Incentivizing novelty in antibiotic development

Gosia Majewska*
Toulouse School of Economics

November 2022

Job Market Paper - click here for the latest version

Abstract

In antibiotics, a constant supply of new products is needed as bacteria become resistant to the existing drugs. I estimate the effectiveness of innovation incentives for antibiotics, introduced in 2012. In a difference-in-differences framework, I find that the incentives have a positive effect on clinical trial success rates, but only for projects using known technologies. To assess the long-term effect of the incentives on market entry, I set up a dynamic structural model of pharmaceutical innovation. The multi agent setting of the model allows the firm decisions to depend not only on the projects' expected cost and profit, but also on the outcomes of technologically close projects. Counterfactual simulations show a 20% increase in the number of market entries due the current incentive scheme, driven mostly by research subsidies.

*Toulouse School of Economics, University of Toulouse Capitole, gosia.majewska@tse-fr.eu. I thank my advisor Pierre Dubois for his support and advice, as well as participants of internal workshops at TSE, the Workshop on the Economics of Antibiotics at TSE and the INAMRSS/CeBIL AMR Workshop in Copenhagen. The author gratefully acknowledges funding from the public funding under the 'PSPC' call for projects operated on behalf of the French government by Bpifrance as part of the 'Investments for the Future' program (PIA) ARPEGE and from the TSE-P Health Center whose lists of sponsors can be found at https://www.tse-fr.eu/health

1 Introduction

Antibiotics are a fundamental element of modern medicine, substantially improving the safety of treatments that expose patients to infection risk, e.g., by requiring surgical intervention or suppressing the immune system. However, bacteria, which antibiotics are supposed to neutralize, are sophisticated enough to develop defense mechanisms – a problem known as antimicrobial resistance (AMR) – making antibiotics less effective over time. Resistant infections are a major health problem and one of the top priorities in the WHO's agenda. Recent estimates attribute to AMR more than 33 000 deaths in the EU in 2015 (Cassini et al., 2019), and 1.27 million deaths globally in 2019 (Murray et al., 2022), a magnitude comparable to HIV and malaria.

AMR creates an urgent need for new antibiotics. At the same time, managing resistance requires that the new drugs remain reserved for cases where other treatment options fail, and their use is limited to a strict minimum, making the potential market size small. Recognizing that the market size is insufficient to stimulate enough innovation, additional incentives have been provided since 2012 for the development of new antibiotics. The existing incentives address both the high cost of developing new antibiotics by providing additional funding for firms (push incentive) and the low reward for successfully bringing a new antibiotic to the market by increasing the market exclusivity period (pull incentive).

I estimate the effect of the push and pull incentives on antibiotic innovation in the short and long term. The analysis follows in two steps. First, I identify the causal effect of the incentives on a short-term innovation outcome – clinical trial success. Then, I set up a structural model of pharmaceutical innovation which allows me to consider the long-term effects of the incentives. Pharmaceutical innovation is a long process, as clinical trials alone take on average 10 to 12 years. Using the model estimates, I simulate a longer time period and estimate the effect on new market entries.

In a difference-in-differences setting, I find that the incentives have successfully invigorated the antibiotic pipeline by encouraging more investment in projects eligible for the pull incentive, increasing their success rates in clinical trials by 21 pp. compared to projects predicted to be eligible but active before 2012. For the push incentive, I find

that an additional million of dollars allocated to a firm increases the success probability of its projects in clinical trials by 1.3 pp., suggesting that antibiotic innovators are facing financial constraints that can be alleviated by the incentive.

To model the innovation process and estimate the long-term effect of the incentives on market entry, I adapt a continuous-time dynamic discrete choice model, developed by Arcidiacono et al. (2016) and first used in the pharmaceutical innovation context by Khmelnitskaya (2021). In the model, firms make decisions to continue or terminate the development of an antibiotic following its value, which is driven by the development cost and expected profit – affected by incentives – and the outcomes of other projects within the same technological class, which can be informative about the probability of success. The results reveal that research subsidies are crucial for progress through the early phases of clinical trials, while the effect of the pull incentive on the expected profit is modest and affects firms' decisions only at the final development stage.

New antibiotics can fall into two categories. First, they can be new variations of already existing molecules, new-in-class innovation. Second, new molecules can establish their own classes. The latter is the truly breakthrough innovation that brings real progress in the fight against AMR, while the new in class molecules bring only marginal benefits. As bacteria might already have developed defense mechanisms against the existing antibiotics, resistance against new in class drugs might emerge much quicker than against a completely new antibiotic (WHO, 2021).

