^{b}	0.465^{a}	0.261^{a}	0.185^{c}	0.331^{b}	0.436^{a}
	(0.077)	(0.105)	(0.146)	(0.139)	(0.081)	(0.097)	(0.146)	(0.138)
D1		0.362^{b}	0.307	0.364		0.200	0.135	0.027
		(0.161)	(0.192)	(0.241)		(0.136)	(0.167)	(0.194)
				Marginal effects: $\partial\lambda/\partial S$ (×100)				
Small	0.402^{b}	0.242	0.660	0.911	0.807^{b}	1.016^{a}	1.399^{b}	1.604^{b}
	(0.203)	(0.284)	(0.525)	(0.852)	(0.407)	(0.391)	(0.691)	(0.734)
Medium	-1.281^{b}	-2.456^{a}	-3.489^{a}	-4.072^{a}	-0.563	-0.692	-0.851	-0.635
	(0.545)	(0.766)	(1.211)	(1.346)	(0.676)	(0.827)	(1.227)	(1.572)
Large	-0.967	-1.445	-1.934	-4.563	-2.484	-2.575	-4.977	-11.482^{b}
	(0.600)	(1.54)	(2.853)	(3.207)	(1.557)	(2.791)	(4.604)	(4.714)
Includes X_j?	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Originators	430	386	312	312	410	364	312	312
Entry events	137	137	104	104	131	131	104	104
Observations	13,456	12,063	5,961	5,961	11,848	10,444	5,961	5,961
Log likelihood	-664	-606	-412	-413	-623	-563	-415	-415

Notes. Clustered standard errors are in parenthesis and superscripts a, b, c indicate significance at 1%, 5% and 10% levels, respectively. All models include duration dummies and ATC2 dummies included in all but columns (1) and (5). Table B-9 provides similar estimates but controlling for $D2$ instead of $D1$.

coefficients on these interaction terms imply that the probability of entry does not increase in medium and large markets as much as in the base case of small markets. Whether the probability actually decreases or not depends on the magnitude of these negative coefficients relative to the positive magnitude of the coefficient on S for small markets. In fact the probability decreases in all cases for both the medium and large markets, but we postpone that discussion until we discuss marginal effects.

Columns (3 and 7). We next removed the 30 originators that experienced entry before the LoE period and retained all other variables in the specifications in columns (2) and (5). Since in this case there is no entry event in the first ten years for any originator, it also required dropping all observations for these years in the hazard model as non-entry is predicted perfectly. This reduced the sample size considerably from ∼12k and ∼10.4k to ∼5.9k. The results are given in columns (3) and (7) for $S1$ and $S2$ respectively. Compared to the previous case, all coefficients retain their sign, and most increase in magnitude with some exceptions. Also the coefficient for $D1$ is no longer significant. Overall however we see a similar pattern for the interaction terms indicating as before that the probability of entry does not increase with S as fast in medium or large markets as in small markets (and in fact once again, it actually decreases with S in these markets).

Columns (4 and 8). Our final and preferred specification uses the same sample as in the previous case, but now measures the value of $S = \{S1, S2\}$, $D1$ and ln(sales) not as one period lagged values, but instead are time invariant values equal to their average value two years prior to the LoE. Note that this could not have been done in the first two specifications as those also included observations from the first ten years when originators were considered at risk. Results from this change are given in columns (4) and (8) and are very similar to those reported earlier. The main difference is that coefficients generally increase in magnitude, and particularly for the interaction terms.

Marginal Effects. Because of the interaction terms, the marginal effect of a variable may not have the same sign as that of the coefficient on the interaction

term. If we rewrite (2) as $\lambda_{jt} = 1 - exp\{-exp(I_{jt})\}$, where $I_{jt} = Z'_{jt}\beta + \alpha_t$, then the marginal effect with respect to S, $\partial\lambda_{jt}/\partial S_{jt}$, is given by $(1 - \lambda_{jt})(\beta_2 + M_j\beta_5 + L_j\beta_6)$, where the sign of the marginal effect depends on the sum of the coefficients $(\beta_2 + \beta_5)$ or $(\beta_2 + \beta_6)$ in medium and large markets respectively. The marginal effect can be computed at either the mean of the sample or for each data point separately, and the standard error can be computed using the delta method. A difficulty in the first case is that there is a very large number of dummy variables in our specifications. These include not only the duration dummies, but also for ATC2 classes and most of the other variables listed in Table 2, and hence either the predicted probability or the marginal effect at the mean can be difficult to interpret. Thus instead we provide the mean marginal effects with respect to $S1$ or $S2$ in the lower part of Table 5. The marginal effect with respect to $S1$ or $S2$ is negative in all cases for medium and large markets and positive in small markets. Particularly, the marginal effect for $S1$ in medium-sized markets is negative and significant in all specifications, indicating that entry is less likely in these markets if originator has more products and patients are spread more evenly across all of the originator's products. However, the marginal effect for $S2$ is negative and significant in large markets for only the last specification, and remains positive and significant for small markets. In turn, it implies that perhaps product hopping, where most patients are switched over to the newer drug, is successful in preventing entry in only large markets, while in smaller markets there may be other demand driven reasons for switching to newer formulations. Regardless, the evidence of its effectiveness in large markets is not robust to specifications, as the marginal effect is not significant in the initial three cases (though the interaction terms are significant).

Figure 3 additionally plots the mean of the predicted values of λ against values of $S1$ and $S2$ in medium and large markets. The slope of the plotted line is equal to the marginal effect, $\partial\lambda/\partial S_j$. To be clear, we computed λ based on the coefficients from the last specification for each data point, and where $S1$ and $S2$ were varied over the specified range in the graph, but other variables were held at their observed value. The error bars are equal to the 95% confidence

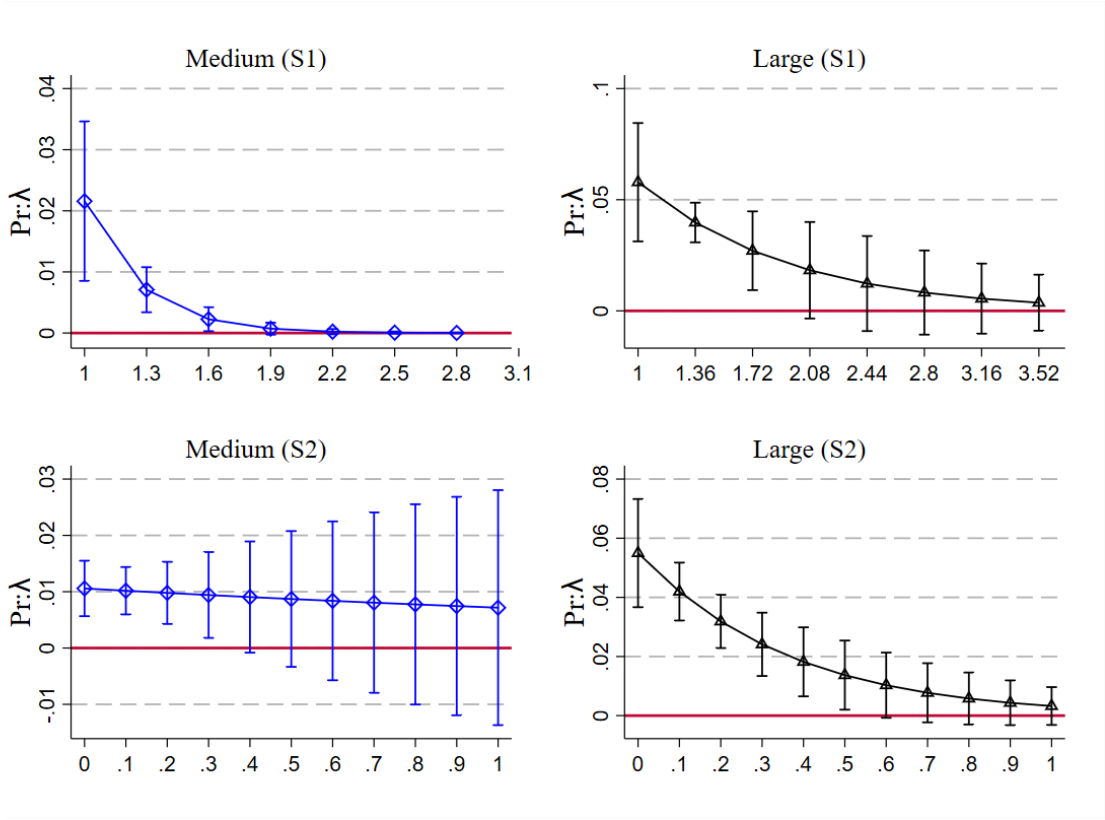

FIGURE 3. Predicted entry probability

interval. The slopes are negative over the plotted range but as in the case for mean marginal effects, not always significantly different from zero.

Robustness. While we have already included several robustness checks above, our main results are also robust to several alternative specifications that we did not discuss. For instance, our results hold up if we change the level of aggregation from quarterly to monthly observations, include or drop some of the control variables, or use different levels of ATC classifications. In the hazard models we also experimented with changing the definition of S2 variable to instead be the share of drugs introduced after three or seven years (instead of five), or to be equal to an indicator variable if the share of the newer drugs was greater than or equal to 0.5, 0.6 or 0.7 with fairly similar results with varying degree of statistical significance. We have not included these to keep the length of the paper manageable.

Discussion. Entry is deterred in medium-sized markets with higher values of $S1$, i.e., if the originator has more products, and patients are evenly spread across these additional variants. This is in line with entry deterrence motives in medium-sized markets, as well as our earlier result that the slow down in the launch rate after LoE among those that experience entry relative to those that do not, is largest in medium-sized markets. However, the alternative measure $S2$, associated with hopping, deters entry only in large markets, but puzzlingly is also positively correlated with entry in small markets.

Note that product hopping is relatively rare. Of the 430 originators in our sample, the variable $S2 > 0.5$ for only 19 cases spread as 6, 8 and 5 in small, medium and large markets. And of these cases, entry occurred 3, 2 and 1 times in small, medium, and large markets for a total of 6/19 (32%) cases. By contrast, in the 411 cases when $S2 \leq 0.5$, entry took place 131 times (131/411 = 32%), but this time, most of the entry was in large markets. See Table 6 below.

TABLE 6. Entry conditional on product hopping

	$S2 \leq 0.5$			$S2 > 0.5$		
	Originator	Entry	Percent	Originator	Entry	Percent
Small	138	9	7%	6	3	50%
Medium	135	35	26%	8	2	25%
Large	138	87	63%	5	1	20%
Total	411	131	32%	19	6	32%

We conjecture that positive and negative correlations with product hopping in small and large markets respectively is because product hopping takes place when an originator cannot maintain multiple product lines. Further, in small markets, they move to the newer variant of the drug without necessarily engaging in significant detailing effort. In turn, this creates an entry opportunity for others. In larger markets, it is more likely that the originator undertakes significant detailing efforts and convinces patients and their physicians that the newer variant is of superior quality. As mentioned earlier, product line extensions do not obtain additional data or marketing exclusivity by the drug approval authorities in the EU. Thus the LoE for these additional products is

the same as that for the original drug, but they may still be protected due to any additional patents. If so, this makes it difficult for competitors to launch generic versions of these newer drugs, and if the originator has successfully moved patients to the newer drug, then entry becomes difficult.