In my analysis, I find no effect of either incentive on clinical trial success rates of novel antibiotics. In the structural model, early stage decisions are affected almost exclusively by the push incentive, meaning that in the current incentive scheme, novel projects could only be promoted if research subsidies are directed explicitly to those projects.

This paper contributes to the literature on innovation in the pharmaceutical industry. Existing papers focus mostly on the effects of market size (Acemoglu and Linn, 2004; Dubois et al., 2015), mergers (Grabowski and Kyle, 2008; Ornaghi, 2009; Haucap et al., 2019; Cunningham et al., 2021), and recently, on strategic considerations in drug development (Rao, 2020; Khmelnitskaya, 2021). Innovation incentives in the pharmaceutical

industry have so far been studied market-wide for tax credits (McCutchen Jr, 1993)) and for the cases of incentives targeting rare diseases (Sarpatwari et al., 2018; Yin, 2008) and neglected diseases (Grace and Kyle, 2009), but not antibiotics. As the degree of innovativeness is particularly relevant for new antibiotics, and an important dimension in my analysis, this paper contributes also to the strand of literature studying novelty in drug development (Krieger et al., 2021, 2022).

This paper extends the nascent empirical literature on the economics of antibiotics by looking at the supply side issues related to AMR. So far, the focus in the area has been on the demand side, studying the interactions between resistance and physician prescriptions behavior (Huang and Ullrich, 2021; Huang et al., 2022; Adda, 2020; Ribers and Ullrich, 2022).

This exercise addresses not only the very relevant question of how to bring to the market new, effective antibiotics but can also guide future policy in pharmaceutical innovation more broadly. With low-hanging fruit seemingly picked in pharmaceutical R&D, continuing on relying on monopoly profits from successful projects might make the incentives of the industry diverge from the priorities of the consumers. Antibiotics constitute a useful case study of a market where additional incentives have been introduced, and their careful evaluation can inform the design of better, future incentive schemes for other drug classes.

This paper is structured as follows. Section 2 overviews the innovation process in pharmaceuticals and the existing innovation incentives for antibiotics. Section 3 describes the dataset used and the definitions of the key variables as well as some summary statistics. Section 4 presents the reduced form analysis and the structural model with simulations. Section 5 concludes.

2 Background

Pharmaceuticals are a tightly regulated industry, and, typically, new products must obtain approval to enter a market. The steps required for approval are similar across the major national markets. In this paper, I focus on the procedures in the US handled by the Food

and Drug Administration (FDA). The US market is the largest and typically the first market on which a new drug is launched.

2.1 Pharmaceutical Innovation

New drugs go through a long process of clinical trials to gain regulatory approval and enter the market. Pharmaceutical labs identify biological targets and test how different agents affect them in vitro and in animal trials. From this preclinical stage, the most promising molecules move on to (human) clinical trials, where the safety and efficacy of drug candidates are tested in 3 phases. Clinical trials are long and expensive, and their outcomes are uncertain, with less than 15 percent of molecules entering Phase I being launched on the market. Based on data from successful clinical trials, a pharmaceutical company can request approval from a government regulator (e.g., the FDA in the US or the European Medicine Agency (EMA) in the EU). Upon approval, the pharmaceutical company is granted marketing exclusivity by the regulator (5 years in the US, in parallel to any patent protection the molecule might have) and becomes a monopolist for its drug. After the exclusivity period expires, other manufacturers (subject to a simplified approval process) are allowed to produce the same molecule and compete with the originator.

2.2 Innovation incentives for antibiotics

2.2.1 Push incentives

Initiatives providing funding for AMR research started to gain momentum in the 2010s. Simpkin et al. (2017) reviews extensively the main organizations providing push incentives for antibiotic development. The most important funders are the governments of the United States (through National Institute of Health grants and a specialized agency, the Biomedical Advanced Research and Development Authority –BARDA, which has had an antibiotic focus since 2008), the UK and the EU (through the European Commission's Directorate-General for Research and Innovation as well as the Innovative Medicines Initiative's (IMIs) New Drugs for Bad Bugs (ND4BB) program launched in 2012, which is a public private partnership with the European Federation of Pharmaceutical Industries and Associations).

Since 2016, there have been two important multilateral organizations: the Global Antibiotic Research and Development Partnership (GARDP) and Combating Antibiotic Resistant Bacteria Biopharmaceutical Accelerator (CARB-X). In addition to providing push funding, the GARDP and CARB-X offer expertise and technical assistance to grantees. Moreover, CARB-X focuses explicitly on antimicrobials targeting high-priority pathogens.

To my knowledge, push incentives for antibiotics have not been evaluated, but the literature on funding programs in other areas can be informative. Howell (2017) finds that subsidies at early innovation stages have a large impact on the survival of the funded projects and allow them to secure more venture capital funding than the control projects.

2.2.2 GAIN Act

The Generating Antibiotic Incentives Now (GAIN) was passed in 2012 as a part of the 2012 Food and Drug Administration Safety and Innovation Act, and it introduced the Qualified Infectious Disease Product (QIDP) designation, a special status for antibiotics targeting resistant infections. The QIDP status provides two main benefits: an extended market exclusivity period and privileges in the approval process.

The GAIN Act defines eligibility for the QIDP designation as follows: *"antibacterial or antifungal drug for human use intended to treat serious or life-threatening infections"*. Drugs meeting these criteria can apply for the QIDP designation at any development stage. Once granted the designation, the drugs are additionally eligible for the fast track designation, which means more frequent exchanges with FDA officials, including regarding the design of the clinical trials, and as a consequence, a decrease in uncertainty over the result of the regulatory review process. QIDP products are also granted priority review, meaning that the FDA must give its decision in 6 rather than 10 months at the approval stage. Finally, successful QIDP products are given an additional 5 years of market exclusivity.

The literature is largely skeptical about the GAIN Act's potential to fulfill its role. Ambrose (2011) is an early critique of the GAIN Act, saying it is insufficient to incentivize firms to innovate more, proposing further measures (e.g., R&D tax credits, advance market commitments, longer exclusivity). Outterson (2013) criticizes the GAIN Act for its

lack of provisions concerning stewardship and the appropriate use of the new antibiotics. Outterson et al. (2015) expresses doubts about whether the GAIN Act provides a large enough incentive for firms to invest in antibiotic development and underlines that there is no prioritization among the pathogens to be targeted by QIDP products.

On the other hand, the GAIN Act resembles the Orphan Drug Act of 1983, which has been shown to be largely successful. The Orphan Drug Act targets rare diseases (affecting less than 200 000 people in the US) and provides 7 additional years of marketing exclusivity for the rare disease indication. Sarpatwari et al. (2018) find that after passing the Orphan Drug Act, the number of patents associated with orphan drug indications increased, as did exclusivity periods of the products with the designation. Yin (2008) shows an increase in clinical trials associated with rare diseases after the introduction of the Orphan Drug Act, especially for the more prevalent among rare conditions.

3 Data

3.1 Data Sources

3.1.1 Innovation

To track the development of antibiotics, I use the Pharmaprojects dataset from Citeline. The data contain information on more than 81 000 pharmaceutical projects starting from the 1980s and up to the end of 2021. Pharmaprojects tracks their development histories from the preclinical stage to market launch (or discontinuation if the project is unsuccessful). Additionally, the dataset provides detailed information about the project: its chemical and biological properties, targeted indications and therapeutic classes, licensing and marketing information. Importantly for my analysis, the data mention whether the drug candidate has been granted the QIDP status.

The *originator* field in the dataset reflects the owner of the project at the end of 2021. As ownership changes can affect innovation outcomes, both as a direct effect of the transaction and the effect of the size of the firm, it is important to recreate the ownership history of each project. The dataset contains information on the identities of the previous owners but not on the timing of ownership changes. However, using Citeline's Medtrack,

a comprehensive list of M&A deals in the pharmaceutical industry, I infer the ownership history of each project and reassign the originator in each year the project was active.

3.1.2 Research grants

To track investments in antibiotic R&D, I use the data collected by the Global AMR R&D Hub – a partnership of countries, nongovernmental donor organizations and intergovernmental organizations to address challenges and to improve coordination and collaboration in global AMR R&D. I use data only on investments targeted at developing therapeutics, the second most-funded research area after basic research. For each payment, I know the funding dates and the identities of the grantee and the funder. I match the funding data to the innovation dataset by aggregating to the grantee-year level (by starting date of the investment) and matching to the grantee projects.

Figure 1 presents the development of the push incentives matched to the main dataset over time. In the early years, 2012-2014, the investments were few and the amounts were smaller than they were in the subsequent years, with an average of $2 million. Between 2015 and 2019, the number of payments made steadily increased and the total amount invested ranged between $100 million and $200 million, with an average investment at approximately $5 million.