Thus while we cannot check detailing efforts, we additionally verified whether indeed entry into the newer versions launched by the originator is less common.

TABLE 7. Entry type by competitors

	Incumbent	Initial Entrant		All Entrants	
137 molecules					
Original formulations	148	117	(81%)	121	(74%)
New formulations	85	27	(19%)	42	(26%)
Total	233	144	(100%)	163	(100%)

Starting with the 137 (=131+6) originators that experienced entry, we classified their 233 $D1$ formulations launched before entry into two groups, 148 original, and 85 follow-on formulations. We then checked the formulation type of the drugs launched by competitors within the first year of their entry, and whether these drugs matched the originators' original formulation, or the originators' follow-on formulations. Results are summarized in Table 7, and show that in 81% of cases, entry was in the originators' original drug formulation and 19% it was for the newer follow-on formulations, i.e., competitors initially enter original formulations more often than new formulations by the originators. If we do not restrict to generic drugs that entered within the first year of any generic entry, the percentages change to 74% and 26% respectively.

Limitations. Our work has three main limitations. First, launch of a product line extension is a Europe-wide decision if not a worldwide one, rather than a decision based solely on entry prospects by a competitor in the UK market. We have relied on the fact that trade within Europe is easier, and the launch of an additional product by the originator, or an entry event by a competitor is a Europe-wide phenomena. But firms can choose to launch products in limited national markets. While the UK is an important market, and most

firms would launch here as well if they were entering or introducing a new variant in other parts of Europe (or other parts of the world), future work should attempt to overcome this difficulty. Second, we do not have access to physician detailing and other marketing data by originators. This is a choice variable, and some of our explanations rely on differences in marketing efforts by market size. We do not actually observe if, for instance, detailing is less in small markets compared to large markets around the time of product line extension. Ideally this variable should be included in the analysis. Finally, as already emphasized in earlier parts of the paper, allocation into subsamples with and without entry is not random, and hence results in the first part should be interpreted as being consistent with strategic deterrence rather than having fully identified the effect.

5. CONCLUSIONS

There is a long standing interest in entry deterrence in the theoretical literature, but there are relatively few empirical studies, primarily due to difficulties in identifying deterrence from other unilateral actions. Our paper adds to that sparse but growing empirical literature. Using data from UK pharmaceuticals, we test for changes in product line extension rate by originators over time, but before any entry takes place. The threat of entry changes after the loss of exclusivity, where some originators may find out that entry is imminent in the near future, and hence may change their rate of product line extensions. To that end, we compare the product launch rate before and after the loss of exclusivity period for all originators, by subsamples of originators that experience entry or not, and again the same by market size. An important assumption we make is that originators know the likelihood of entry based on their past sales/size of the market, therapy class, formulation and other characteristics of their original drug, and may also observe whether a competitor has filed for generic entry with the EMA or other national authorities.

We find that for firms that eventually experience entry, there is a sharp decline in their product launch rate after the loss of exclusivity, and before any

entry. This drop in rate is significant even when we compare it to originators with no entry who also experience the loss of exclusivity period for their drugs. The effect is larger in medium-size markets. We conclude from this that entry deterrence is a strong motive for launching product line extensions in the pharmaceuticals. We also find that product line extension is a successful strategy to deter entry in medium-sized markets if the originators can spread their patient base over the old and newer formulations. Since such a move requires expensive physician detailing, it probably makes it credible that the originator will not necessarily withdraw after an entry takes place. This does not appear to act as a deterrent in large or small markets. An alternative strategy, called 'product hopping' is to shift almost all the patients to the newer formulation prior to any generic entry. However, the evidence on its success is not very robust across different specifications but appears to deter entry in large markets.

REFERENCES

Ai, Churong and Edward C. Norton, "Interaction terms in logit and probit models," *Economics Letters*, July 2003, *80* (1), 123–129.

Appelt, Silvia, "Authorized generic entry prior to patent expiry: reassessing incentives for independent generic entry," *The Review of Economics and Statistics*, July 2015, *97* (3), 654–666.

Athey, Susan and Guido W. Imbens, "Identification and inference in nonlinear difference-in-differences models," *Econometrica*, March 2006, *74* (2), 431–497.

Bergman, Mats A and Niklas Rudholm, "The relative importance of actual and potential competition: empirical evidence from the pharmaceuticals market," *The Journal of Industrial Economics*, 2003, *51* (4), 455–467.

Berndt, Ernst R, Richard Mortimer, and Andrew Parece, "Do authorized generic drugs deter paragraph IV certifications? Recent evidence," 2007. Analysis Group Inc., Working Paper.

Bokhari, Farasat A. S. and Gary M. Fournier, "Entry in the ADHD drugs market: welfare impact of generics and me-toos," *Journal of Industrial Economics*, 2013, *61* (2), 340–393.

Bokhari, Farasat A.S., "Managed care and the adoption of hospital technology: The case of cardiac catheterization," *International Journal of Industrial Economics*, March 2009, *27* (2), 223–237.

Cameron, A. Colin and Pravin K. Trivedi, *Microeconometrics: methods and applications*, Cambridge: Cambridge University Press, 2005.

_____, **Jonah B. Gelback, and Douglas L Miller**, "Bootstrap-based improvements for inference with clustered errors," *The Review of Economics and Statistics*, August 2008, *90* (3), 414–427.

Carrier, Michael A and Steve D Shadowen, "Product hopping: a new framework," *Notre Dame Law Review*, 2016, *92*, 167.

Castanheira, Micael, Carmine Ornaghi, and Georges Siotis, "The unexpected consequences of generic entry," *Journal of Health Economics*, December 2019, *68* (102243), 1–22.

Caves, Richard E, Michael D Whinston, Mark A Hurwitz, Ariel Pakes, and Peter Temin, "Patent expiration, entry, and competition in

the US pharmaceutical industry," *Brookings Papers on Economic Activity. Microeconomics*, 1991, *1991*, 1–66.

Cook, Trevor, Catherine Doyle, and David Jabbari, *Pharmaceuticals, biotechnology and the law*, Springer, 1991.

Cookson, J Anthony, "Anticipated entry and entry deterrence: evidence from the American casino industry," *Management Science*, 2017.

Dafny, Leemore S, "Games hospitals play: entry deterrence in hospital procedure markets," *Journal of Economics & Management Strategy*, 2005, *14* (3), 513–542.

Danzon, Patricia M and Michael F Furukawa, "Cross-national evidence on generic pharmaceuticals: pharmacy vs. physician-driven markets," Technical Report, National Bureau of Economic Research 2011. NBER Working Paper 17226.

Dixit, Avinash, "A model of duopoly suggesting a theory of entry barriers," *The Bell Journal of Economics*, Spring 1979, *10* (1), 20–32.

_____, "The role of investment in entry-deterrence," *The Economic Journal*, 1980, *90* (357), 95–106.

EC, "Pharmaceutical Sector Inquiry, Final Report," European Commission (Directorates General, Competition) Brussels, Belgium July 2009.

Ellison, Glenn and Sara Fisher Ellison, "Strategic Entry Deterrence and the Behavior of Pharmaceutical Incumbents Prior to Patent Expiration," *American Economic Journal: Microeconomics*, 2011, *3* (1), 1–36.

FTC, "Authorized generic drugs: short-term effects and long-term impact," Technical Report, Federal Trade Commission 2011.

Goolsbee, Austan and Chad Syverson, "How do incumbents respond to the threat of entry? Evidence from the major airlines," *The Quarterly Journal of Economics*, 2008, *123* (4), 1611–1633.

Grabowski, Henry G and Margaret Kyle, "Generic competition and market exclusivity periods in pharmaceuticals," *Managerial and Decision Economics*, 2007, *28* (4-5), 491–502.

Hemphill, C Scott and Bhaven N Sampat, "Evergreening, patent challenges, and effective market life in pharmaceuticals," *Journal of Health Economics*, 2012, *31* (2), 327–339.

Hollis, Aidan, "The anti-competitive effects of brand-controlled 'pseudo-generics' in the Canadian pharmaceutical market," *Canadian Public Policy/Analyse de Politiques*, 2003, pp. 21–32.

Hollis, Aidian and Bryan A Liang, "An assessment of the effect of authorized generics on consumer prices," *Journal of Biolaw and Business*, 2007, *10* (1), 10.

Huskamp, Haiden A, Julie M Donohue, Catherine Koss, Ernst R Berndt, and Richard G Frank, "Generic entry, reformulations and promotion of SSRIs in the US," *Pharmacoeconomics*, 2008, *26* (7), 603–616.

Judd, Kenneth L, "Credible spatial preemption," *The RAND Journal of Economics*, 1985, pp. 153–166.

Kyle, Margaret, "Competition law, intellectual property, and the pharmaceutical sector," *Antitrust Law Journal*, 2016, *81* (1), 1–36.

Lieberman, Marvin B, "Excess capacity as a barrier to entry: an empirical appraisal," *The Journal of Industrial Economics*, 1987, pp. 607–627.

Morton, Fiona M Scott, "Entry decisions in the generic pharmaceutical industry," *The RAND Journal of Economics*, 1999, *30* (3), 421–440.

———, "Barriers to entry, brand advertising, and generic entry in the US pharmaceutical industry," *International Journal of Industrial Organization*, 2000, *18* (7), 1085–1104.

Puhani, Patrick A., "The treatment effect, the cross difference, and the interaction term in nonlinear "difference-in-differences" models," *Economic Letters*, April 2012, *115* (1), 85–87.

Reiffen, David and Michael R Ward, "Generic drug industry dynamics," *The Review of Economics and Statistics*, 2005, *87* (1), 37–49.

——— **and** ———, "'Branded Generics' as a strategy to limit cannibalization of pharmaceutical markets," *Managerial and Decision Economics*, 2007, *28* (4-5), 251–265.

Reinganum, Jennifer F., "The timing of innovation: Research, development, and diffusion," in R. Schmalensee and R.D. Willig, eds., *Handbook of Industrial Organization*, Vol. 1, North-Holland: Elsevier Science B.V., 1989.

Schmalensee, Richard, "Entry deterrence in the ready-to-eat breakfast cereal industry," *The Bell Journal of Economics*, 1978, pp. 305–327.

Seamans, Robert C., "Fighting City Hall: Entry deterrence and technology upgrades in cable tv markets," *Management Science*, March 2012, *58* (3), 461–475.

Smiley, Robert, "Empirical evidence on strategic entry deterrence," *International Journal of Industrial Organization*, 1988, *6* (2), 167–180.

Spence, Michael A, "Entry, capacity, investment and oligopolistic pricing," *The Bell Journal of Economics*, Authumn 1977, *8* (2), 534–544.

Weiman, David F and Richard C Levin, "Preying for monopoly? The case of Southern Bell Telephone Company, 1894-1912," *Journal of Political Economy*, 1994, *102* (1), 103–126.

Wilson, Robert, "Strategic models of entry deterrence," in Robert Aumann and Sergiu Hart, eds., *Handbook of Game Theory with Economic Applications*, Vol. 1, Amsterdam, North-Holland: Elsevier Science Publishers, 1992, chapter 10, pp. 305–329.

APPENDIX A. BACKGROUND AND DATA

A.1. **Legal protection in the EU and UK.** Market authorizations and patents protect originators from generic competition for a limited period. Since 1965, all pharmaceutical products require market authorization (MA) prior to launch, to ensure safety and effectiveness based on the Council Directive 65/65/EEC (and Medicines Act 1968 in the UK). In order to obtain MA, all applicants (originators) normally have to provide information from pre-clinical test and human clinical trials. However, given that replication of such data can be expensive, generic entrants are exempt from such requirements and can refer to the originators' data when applying for market authorization of their generic versions of the same molecule - as long as they can prove that their generic version is bio-equivalent to the originator.

Furthermore, the intellectual property rights, based on Article 39.3 of the TRIPS Agreement, protects the data supplied by the originators against 'unfair commercial use'. It implies that in some countries such data should not be used to authorize generic versions. Test and clinical trial data were protected as trade secrets until 1987 in the European Community, when the 87/21/EEC Directive and the 65/65/EEC Directive Amendment were introduced. The Amendment protects the originator's data for a pre-determined period, during which generic entrants cannot refer to the originator's data to receive market authorization. This data exclusivity period varies between 6 and 10 years across European countries. In the UK, data exclusivity is 10 years, (Cook et al., 1991). The period of data exclusivity starts from the date of first market authorisation registered anywhere in the European Community. Although this data exclusivity runs in parallel and irrespective of patent production, it often extends the monopoly position of the originator beyond the patent expiration, as 10 or more years can elapse between the filing of the primary patent and the launch date (Cook et al., 1991, Kyle, 2016). Moreover, data exclusivity only protects novel substances (molecules), while subsequent improvements to a drug, such as new therapeutic indications, dosage strength, or formulations, are not granted for an additional period of protection.[1]

Pharmaceutical companies can obtain licences either from EU member states' national authorities or the central European Medicines Agency (EMA), established in 1995. The difference between the centralized and decentralized licensing regime is that drugs can be sold in all member states if they are licensed from EMA, while they can only be sold in a specific country if they are licensed by a national agency. In addition, under the mutual recognition process, countries that receive an MA application do not have to start their own review, but can refer to the decision by the first agency that approved

[1]The Queen v The Licensing Authority established by the Medicines Act 1968 (acting by The Medicines Control Agency), ex parte Generics (UK) Ltd, The Wellcome Foundation Ltd and Glaxo Operations UK Ltd and Others. Case C-368/96. European Court Reports 1998 I-07967.

the drug (Kyle, 2016). In the UK, the Medicines and Healthcare products Regulatory Agency (MHRA) is responsible for medicine market authorization. The National Institute for Health and Care Excellence (NICE) assesses drugs' cost-effectiveness and issues purchase recommendations for the National Health Service (NHS) in England.

One notable change in the EU market authorization system is the harmonization of the '8+2+1' formula introduced in the 2001/83/EC Directive and amended by the follow-on Directive 2004/27/EC and Regulation 726/2004/EC. Market authorization applications made from November 2005 and onwards will follow this new rule. Under this new system, all member states of EU will have harmonized 8 years of data exclusivity from the first authorization date in the EU, followed by 2 years of 'market exclusivity'. This 10-year protection can be extended by one additional year if a 'significant new indication' or 'significant clinical benefit over existing therapies' is granted for this relevant medical product. Although generic entrants cannot market their versions during the data exclusivity and market exclusivity period (and possibly the additional year), they can make use of originators' pre-clinical and human clinical trial data after the first 8 years of data exclusivity. Comparing the old and new systems in the UK, the overall protection period for the originator remains 10 years. However, generics may apply for MA two years in advance under the new system. Although MA cannot be issued before the expiration of market exclusivity, the new system may reduce the gap between the expiration of market exclusivity and the launch of generic products, as they can start preparations for launch two years earlier (Kyle, 2016). Moreover, like the old system, the new system does not consider additional strengths, formulations, administration routes, presentations, and variations and extensions as new sources for another market authorization other than the initial one.

Running in parallel with the market authorization system is patent protection. In the EU, patent life normally lasts 20 years since filing, during which the originator has an exclusive right to prevent generics from marketing their products. However, the effective patent protection period for drugs marketed after MA is generally shorter, as it may take a long time for firms to obtain enough data for MA. In order to compensate for the loss of patent protection, and to protect innovation in the pharmaceutical market, the Supplementary Protection Certificate (SPC) was introduced in 1992 in the EU.[2] SPC offers the same protection as the basic patent (*sui generis*) and it extends the patent life of medicines up to 5 years beyond patent expiration or 15 years since market authorization, whichever is less. Moreover, as noted in Kyle (2016), EU regulators tend to prevent the linkage between patent and exclusivity. This

[2]Council Regulation (EEC) No 1768/92 of 18 June 1992 concerning the creation of a supplementary protection certificate for medicinal products entered into force in 1993. It has been replaced by Regulation (EC) No 469/2009 of the European Parliament and of the Council of 6 May 2009 concerning the supplementary protection certificate for medicinal products.

means that regulators may review generics even if the originator may still have some valid patents. Since investing around a secondary patent is easier than a primary product patent, generics may enter earlier.

One side effect of this de-linkage between patent and market authorization in the UK/EU is that it is not easy to trace when the relevant patents associated with a given drug expire (in the US, relevant patent information for a drug is available from the Food and Drug Administration's 'Orange Book' data base). Although both regulatory data protection and patent aim at protecting the originator's innovation, the interaction between MA and patent (and SPC) is complex, as distinct laws govern them. One medical product can have several patents, while only one MA will be granted. Therefore, how to implement patent protection to the entire product depends on specific conditions, which vary across different cases. Since we do not obtain patent information for our products, we rely on MA information and firms' launch date as recorded in IMS to determine when a molecule (or market) is open for generics to enter.

A.2. **Data construction.** As mentioned in the main text, our data source is sales data from the British Pharmaceutical Index (BPI) series by Intercontinental Marketing Services (IMS) for the period 1996:Q3-2016:Q3. This appendix describes various data cleaning steps.

Zero sales and counts of drugs. For each item, IMS only reports values of positive sales (quantity and value) of shipments from wholesalers to retailers at monthly intervals. If a particular item was not shipped within a given month, it would not appear as a line item in the data. Since we count how many products an originator has in each period based on it being present in the data, zero sales within a given month can lead to erroneously under-counting products lines. However, products may be formally or effectively withdrawn by setting sales to zero. To resolve this problem, we (1) aggregate the data to quarters and (2) check if an individual item ever has a positive sale in any of the future periods. If there is a positive sale in the future, the item is not considered withdrawn from the market. Instead we set its sale value equal to zero for all intermediate periods until it makes a positive sale, and the originator's product count is not decreased. If however there are no future sales, then the product is considered withdrawn from the market and the count of originator's product lines are adjusted accordingly from that point onwards.

Mergers and acquisitions. Merger and acquisition activities during the period are handled by IMS by retroactively reassigning the sales and associated products to the end-of-period corporation that owns them as if they owned it for the entire period. Thus, if two firms merged during the observed period, they appear to be a single firm from the start. Generally, a similar rule applies to product line acquisitions. However, here we found some inconsistencies in the data as we could sometimes observe change of ownership for a given product. Further, the full 1996-2016 data series was obtained from IMS in three parts, and so the above rule was applied to each datum cut separately which lead to further inaccuracies about ownership. We corrected for these by tracking the name of a manufacturer listed against a propriety name for each branded drug that we used in our sample (described below). If the name of the manufacturer changed, we assigned it for the entire period to the last owner in the series. In some cases, IMS makes this task easier by appending an abbreviation of manufacturer's name to a proprietary product name when the ownership changes. Clearly the method does not apply to generics, as they are listed only by non-proprietary names and the name of a generic manufacturer is typically not listed in the data. However, this should not matter since the analysis is centered around originators and whether they experience any generic entry or not. (Identity of the generic firm is not important in our analysis).

Formulations. We used the three digit New Form Code (NFC3) system introduced by EphMRA (version 2016) to distinguish between formulations and to

construct our measure $D1_j$, i.e., count of drugs by the originator that differ by the value of this code. It takes upto 78 different values in our sample. We also constructed a simplified version from these codes, based on first or second digit, and is given in Table A-1). The latter simplified code is used to construct dummy variables for formulation which are then used in our regression analysis. These follow aggregation rule similar to that used in the literature (see Scott Morton, 1999).

TABLE A-1. Formulations

Formulation	Description	NFC Code combinations
Solid	Oral solid (ordinary or long acting) as tablet, coated tablet, or a capsule	NFC1 $\in \{A, B\}$ and NFC2 $\in \{A, B, C\}$
Liquid	Liquid & ressured aerosols	NFC2 $\in \{G, H\}$
Injection	Ampoules, pre-filled syringe, vials, infusions and cartridges/pens	NFC2 $\in \{M, N, P, Q, R\}$
Ointment	Ointments, creams and gels and sols	NFC2 $\in \{S, T, V\}$
Other	All others (i.e., powers /granules, suppositories, medicated dressings and other special forms) or if the originator had multiple original formulations	NFC2 $\in \{E, L, W, Y\}$

Notes. Other NFC codes do not appear in the final version of the data sets we used so are not listed here.

Data Samples. Several drugs were eliminated from the analysis. These included multi-molecule Drugs, such as vitamins and vaccines (J07), as well as those with single 1-digit ATC categories of hospital solutions (K), diagnostic agents (T) and various (V). For such drugs it is not always clear what constitutes a market. Additionally, we focus on Prescription only Medicines (PoM), which count for about 75% of single molecule medicines. Over-the-Counter (OTC) drugs are excluded as they could also be sold in supermarkets, whose sales information is not included in IMS data. From this set, we restricted the analysis to only those originator (branded) drugs that lost exclusivity between 1996 and 2016 (measured as the tenth year from the UK launch date noted in the IMS data). This criteria initially identified 508 originators. However in some of these cases, a generic version of the drug pre-dated entry into the UK by the originator in the same ATC4-molecule class. This may be due to a merger or product acquisition, re-registration with MHRA in the UK, or other errors in the date. We eliminated these cases. This lead to 450 originators. Of these, an additional 11 originators were eliminated because competitors entered before our data series begins in 1996 and two more were discarded because they show zero sales until a competitor enters giving us a sample size of

437 originators. Finally seven more were eliminated as they had less than five total observations in the 20 year data span. This gave us our final full sample of 430 originators of which 137 experienced entry. This data is described in the main text (descriptive statistics are given in Table 2) and used in hazard analysis. A smaller sub-sample was constructed for studying originators' product launches before and after the end of exclusivity, where the period for the end of exclusivity was restricted to be between 2001-2011. This gave 263 originators of which 92 experienced entry by the end of our series. The descriptive statistics are given below in Table A-2 and the mean values of these variables by entry status are given in Table A-3 (for a single cross-section two years before LoE).

TABLE A-2. Originator's characteristics: sub sample (263 originators)

Variable	Description	Mean	Std. Dev.			Min	Max
			overall	between	within		
$D1$	Count based on formulations	1.34	0.67	0.59	0.31	1	5
$D2$	Count based on pack variations	3.21	3.56	3.31	1.38	1	37
$S1$	1/HHI from shares of D1	1.12	0.28	0.23	0.15	1	3.61
$S2$	Share of D1 launched after five years	0.03	0.15	0.12	0.10	0	1
Sales (log)	Sales by originator	10.69	4.46	4.16	2.20	0	18.27
†Monopoly	Originator monopolist in other classes	0.91	0.28	0.27	0.14	0	1
†Nearby	Other monopolists in ATC3 class	0.80	0.40	0.36	0.18	0	1
†Chronic	Chronic disease drug	0.71	0.45			0	1
†SPC	Originator enters after 1993	0.87	0.34			0	1
†1Form	Single original formulation	0.93	0.25			0	1
†Solid	Tablets, capsules, extend release, etc.	0.45	0.50			0	1
†Liquid	Liquids & aerosols	0.09	0.29			0	1
†Injection	Ampules, vials, pre-filled syringes, etc.	0.26	0.44			0	1
†Ointment	Ointments, creams, gels & sols	0.08	0.27			0	1
†Other	All others & multiple formulations	0.13	0.33			0	1

Notes. Summary statistics from unbalanced panel of 263 originators over 40 quarters with 13,559 observations. For time invariant variables, there is no *within* standard deviation and overall standard deviation is the same as *between*. For the larger sample with 430 originators, see Table 2 in the main text.†1/0 Dummy variable, 1 if true.

TABLE A-3. Summary statistics by entry status (Cross-section)

		Entry		No entry		Diff in	P-value
		Mean	sd	Mean	sd	Mean	(2sided)
$D1$	Count based on formulations	1.51	0.85	1.12	0.69	0.39	0.00
$D2$	Count based on pack variations	5.45	4.91	2.25	3.03	3.20	0.00
$S1$	1/HHI from shares of D1	1.13	0.27	0.97	0.43	0.16	0.00
$S2$	Share of D1 launched after five years	0.05	0.19	0.01	0.09	0.04	0.08
Sales (log)	Sales by originator	14.04	2.71	9.13	4.36	4.92	0.00
†Monopoly	Originator monopolist in other classes	0.86	0.35	0.88	0.33	-0.02	0.68
†Nearby	Other monopolists in ATC3 class	0.78	0.41	0.74	0.44	0.05	0.41
†Chronic	Chronic disease drug	0.85	0.36	0.64	0.48	0.21	0.00
†SPC	Originator enters after 1993	0.84	0.37	0.87	0.34	-0.03	0.46
†1Form	Single original formulation	0.95	0.23	0.92	0.27	0.02	0.49
†Solid	Tablets, capsules, extend release, etc.	0.67	0.47	0.35	0.48	0.33	0.00
†Liquid	Liquids & aerosols	0.03	0.18	0.11	0.32	-0.08	0.01
†Injection	Ampules, vials, pre-filled syringes, etc.	0.15	0.36	0.31	0.46	-0.16	0.00
†Ointment	Ointments, creams, gels & sols	0.02	0.15	0.09	0.29	-0.07	0.01
†Other	All others & multiple formulations	0.12	0.33	0.14	0.35	-0.02	0.63
Originators		92		171			

Notes. Summary statistics from a cross-sectional of 263 originators. It has 263 observations.
†1/0 Dummy variable, 1 if true.

TABLE A-4. Summary statistics by entry status in medium-sized market (cross-section)

		Entry		No entry		Diff in	P-value
		Mean	sd	Mean	sd	Mean	(2sided)
$D1$	Count based on formulations	1.45	0.73	1.30	0.52	0.15	0.47
$D2$	Count based on pack variations	3.78	1.93	2.65	2.57	1.13	0.07
$S1$	1/HHI from shares of D1	1.12	0.23	1.13	0.26	-0.01	0.90
$S2$	Share of D1 launched after five years	0.11	0.30	0.01	0.06	0.10	0.22
Sales (log)	Sales by originator	12.26	0.79	11.28	1.40	0.98	0.00
†Monopoly	Originator monopolist in other classes	0.93	0.26	0.89	0.32	0.05	0.54
†Nearby	Other monopolists in ATC3 class	0.87	0.35	0.75	0.43	0.11	0.30
†Chronic	Chronic disease drug	0.67	0.49	0.67	0.47	-0.01	0.97
†SPC	Originator enters after 1993	0.80	0.41	0.87	0.34	-0.07	0.56
†1Form	Single original formulation	0.93	0.26	0.93	0.25	0.00	0.99
†Solid	Tablets, capsules, extend release, etc.	0.53	0.52	0.39	0.49	0.14	0.35
†*Liquid	Liquids & aerosols	0.00	-	0.16	0.37	-0.16	0.10
†Injection	Ampules, vials, pre-filled syringes, etc.	0.40	0.51	0.20	0.40	0.20	0.17
†*Ointment	Ointments, creams, gels & sols	0.00	-	0.15	0.36	-0.15	0.12
†Other	All others & multiple formulations	0.07	0.26	0.10	0.30	-0.03	0.68
Originators		15		61			

Notes. Summary statistics from a cross-sectional of 263 originators. It has 263 observations.†1/0 Dummy variable, 1 if true.

*We assume equal variances of the groups for these variables in t-test, as variance of the group of originators that experienced entry is not available.

TABLE A-5. Balancing test: All markets

Overall	Treated		Control		Diff in	p-value	Std. diff.		Var. ratio	
	mean	sd	mean	sd	mean	(2sided)	Raw	Matched	Raw	Matched
Sales (log)	12.92	3.33	11.57	3.61	1.36	0.12	0.93	-0.07	0.60	0.86
†Monopoly	0.95	0.21	0.96	0.19	0.00	0.92	0.14	-0.11	0.63	1.93
†Nearby	0.88	0.32	0.88	0.31	0.00	0.98	0.30	-0.16	0.60	1.66
†Chronic	0.77	0.43	0.70	0.47	0.06	0.57	0.26	0.16	0.80	0.85
†SPC	0.84	0.37	0.89	0.32	-0.05	0.54	-0.09	-0.21	1.22	1.62
†1Form	0.91	0.29	0.96	0.19	-0.06	0.34	-0.09	-0.18	1.34	1.90
†Solid	0.51	0.51	0.52	0.51	-0.01	0.96	0.34	0.19	1.13	1.03
†Liquid	0.07	0.26	0.04	0.19	0.03	0.55	-0.12	0.22	0.72	2.86
†Injection	0.26	0.44	0.33	0.48	-0.08	0.50	-0.12	-0.44	0.91	0.77
†Ointment	0.02	0.15	0.07	0.27	-0.05	0.37	-0.35	-0.13	0.24	0.51
†Other	0.14	0.35	0.04	0.19	0.10	0.12	0.01	0.32	1.03	2.71
Originators	43		27							

Notes. Treated: markets with entry. Control: markets without entry. †1/0 Dummy variable, 1 if true.

TABLE A-6. Balancing test: Small markets

Small	Treated		Control		Diff in	p-value	Std. diff.		Var. ratio	
	mean	sd	mean	sd	mean	(2sided)	Raw	Matched	Raw	Matched
Sales (log)	5.94	3.54	6.04	3.64	-0.10	0.97	-0.86	-0.03	0.68	0.95
†Monopoly	1.00	0.00	0.80	0.45	0.20	0.35	0.42	0.63	0.00	0.00
†Nearby	0.95	0.11	0.88	0.28	0.08	0.59	0.60	0.35	0.07	0.16
†Chronic	0.80	0.45	0.40	0.55	0.40	0.24	0.32	0.80	0.87	0.67
†SPC	0.80	0.45	0.80	0.45	0.00	1.00	-0.18	0.00	1.75	1.00
†1Form	1.00	0.00	1.00	0.00	0.00	.	0.38	.	0.00	.
†Solid	0.00	0.00	0.00	0.00	0.00	.	-1.02	.	0.00	.
†Liquid	0.00	0.00	0.00	0.00	0.00	.	-0.48	.	0.00	.
†Injection	0.80	0.45	0.80	0.45	0.00	1.00	1.08	0.00	0.93	1.00
†Ointment	0.00	0.00	0.00	0.00	0.00	.	-0.49	.	0.00	.
†Other	0.20	0.45	0.20	0.45	0.00	1.00	0.16	0.00	1.68	1.00
Originators	5		5							

Notes. Treated: markets with entry. Control: markets without entry. †1/0 Dummy variable, 1 if true.

TABLE A-7. Balancing test: Medium markets

Medium	Treated		Control		Diff in	p-value	Std. diff.		Var. ratio	
	mean	sd	mean	sd	mean	(2sided)	Raw	Matched	Raw	Matched
Sales (log)	11.99	1.67	12.43	1.66	-0.44	0.49	0.81	-0.27	0.15	1.18
†Monopoly	0.93	0.26	1.00	0.00	-0.07	0.33	0.05	-0.37	0.92	.
†Nearby	0.87	0.35	0.84	0.37	0.03	0.83	0.26	0.02	0.71	1.01
†Chronic	0.67	0.49	0.62	0.51	0.05	0.79	0.03	0.13	1.04	0.93
†SPC	0.80	0.41	0.85	0.38	-0.05	0.76	-0.18	0.00	1.50	1.00
†1Form	0.93	0.26	0.92	0.28	0.01	0.92	0.01	0.00	1.04	1.00
†Solid	0.53	0.52	0.31	0.48	0.23	0.24	0.38	0.40	1.18	1.12
†Liquid	0.00	0.00	0.23	0.44	-0.23	.	-0.48	-0.68	0.00	.
†Injection	0.40	0.51	0.23	0.44	0.17	0.35	0.19	0.28	1.20	1.23
†Ointment	0.00	0.00	0.08	0.28	-0.08	.	-0.49	-0.37	0.00	.
†Other	0.07	0.26	0.15	0.38	-0.09	0.49	-0.23	-0.22	0.56	0.54
Originators	15		13							

Notes. Treated: markets with entry. Control: markets without entry. †1/0 Dummy variable, 1 if true.

TABLE A-8. Balancing test: Large markets

Large	Treated		Control		Diff in	p-value	Std. diff.		Var. ratio	
	mean	sd	mean	sd	mean	(2sided)	Raw	Matched	Raw	Matched
Sales (log)	14.96	0.77	14.59	1.16	0.37	0.34	1.81	0.01	0.03	0.58
†Monopoly	0.95	0.22	1.00	0.00	-0.05	.	0.13	-0.31	0.66	.
†Nearby	0.90	0.30	1.00	0.00	-0.10	.	0.38	-0.45	0.52	.
†Chronic	0.86	0.36	0.55	0.52	0.31	0.08	0.49	1.23	0.56	0.55
†SPC	0.81	0.40	0.73	0.47	0.08	0.62	-0.16	0.11	1.42	0.85
†1Form	0.90	0.30	0.91	0.30	0.00	0.97	-0.10	0.14	1.41	0.70
†Solid	0.67	0.48	0.45	0.52	0.21	0.27	0.68	0.19	1.03	0.91
†Liquid	0.14	0.36	0.18	0.40	-0.04	0.79	0.12	0.14	1.39	1.42
†Injection	0.05	0.22	0.27	0.47	-0.23	0.14	-0.72	-0.44	0.22	0.29
†Ointment	0.00	0.00	0.00	0.00	0.00	.	-0.49	.	0.00	.
†Other	0.14	0.36	0.09	0.30	0.05	0.67	0.02	0.00	1.08	1.00
Originators	21		11							

Notes. Treated: markets with entry. Control: markets without entry. †1/0 Dummy variable, 1 if true.

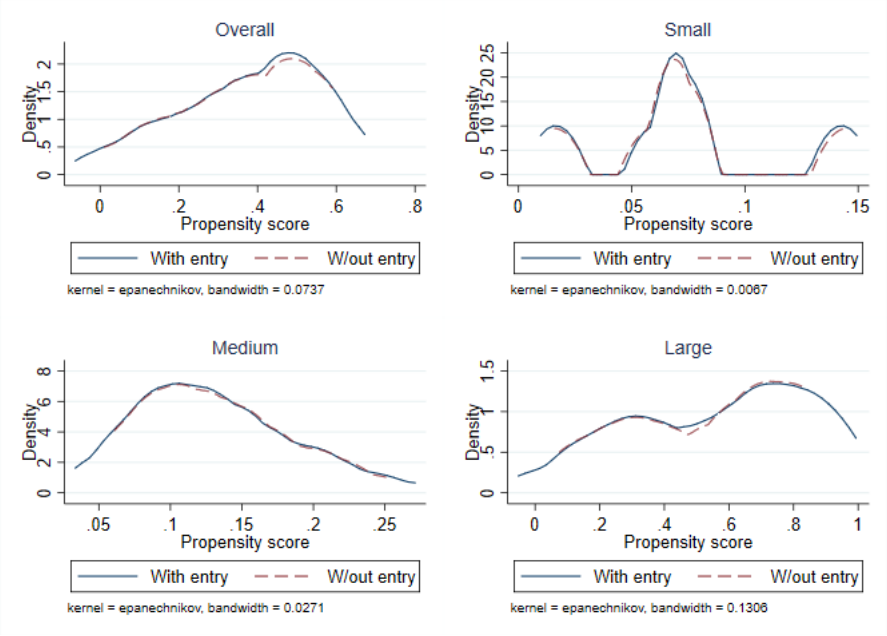

FIGURE A-1. Predicted propensity scores of the matched sample

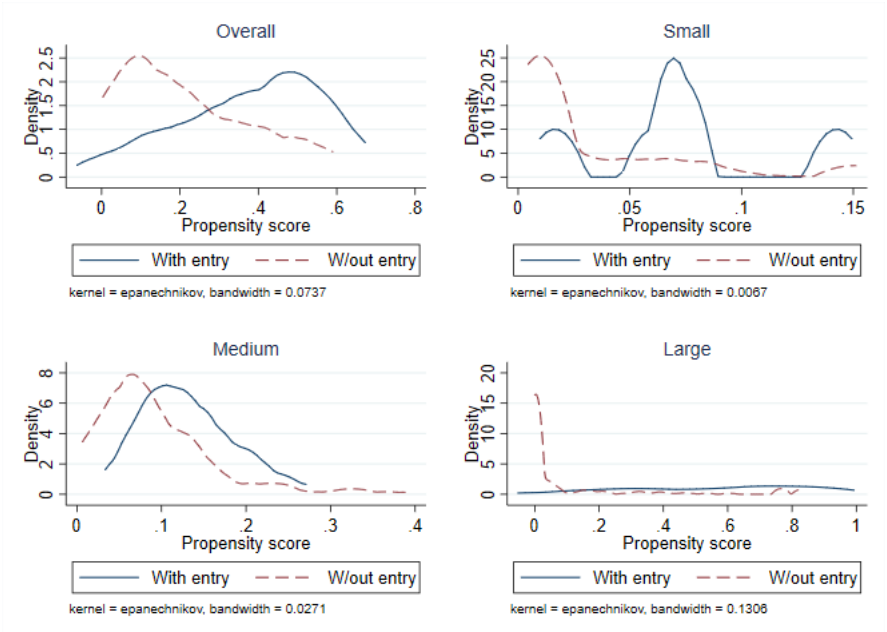

FIGURE A-2. Predicted propensity scores of the raw sample

TABLE A-9. Product launch rate, matched sample

	Overall		Small		Medium		Large	
	With	W/out	With	W/out	With	W/out	With	W/out
D1	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T	0.005	0.008	-0.000	0.001	0.009	-0.005	0.007	0.001
	(0.002)^a	(0.001)^a	(0.002)	(0.006)	(0.003)^a	(0.002)^b	(0.003)^b	(0.001)
	[0.003]	[0.004]^b	[0.001]	[0.005]	[0.007]	[0.004]	[0.004]^c	[0.003]
LoE	-0.013	-0.080	-0.039	-0.003	-0.129	0.115	0.090^c	-0.031
	(0.031)	(0.026)^a	(0.025)	(0.086)	(0.041)^a	(0.044)^a	(0.048)	(0.025)
	[0.023]	[0.022]^a	[0.046]	[0.088]	[0.077]	[0.044]^b	[0.072]	[0.015]^b
LoE×T	-0.008	0.019	0.001	0.008	-0.012	-0.001	0.002	0.013
	(0.003)^a	(0.002)^a	(0.002)	(0.008)	(0.004)^a	(0.004)	(0.005)	(0.003)^a
	[0.006]	[0.009]^b	[0.001]	[0.009]	[0.009]	[0.006]	[0.008]	[0.010]
r^2	0.649	0.936	0.896	0.375	0.863	0.816	0.852	0.957
†$\chi^2(1)$	6.59		0.73		18.47		0.26	
p-value	0.010		0.393		0.000		0.610	
D2	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T	0.027	-0.004	-0.001	-0.001	0.028	-0.053	0.084	0.037
	(0.008)^a	(0.003)	(0.015)	(0.007)	(0.006)^a	(0.013)^a	(0.011)^a	(0.004)^a
	[0.013]^b	[0.006]	[0.023]	[0.008]	[0.013]^b	[0.038]	[0.022]^a	[0.013]^a
LoE	0.240	-0.089	0.195	-0.071	0.039	0.070	0.076	-0.473
	(0.151)	(0.052)^c	(0.168)	(0.102)	(0.152)	(0.153)	(0.145)	(0.073)^a
	[0.139]^c	[0.078]	[0.207]	[0.117]	[0.249]	[0.300]	[0.235]	[0.134]^a
LoE×T	-0.027	0.027	0.008	0.008	-0.048	0.068	-0.065	-0.032
	(0.014)^c	(0.005)^a	(0.016)	(0.009)	(0.011)^a	(0.016)^a	(0.017)^a	(0.006)^a
	[0.028]	[0.012]^b	[0.025]	[0.010]	[0.033]	[0.038]^c	[0.038]^c	[0.022]
r^2	0.722	0.948	0.646	0.874	0.869	0.919	0.911	0.938
†$\chi^2(1)$	24.89		6.78		51.88		8.74	
p-value	0.000		0.009		0.000		0.003	
Observations	1,565	1,682	159	185	550	593	771	804
Unique originators	43	27	5	5	15	13	21	11

Robust standard errors in parentheses followed by cluster standard errors in brackets. Superscripts a, b, c indicate significance at 1%, 5% and 10% levels, respectively. Full set of coefficients are available from authors upon request..

†Test of equality of the interaction term across samples restricted to originator with and without eventual entry.

Appendix B. Additional Results

B.1. **Regression Coefficients for count of Products.** This appendix provides a full set of regressions coefficients for the selected coefficients reported in section 4.

Table B-1 below corresponds to the selected coefficients shown in Table 3.

TABLE B-1. Product launch rate (full version)

	D1 and D2 by samples A,B,C						D1 and D2 by subsamples of C			
	D1			D2			D1		D2	
	(1) A	(2) B	(3) C	(4) A	(5) B	(6) C	C(7) With	C(8) W/out	C(9) With	C(10) W/out
T	0.006	0.006	0.003	0.039	0.039	0.028	0.009	0.001	0.088	-0.005
	$(0.001)^a$	$(0.001)^a$	$(0.001)^b$	$(0.003)^a$	$(0.003)^a$	$(0.008)^a$	$(0.003)^a$	(0.001)	$(0.013)^a$	(0.008)
	$[0.001]^a$	$[0.002]^a$	$[0.002]^c$	$[0.007]^a$	$[0.007]^a$	$[0.010]^a$	$[0.003]^a$	[0.002]	$[0.018]^a$	[0.007]
LoE	0.017	-0.001	0.018	-0.030	-0.001	0.182	-0.004	0.021	0.315	0.070
	(0.018)	(0.019)	(0.025)	(0.103)	(0.104)	(0.139)	(0.052)	(0.022)	(0.228)	(0.136)
	[0.026]	[0.026]	[0.018]	[0.134]	[0.126]	$[0.091]^b$	[0.038]	[0.018]	[0.193]	[0.079]
LoE×T	-0.005	-0.005	-0.001	-0.056	-0.054	-0.053	-0.008	0.003	-0.098	0.001
	$(0.001)^a$	$(0.001)^a$	(0.002)	$(0.004)^a$	$(0.004)^a$	$(0.011)^a$	$(0.005)^c$	$(0.002)^c$	$(0.020)^a$	(0.012)
	$[0.002]^c$	$[0.002]^c$	[0.003]	$[0.012]^a$	$[0.012]^a$	$[0.015]^a$	[0.007]	[0.003]	$[0.029]^a$	[0.012]
Sales (log)	0.058	0.060	0.062	0.340	0.350	0.381	0.042	0.050	0.359	0.228
	$(0.001)^a$	$(0.001)^a$	$(0.002)^a$	$(0.008)^a$	$(0.009)^a$	$(0.011)^a$	$(0.006)^a$	$(0.002)^a$	$(0.041)^a$	$(0.011)^a$

This table is an extended version of regression coefficients shown in Table 3 in the main text. Robust standard errors in parentheses followed by cluster standard errors in brackets. Superscripts a, b, c indicate significance at 1%, 5% and 10%, respectively. All regressions include other controls and ATC2 dummies. Sample (A) is all initial 263 originators, (B) restricts to 212 originators with sales observed before/after LoE, and (C) is same as (B) but with observations restricted to within 5 years of the LoE.

continued

TABLE B-1. Product launch rate (full version)

	D1 and D2 by samples A,B,C						D1 and D2 by subsamples of C			
	D1			D2			D1		D2	
	(1) A	(2) B	(3) C	(4) A	(5) B	(6) C	C(7) With	C(8) W/out	C(9) With	C(10) W/out
	[0.008]^a	[0.008]^a	[0.009]^a	[0.052]^a	[0.054]^a	[0.061]^a	[0.027]	[0.007]^a	[0.167]^b	[0.059]^a
†Solid	-0.204	-0.235	-0.205	-1.170	-1.331	-1.387	-0.161	-0.516	-0.034	-3.092
	(0.033)^a	(0.035)^a	(0.048)^a	(0.257)^a	(0.285)^a	(0.396)^a	(0.022)^a	(0.062)^a	(0.228)	(0.512)^a
	[0.232]	[0.257]	[0.287]	[1.918]	[2.137]	[2.457]	[0.087]^c	[0.373]	[1.055]	[3.190]
†Liquid	0.011	-0.007	-0.098	-2.370	-2.579	-3.481	-0.615	-0.118	-6.081	-4.120
	(0.044)	(0.049)	(0.065)	(0.317)^a	(0.362)^a	(0.518)^a	(0.097)^a	(0.076)	(0.557)^a	(0.617)^a
	[0.301]	[0.341]	[0.380]	[2.317]	[2.660]	[3.163]	[0.481]	[0.450]	[2.886]^b	[3.784]
†Injection	0.250	0.233	0.225	-0.195	-0.231	-0.463	-0.435	0.050	-1.417	-1.512
	(0.031)^a	(0.034)^a	(0.046)^a	(0.230)	(0.256)	(0.357)	(0.062)^a	(0.057)	(0.375)^a	(0.445)^a
	[0.210]	[0.233]	[0.262]	[1.690]	[1.893]	[2.186]	[0.319]	[0.330]	[1.876]	[2.748]
†Ointment	0.113	0.167	0.174	-0.244	0.252	0.186	0.887	-0.205	0.852	-2.481
	(0.036)^a	(0.044)^a	(0.059)^a	(0.255)	(0.330)	(0.462)	(0.103)^a	(0.068)^a	(0.685)	(0.538)^a

This table is an extended version of regression coefficients shown in Table 3 in the main text. Robust standard errors in parentheses followed by cluster standard errors in brackets. Superscripts a, b, c indicate significance at 1%, 5% and 10%, respectively. All regressions include other controls and ATC2 dummies. Sample (A) is all initial 263 originators, (B) restricts to 212 originators with sales observed before/after LoE, and (C) is same as (B) but with observations restricted to within 5 years of the LoE.

continued

TABLE B-1. Product launch rate (full version)

	D1 and D2 by samples A,B,C						D1 and D2 by subsamples of C			
	D1			D2			D1		D2	
	(1) A	(2) B	(3) C	(4) A	(5) B	(6) C	C(7) With	C(8) W/out	C(9) With	C(10) W/out
	[0.233]	[0.295]	[0.335]	[1.803]	[2.353]	[2.758]	[0.519]^{c}	[0.392]	[3.381]	[3.291]
†1Form	-0.499	-0.490	-0.498	-1.020	-0.978	-1.297	-0.200	-0.298	-5.590	1.923
	(0.039)^{a}	(0.042)^{a}	(0.057)^{a}	(0.281)^{a}	(0.312)^{a}	(0.431)^{a}	(0.123)	(0.065)^{a}	(0.545)^{a}	(0.497)^{a}
	[0.254]^{c}	[0.280]^{c}	[0.314]	[1.972]	[2.218]	[2.583]	[0.645]	[0.371]	[2.626]^{b}	[3.102]
†Chronic	0.237	0.251	0.275	-0.046	0.111	0.178	-0.754	0.378	-3.136	0.818
	(0.016)^{a}	(0.017)^{a}	(0.022)^{a}	(0.072)	(0.077)	(0.103)^{c}	(0.140)^{a}	(0.021)^{a}	(0.552)^{a}	(0.111)^{a}
	[0.095]^{b}	[0.101]^{b}	[0.120]^{b}	[0.492]	[0.526]	[0.608]	[0.800]	[0.110]^{a}	[3.249]	[0.670]
†SPC	-0.219	-0.241	-0.307	-1.105	-1.347	-1.612	-0.404	-0.160	-1.930	-0.525
	(0.021)^{a}	(0.022)^{a}	(0.026)^{a}	(0.113)^{a}	(0.118)^{a}	(0.144)^{a}	(0.032)^{a}	(0.026)^{a}	(0.288)^{a}	(0.133)^{a}
	[0.126]^{c}	[0.134]^{c}	[0.150]^{b}	[0.677]	[0.707]^{c}	[0.797]^{b}	[0.156]^{b}	[0.145]	[1.340]	[0.814]
†Monopoly	-0.105	-0.102	-0.111	-0.247	-0.250	0.335	-0.348	-0.095	-0.799	0.561

This table is an extended version of regression coefficients shown in Table 3 in the main text. Robust standard errors in parentheses followed by cluster standard errors in brackets. Superscripts a, b, c indicate significance at 1%, 5% and 10%, respectively. All regressions include other controls and ATC2 dummies. Sample (A) is all initial 263 originators, (B) restricts to 212 originators with sales observed before/after LoE, and (C) is same as (B) but with observations restricted to within 5 years of the LoE.

continued

TABLE B-1. Product launch rate (full version)

	D1 and D2 by samples A,B,C						D1 and D2 by subsamples of C			
	D1			D2			D1		D2	
	(1) A	(2) B	(3) C	(4) A	(5) B	(6) C	C(7) With	C(8) W/out	C(9) With	C(10) W/out
	(0.013)^a	(0.015)^a	(0.018)^a	(0.074)^a	(0.088)^a	(0.125)^a	(0.046)^a	(0.018)^a	(0.156)^a	(0.156)^a
	[0.052]^b	[0.060]^c	[0.075]	[0.382]	[0.473]	[0.687]	[0.085]^a	[0.090]	[0.489]	[0.862]
†Nearby	0.078	0.082	0.061	0.406	0.411	0.339	-0.166	0.122	-1.436	1.075
	(0.017)^a	(0.019)^a	(0.028)^b	(0.102)^a	(0.112)^a	(0.165)^b	(0.046)^a	(0.022)^a	(0.245)^a	(0.167)^a
	[0.087]	[0.095]	[0.137]	[0.462]	[0.506]	[0.709]	[0.218]	[0.078]	[1.085]	[0.559]^c
Constant	1.330	1.298	1.392	3.568	2.646	2.880	2.973	1.227	15.140	-0.516
	(0.049)^a	(0.055)^a	(0.071)^a	(0.257)^a	(0.289)^a	(0.392)^a	(0.234)^a	(0.051)^a	(1.172)^a	(0.313)^c
	[0.264]^a	[0.272]^a	[0.355]^a	[1.358]^a	[1.442]^c	[1.992]	[1.224]^b	[0.228]^a	[6.118]^b	[1.607]
Observations	13,559	12,560	8,052	13,559	12,560	8,052	2,325	5,727	2,325	5,727
R^2	0.410	0.421	0.425	0.384	0.396	0.423	0.551	0.533	0.666	0.383
Originators	263	212	212	263	212	212	66	146	66	146

This table is an extended version of regression coefficients shown in Table 3 in the main text. Robust standard errors in parentheses followed by cluster standard errors in brackets. Superscripts a, b, c indicate significance at 1%, 5% and 10%, respectively. All regressions include other controls and ATC2 dummies. Sample (A) is all initial 263 originators, (B) restricts to 212 originators with sales observed before/after LoE, and (C) is same as (B) but with observations restricted to within 5 years of the LoE.

end

Table B-2 below corresponds to the selected coefficients shown in the *upper block* of Table 4 in the main text.

TABLE B-2. Product launch rate by market size (for D1)

	Small		Medium		Large	
	With	W/out	With	W/out	With	W/out
	(1)	(2)	(3)	(4)	(5)	(6)
T	0.000	-0.002	0.013	-0.002	0.002	0.008
	(0.003)	(0.001)	(0.003)^a	(0.002)	(0.003)	(0.003)^a
	[0.001]	[0.002]	[0.010]	[0.003]	[0.004]	[0.008]
LoE	-0.057	0.021	-0.122	0.035	0.057	-0.005
	(0.038)	(0.020)	(0.045)^a	(0.026)	(0.055)	(0.038)
	[0.060]	[0.024]	[0.065]^c	[0.026]	[0.049]	[0.052]
LoE×T	0.001	0.001	-0.019	0.008	0.002	-0.001
	(0.003)	(0.002)	(0.004)^a	(0.002)^a	(0.005)	(0.004)
	[0.001]	[0.003]	[0.015]	[0.004]^c	[0.007]	[0.011]
Sales (log)	0.004	0.024	-0.001	0.043	0.330	0.032
	(0.005)	(0.003)^a	(0.002)	(0.004)^a	(0.021)^a	(0.006)^a
	[0.004]	[0.009]^a	[0.005]	[0.013]^a	[0.088]^a	[0.022]
†Solid		0.148	-0.653	0.254	-0.401	-3.523
		(0.027)^a	(0.076)^a	(0.035)^a	(0.035)^a	(0.101)^a
		[0.091]	[0.007]^a	[0.184]	[0.142]^a	[0.061]^a
†Liquid		-0.292		1.284	-0.102	-3.278
		(0.052)^a		(0.064)^a	(0.063)	(0.174)^a
		[0.238]		[0.339]^a	[0.277]	[0.090]^a
†Injection	0.048	0.172	-0.643	0.675	-0.176	-3.327
	(0.024)^b	(0.032)^a	(0.080)^a	(0.060)^a	(0.059)^a	(0.116)^a
	[0.039]	[0.105]	[0.020]^a	[0.279]^b	[0.114]	[0.064]^a
†Ointment		0.090		0.817	1.199	-1.651
		(0.050)^c		(0.065)^a	(0.074)^a	(0.134)^a
		[0.163]		[0.259]^a	[0.282]^a	[0.117]^a
†1Form		-0.557		-1.250	-0.069	-1.149
		(0.053)^a		(0.046)^a	(0.135)	(0.180)^a
		[0.205]^a		[0.240]^a	[0.694]	[0.262]^a
†Chronic		0.082	1.483	0.336	-1.073	1.776

This table is an extended version of regression coefficients shown in the *upper block* of Table 4 in the main text. Robust errors are in first parentheses, and cluster adjusted standard errors in second square parentheses, with clustering at the originator level. Superscripts a, b, c indicate significance at 1%, 5% and 10%, respectively. All regressions include ATC2 dummies. †1/0 Dummy variable, 1 if true.

continued

TABLE B-2. Product launch rate by market size (for D1)

	Small		Medium		Large	
	With	W/out	With	W/out	With	W/out
	(1)	(2)	(3)	(4)	(5)	(6)
		(0.027)^a	(0.081)^a	(0.035)^a	(0.177)^a	(0.078)^a
		[0.100]	[0.012]^a	[0.195]^c	[1.031]	[0.151]^a
†SPC	-0.934	-0.489	-1.043	0.146	-0.182	0.903
	(0.031)^a	(0.038)^a	(0.022)^a	(0.030)^a	(0.032)^a	(0.090)^a
	[0.051]^a	[0.242]^b	[0.043]^a	[0.167]	[0.113]	[0.145]^a
†Monopoly		0.041	0.107	-0.158	-0.320	0.231
		(0.015)^a	(0.049)^b	(0.044)^a	(0.072)^a	(0.074)^a
		[0.074]	[0.060]^c	[0.199]	[0.030]^a	[0.103]^b
†Nearby	0.018	0.059	0.008	-0.122	-0.114	-0.004
	(0.017)	(0.011)^a	(0.018)	(0.025)^a	(0.044)^a	(0.054)
	[0.023]	[0.035]^c	[0.013]	[0.119]	[0.106]	[0.123]
Constant	1.922	1.746	3.697	1.442	-1.752	2.328
	(0.060)^a	(0.071)^a	(0.111)^a	(0.095)^a	(0.370)^a	(0.113)^a
	[0.066]^a	[0.329]^a	[0.181]^a	[0.435]^a	[1.647]	[0.309]^a
Observations	159	2,272	550	2,463	1,616	992
R^2	0.859	0.446	0.862	0.641	0.700	0.905
Originators	5	60	15	61	46	25

This table is an extended version of regression coefficients shown in the *upper block* of Table 4 in the main text. Robust errors are in first parentheses, and cluster adjusted standard errors in second square parentheses, with clustering at the originator level. Superscripts a, b, c indicate significance at 1%, 5% and 10%, respectively. All regressions include ATC2 dummies. †1/0 Dummy variable, 1 if true.

end

Table B-3 below corresponds to the selected coefficients shown in the *lower block* of Table 4 in the main text.

TABLE B-3. Product launch rate of by market size (for D2)

	Small		Medium		Large	
	With	W/out	With	W/out	With	W/out
	(1)	(2)	(3)	(4)	(5)	(6)
T	0.020	-0.011	0.032	-0.021	0.060	0.030
	(0.010)	(0.004)^a	(0.006)^a	(0.008)^a	(0.014)^a	(0.007)^a
	[0.020]	[0.008]	[0.015]	[0.012]^c	[0.023]^b	[0.016]^c
LoE	0.020	0.033	-0.019	0.092	0.577	0.137
	(0.156)	(0.056)	(0.117)	(0.101)	(0.235)^b	(0.121)
	[0.198]	[0.062]	[0.179]	[0.093]	[0.239]^b	[0.231]
LoE×T	-0.018	0.002	-0.052	0.028	-0.049	-0.043
	(0.012)	(0.005)	(0.009)^a	(0.010)^a	(0.022)^b	(0.011)^a
	[0.026]	[0.009]	[0.032]	[0.015]^c	[0.042]	[0.037]
Sales (log)	0.097	0.061	-0.028	0.135	2.390	0.102
	(0.017)^a	(0.008)^a	(0.009)^a	(0.015)^a	(0.118)^a	(0.013)^a
	[0.015]^a	[0.023]^b	[0.020]	[0.057]^b	[0.472]^a	[0.039]^b
†Solid		1.544	0.807	0.660	-2.240	-28.915
		(0.082)^a	(0.125)^a	(0.135)^a	(0.293)^a	(0.531)^a
		[0.396]^a	[0.026]^a	[0.663]	[1.129]^c	[0.107]^a
†Liquid		0.170		2.587	-2.882	-28.771
		(0.136)		(0.233)^a	(0.369)^a	(0.602)^a
		[0.514]		[1.176]^b	[1.570]^c	[0.157]^a
†Injection	-0.455	1.250	0.761	2.570	-0.665	-29.697
	(0.126)^a	(0.079)^a	(0.149)^a	(0.220)^a	(0.647)	(0.537)^a
	[0.258]	[0.359]^a	[0.082]^a	[0.980]^b	[2.312]	[0.114]^a
†Ointment		-0.062		1.178	0.330	-29.056
		(0.139)		(0.196)^a	(0.633)	(0.546)^a
		[0.570]		[0.868]	[2.663]	[0.197]^a
†1Form		-0.914		-1.202	-5.622	28.175
		(0.132)^a		(0.183)^a	(0.480)^a	(0.679)^a
		[0.700]		[0.816]	[1.607]^a	[0.755]^a

This table is an extended version of regression coefficients shown in the *lower block* of Table 4 in the main text. Robust errors are in first parentheses, and cluster adjusted standard errors in second square parentheses, with clustering at the originator level. Superscripts a, b, c indicate significance at 1%, 5% and 10%, respectively. All regressions include ATC2 dummies. †1/0 Dummy variable, 1 if true.

continued

TABLE B-3. Product launch rate of by market size (for D2)

	Small		Medium		Large	
	With	W/out	With	W/out	With	W/out
	(1)	(2)	(3)	(4)	(5)	(6)
†Chronic		1.037	1.996	-0.430	2.421	-0.520
		$(0.087)^a$	$(0.123)^a$	$(0.105)^a$	$(0.443)^a$	$(0.169)^a$
		$[0.429]^b$	$[0.046]^a$	[0.600]	[1.988]	$[0.257]^c$
†SPC	-0.714	-0.710	-2.229	0.828	-2.556	-1.374
	$(0.098)^a$	$(0.064)^a$	$(0.148)^a$	$(0.109)^a$	$(0.413)^a$	$(0.180)^a$
	$[0.227]^b$	$[0.293]^b$	$[0.125]^a$	[0.569]	[1.765]	$[0.249]^a$
†Monopoly		0.115	-0.486	-0.528	-0.401	0.348
		$(0.048)^b$	$(0.185)^a$	$(0.133)^a$	(0.263)	$(0.140)^b$
		[0.184]	$[0.168]^b$	[0.430]	[0.754]	[0.212]
†Nearby	-0.667	0.396	-2.757	-0.255	-1.734	1.563
	$(0.156)^a$	$(0.045)^a$	$(0.351)^a$	$(0.083)^a$	$(0.311)^a$	$(0.262)^a$
	$[0.264]^c$	$[0.146]^a$	$[0.311]^a$	[0.393]	$[0.718]^b$	$[0.724]^b$
Constant	3.276	1.044	9.704	2.633	-21.614	1.785
	$(0.284)^a$	$(0.173)^a$	$(0.444)^a$	$(0.325)^a$	$(1.765)^a$	$(0.301)^a$
	$[0.424]^a$	[0.827]	$[0.519]^a$	$[1.286]^b$	$[6.915]^a$	$[0.750]^b$
Observations	159	2,272	550	2,463	1,616	992
R^2	0.647	0.621	0.899	0.765	0.798	0.976
Originators	5	60	15	61	46	25

This table is an extended version of regression coefficients shown in the *lower block* of Table 4 in the main text. Robust errors are in first parentheses, and cluster adjusted standard errors in second square parentheses, with clustering at the originator level. Superscripts a, b, c indicate significance at 1%, 5% and 10%, respectively. All regressions include ATC2 dummies. †1/0 Dummy variable, 1 if true.

end

Table B-4 is analogous to Table 3. Primary difference is that it estimates random coefficients model instead of pooled OLS.

TABLE B-4. Product launch rate (random effects)

	D1 and D2 by samples A,B,C						D1 and D2 by subsamples of C			
	D1			D2			D1		D2	
	(1) A	(2) B	(3) C	(4) A	(5) B	(6) C	C(7) With	C(8) W/out	C(9) With	C(10) W/out
T	0.005^a	0.006^a	0.003^a	0.035^a	0.034^a	0.027^a	0.011^a	-0.000	0.102^a	-0.007^a
	(0.000)	(0.000)	(0.001)	(0.001)	(0.002)	(0.003)	(0.001)	(0.001)	(0.007)	(0.002)
LoE	0.023^b	0.013	0.012	0.115^a	0.146^a	0.124^b	-0.007	0.020	0.195	0.072^c
	(0.010)	(0.010)	(0.011)	(0.043)	(0.045)	(0.048)	(0.023)	(0.013)	(0.122)	(0.039)
LoE×T	-0.005^a	-0.005^a	0.001	-0.056^a	-0.055^a	-0.035^a	-0.005^b	0.004^a	-0.097^a	0.003
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.004)	(0.002)	(0.001)	(0.011)	(0.003)
Obs	13,559	12,560	8,052	13,559	12,560	8,052	2,325	5,727	2,325	5,727
Originators	263	212	212	263	212	212	66	146	66	146

Standard errors in parentheses. Superscripts a, b, c indicate significance at 1%, 5% and 10%, respectively. All regressions include other controls and ATC2 dummies. Sample (A) is all initial 263 originators, (B) restricts to 212 originators with sales observed before/after LoE, and (C) is same as (B) but with observations restricted to within 5 years of the LoE.

Table B-5 is analogous to Table 4. Primary difference is that it estimates random coefficients model instead of pooled OLS.

TABLE B-5. Product launch rate by market size (random effects)

	Small		Medium		Large	
	With	W/out	With	W/out	With	W/out
D1	(1)	(2)	(3)	(4)	(5)	(6)
T	0.000	-0.002^{b}	0.013^{a}	-0.002^{b}	0.005^{a}	0.008^{a}
	(0.003)	(0.001)	(0.003)	(0.001)	(0.002)	(0.002)
LoE	-0.057	0.026	-0.122^{a}	0.031	0.043	-0.005
	(0.053)	(0.017)	(0.045)	(0.019)	(0.028)	(0.040)
LoE×T	0.001	0.002	-0.019^{a}	0.009^{a}	0.005^{c}	-0.001
	(0.005)	(0.002)	(0.004)	(0.002)	(0.003)	(0.003)
D2	(1)	(2)	(3)	(4)	(5)	(6)
T	0.020^{b}	-0.012^{a}	0.032^{a}	-0.022^{a}	0.082^{a}	0.030^{a}
	(0.010)	(0.003)	(0.007)	(0.004)	(0.010)	(0.007)
LoE	0.020	0.029	-0.019	0.086	0.360^{b}	0.137
	(0.165)	(0.044)	(0.108)	(0.062)	(0.158)	(0.120)
LoE×T	-0.018	0.000	-0.052^{a}	0.028^{a}	-0.062^{a}	-0.043^{a}
	(0.016)	(0.004)	(0.010)	(0.005)	(0.015)	(0.010)
Observations	159	2,272	550	2,463	1,616	992
Originators	5	60	15	61	46	25

Standard errors in parentheses. Superscripts a, b, c indicate significance at 1%, 5% and 10%, respectively.

Table B-6 gives results for truncated Poisson model and can be compared to Table 3 for the pooled OLS regressions for D1 and D2.

TABLE B-6. Product launch rate (truncated poisson regression)

	D1 and D2 by samples A,B,C						D1 and D2 by subsamples of C			
	D1			D2			D1		D2	
	(1) A	(2) B	(3) C	(4) A	(5) B	(6) C	C(7) With	C(8) W/out	C(9) With	C(10) W/out
T	0.017^a	0.015^a	0.005	0.012^a	0.011^a	0.007^a	0.010	0.000	0.014^a	-0.001
	(0.006)	(0.005)	(0.005)	(0.002)	(0.002)	(0.002)	(0.006)	(0.006)	(0.003)	(0.003)
LoE	-0.058	-0.079	0.022	-0.002	0.016	0.063^b	-0.002	0.033	0.034	0.031
	(0.061)	(0.055)	(0.037)	(0.040)	(0.040)	(0.026)	(0.048)	(0.053)	(0.029)	(0.039)
LoE×T	-0.007	-0.005	0.003	-0.015^a	-0.013^a	-0.011^a	-0.008	0.010	-0.012^b	-0.000
	(0.007)	(0.006)	(0.007)	(0.004)	(0.004)	(0.003)	(0.009)	(0.008)	(0.005)	(0.005)
Observations	13,559	12,560	8,052	13,559	12,560	8,052	2,325	5,727	2,325	5,727
Originators	263	212	212	263	212	212	66	146	66	146

Clustered standard errors are in parentheses. Superscripts a, b, c indicate significance at 1%, 5% and 10%, respectively. All regressions include other controls and ATC2 dummies. Sample (A) is all initial 263 originators, (B) restricts to 212 originators with sales observed before/after LoE, and (C) is same as (B) but with observations restricted to within 5 years of the LoE. Column (1) uses dependent variable (D1-1) and estimates a poisson model instead of truncated poisson as the latter did not converge.

Table B-7 gives results for truncated Poisson model and can be compared to results for D2 in Table 4 for the pooled OLS regressions.

TABLE B-7. Product launch rate by market size (truncated poisson regression)

	Small		Medium		Large	
	With	W/out	With	W/out	With	W/out
D2	(1)	(2)	(3)	(4)	(5)	(6)
T	0.016^{c}	-0.004	0.010^{a}	-0.007	0.010^{a}	0.009^{a}
	(0.009)	(0.006)	(0.004)	(0.005)	(0.003)	(0.003)
LoE	-0.032	0.011	-0.009	0.078	0.047	0.019
	(0.099)	(0.044)	(0.052)	(0.063)	(0.032)	(0.046)
LoE×T	-0.015^{c}	-0.008	-0.016^{b}	0.012^{c}	-0.006	-0.013^{b}
	(0.009)	(0.008)	(0.008)	(0.007)	(0.006)	(0.006)
Observations	159	2,272	550	2,463	1,616	992
Originators	5	60	15	61	45	25
Log pseudo LL	-194.24	-1813.95	-812.50	-2975.52	-3137.37	-1316.40
Pseudo R2	0.1387	0.4038	0.2759	0.4386	0.4763	0.6473

Clustered standard errors in parentheses. Superscripts a, b, c indicate significance at 1%, 5% and 10%, respectively.

B.2. **Monotonicity test.** This appendix implements a test based on a monotonicity argument proposed in Ellison and Ellison (2011, 2000 working paper) and in Dafny (2005). The general idea is that firms' investments may be monotonically related to profitability/size of the market, but the relation may not be monotone when there are strategic considerations. For instance, a firm may introduce more product varieties to match consumer tastes more closely in larger markets than in smaller markets. Thus, absent any strategic considerations, there might be a monotonically increasing relationship between the size of the market and the number of products launched by the originator. But size/profitability may also be correlated with the risk of entry, and the incumbent may change their investment if they can deter entry. In small markets, high entry costs relative to profits may mean that entry is blocked, while in large markets deterrence may not be feasible, and hence originators may not take any deterring action in either case. By contrast, in medium size markets there may be incentives to over-invest in product launches, thus breaking the monotone increasing relation. To this end, we estimated the specification suggested in Ellison and Ellison (2011), given by

$$D\#_j = \beta_0 + \beta_1 \ln(\text{Sales})_j + \beta_2(\ln(\text{Sales})_j - \overline{\ln(\text{Sales})})^2 + X_j\gamma + \epsilon_j. \qquad (4)$$

In the equation above, $D\#_j$ is one of the two measures of count of products launched by the originator, and the term $(\ln(\text{Sales})_j - \overline{\ln(\text{Sales})})^2$ captures the deviation in sales of the jth originator relative to the mean value of sales for all originators. We used cross-sectional observations where values for all variables were computed using the average value from two years prior to the LoE (and in the handful of cases where entry took place before the LoE, we used average values from two years prior to entry or by dropping those observations). A negative and significant value of β_2 would indicate a break from the monotonic relationship where originators in medium size markets (or closer to the mean) launch more products than those in small or large markets.

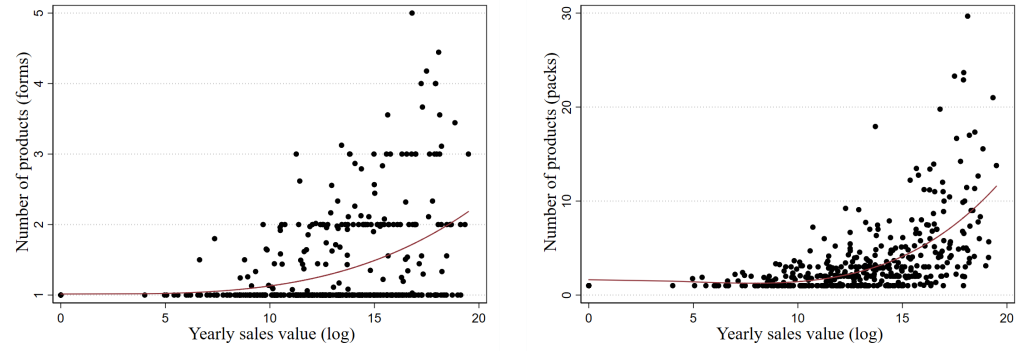

FIGURE B-1. Count of products and market size

We estimated the model using $D1$ and $D2$ measures of count of products on both the smaller and the larger sample of originators, as well as on a different variant where size of the market measured as log of sales was replaced by the likelihood of entry using a simple probit model of entry. Results from these six cases are given in the appendix in Table B-8, but based on this test, we did not find any evidence consistent with entry deterrence based on these tests. In fact the coefficient β_2 turned out to be positive and significant in most cases, as would be the case if the underlying relation is very steep or convex. The convexity can be seen in Figure B-1, in which case the test is not applicable, and hence we do not discuss it further.

TABLE B-8. Non-monotonicity tests, full model

	Size (Sub sample)		Size (full sample)		‡Entry risk (full sample)		
	(1 - D1)	(2 - D2)	(3 - D1)	(4 - D2)	(5 - D1)	(6 - D2)	($P(E)_j = 1$)
$\ln(S)_j$	0.102^a	0.700^a	0.092^a	0.652^a			0.09
	(0.013)	(0.075)	(0.011)	(0.056)			(0.428)
$(\ln(S)_j - \overline{\ln(S)})^2$	0.009^a	0.064^a	0.007^a	0.054^a			
	(0.002)	(0.010)	(0.001)	(0.007)			
$P(E)_j$					1.11^a	5.823^a	
					(0.25)	(1.319)	
$(P(E)_j - \overline{P(E)})^2$					-0.211	3.499	
					(0.595)	(2.999)	
†SPC	-0.246^b	-1.899^a	-0.241^a	-1.221^a	-0.081	-0.214	-0.844^a
	(0.122)	(0.687)	(0.072)	(0.370)	(0.1)	(0.572)	(0.272)
†Nearby	0.310^b	1.266	0.222^b	1.005^c	0.215^b	0.882^c	0.236
	(0.137)	(0.771)	(0.103)	(0.532)	(0.1)	(0.498)	(0.506)
†Monopoly	-0.055	0.735	-0.071	0.706	-0.072	0.634	-0.167
	(0.184)	(1.037)	(0.136)	(0.699)	(0.109)	(0.518)	(1.42)
†Chronic	0.240^b	-0.293	0.160^c	-0.285	0.113	-0.395	0.401
	(0.122)	(0.684)	(0.087)	(0.451)	(0.111)	(0.473)	(0.379)
†Solid	-0.008	0.082	0.129	0.581	0.076	0.34	0.327
	(0.175)	(0.986)	(0.121)	(0.625)	(0.16)	(0.958)	(0.579)
†Liquid	0.058	-1.402	0.439^b	0.327	0.451^b	0.254	-0.468
	(0.290)	(1.634)	(0.201)	(1.036)	(0.226)	(1.349)	(2.11)
†Injection	0.432^b	1.059	0.381^a	1.093	0.345^b	0.844	0.001
	(0.193)	(1.083)	(0.132)	(0.681)	(0.164)	(0.927)	(0.648)
†Ointment	0.300	1.522	0.357	0.847	0.547^b	1.933	-7.139
	(0.355)	(1.996)	(0.242)	(1.245)	(0.254)	(1.433)	(6.192)
†1Form	-0.583^b	-2.883^b	-0.838^a	-2.823^a	-0.869^a	-3.089^c	-0.103
	(0.232)	(1.302)	(0.176)	(0.909)	(0.249)	(1.441)	(0.886)
Constant	0.267	-2.436	0.581	-3.013	1.405^a	2.983	-3.239
	(0.600)	(3.377)	(0.506)	(2.607)	(0.367)	(2.102)	(4.186)
$(\ln(\text{Sales})_j)^2$							0.011
							(0.015)
Observations	263	263	430	430	430	430	430

All regressions include dummies for ATC2 class. Superscripts a, b, c indicate significance at 1%, 5% and 10%, respectively. †1/0 Dummy variable, 1 if true. ‡ This is a two-step model where in step 1, probability of entry for the j-th originator is computed via a probit (results shown in the last column), and in step 2, size of the market is replaced with the probability. Standard errors for the two-step model are computed using bootstraps with 500 replications.

B.3. **Regression coefficients for the hazard models.** This appendix replicates results from Table 5 which controlled from D1, but instead now controls for D2.

TABLE B-9. Discrete time hazard models for S1 & S2 (controlling for D2)

	S=S1				S=S2			
	(1-A)	(2-B)	(3-C)	(4-D)	(5-A)	(6-B)	(7-C)	(8-D)
S	1.337^b	1.038	1.628	2.587	2.379^a	2.593^a	2.719^a	3.182^a
	(0.671)	(0.829)	(1.034)	(1.724)	(0.826)	(0.832)	(1.049)	(1.057)
Medium	3.253^a	4.320^a	5.036^a	6.471^a	-0.060	0.563	0.378	0.407
	(1.156)	(1.379)	(1.710)	(2.418)	(0.361)	(0.543)	(0.669)	(0.774)
Large	2.168^b	2.149^c	2.357	3.599	0.488	1.607^b	1.095	1.043
	(0.944)	(1.219)	(1.555)	(2.277)	(0.532)	(0.769)	(1.008)	(1.146)
Medium×S	-3.159^a	-3.741^a	-4.620^a	-6.001^a	-3.187^b	-3.396^a	-3.576^b	-3.888^b
	(0.990)	(1.179)	(1.435)	(2.017)	(1.277)	(1.244)	(1.440)	(1.698)
Large×S	-1.816^b	-1.243	-1.772	-3.182^c	-3.457^a	-3.511^a	-3.890^a	-5.943^a
	(0.749)	(0.949)	(1.140)	(1.755)	(1.069)	(1.316)	(1.483)	(1.472)
Sales (log)	0.274^a	0.234^b	0.351^b	0.465^a	0.261^a	0.172^c	0.309^b	0.414^a
	(0.077)	(0.104)	(0.146)	(0.142)	(0.081)	(0.093)	(0.141)	(0.140)
D2		0.049^c	0.044	0.029		0.045^c	0.043	0.023
		(0.026)	(0.031)	(0.032)		(0.025)	(0.029)	(0.029)
				Marginal effects: $\partial\lambda/\partial S$ (×100)				
Small	0.402^b	0.345	0.807	1.282	0.807^b	0.979^b	1.342^c	1.566^b
	(0.203)	(0.275)	(0.523)	(0.807)	(0.407)	(0.397)	(0.702)	(0.733)
Medium	-1.281^b	-2.077^a	-3.026^a	-3.451^a	-0.563	-0.621	-0.868	-0.714
	(0.545)	(0.693)	(1.118)	(1.211)	(0.676)	(0.792)	(1.200)	(1.532)
Large	-0.967	-0.465	-0.597	-2.464	-2.484	-2.440	-4.860	-11.427^b
	(0.600)	(1.167)	(2.245)	(2.278)	(1.557)	(2.902)	(4.715)	(4.53)
Includes X_j?	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Originators	430	386	312	312	410	364	312	312
Entry events	137	137	104	104	131	131	104	104
Observations	13,456	12,063	5,961	5,961	11,848	10,444	5,961	5,961
Log likelihood	-664	-607	-412	-414	-623	-562	-414	-414

Notes. Clustered standard errors are in parenthesis and superscripts a, b, c indicate significance at 1%, 5% and 10%, respectively. All models include duration dummies and ATC2 dummies included in all but columns (1) and (5). Table 5 in the main text gives coefficients when controlling for $D1$.