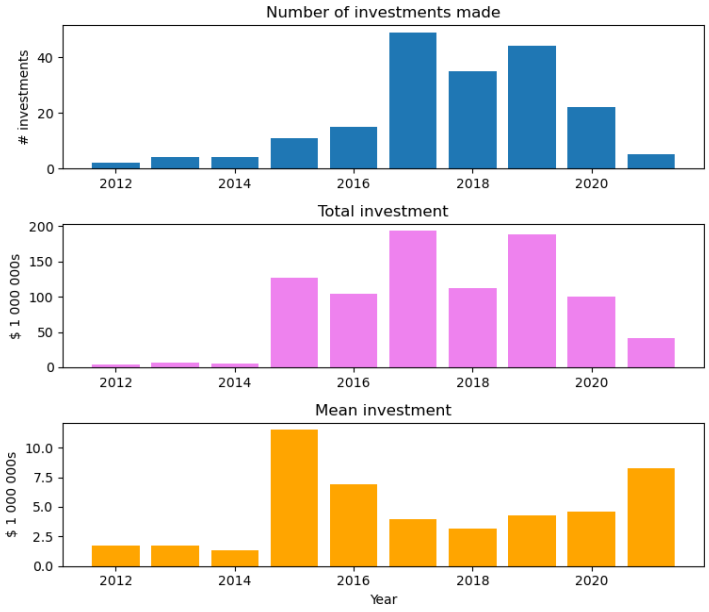
3.2 Antibiotics

While the Pharmaprojects dataset distinguishes 'antibiotics' and 'antibacterials' as separate therapeutic classes, those designations are not enough to cover either all products developed to fight infections caused by bacteria or all QIDP projects.

I construct a broader antibiotic class, considering their therapeutic classes and disease targets. I use keywords such as 'antibiotic", 'antibacterial', 'anti-infective', and 'infection' as well as a redacted list of indications of the QIDP projects, and I exclude projects targeting viral infections and vaccines.

Within antibiotics, I also create dummies for projects targeting specific, resistant bacteria. To do that, I search the 'summary' field of the dataset for the names of resistant

Figure 1: *Push incentive evolution over time*



Aggregate statistics from the Global AMR R&D Hub data on investments targeted at developing new antibiotics, using only values matched to firms present in the PharmaProjects innovation dataset.

bacteria[1], allowing for variations in spelling (e.g., for the bacteria name 'Escherichia coli', I also allow 'E. coli')

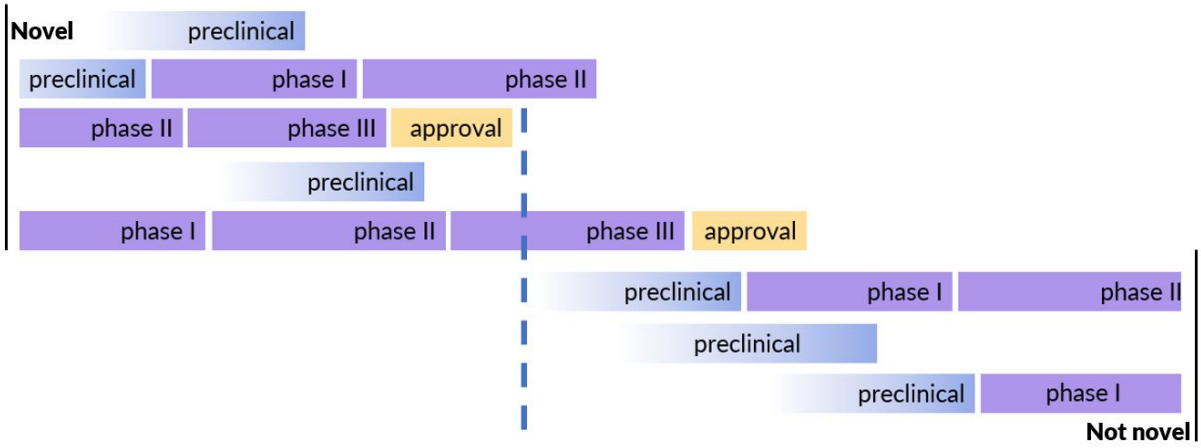
3.3 Novelty

In antibiotics, it is particularly important that new products represent true novelty and differ substantially from existing drugs to delay the emergence of resistance. In the case of drugs, novelty can have two dimensions: structural (chemical structure of the molecule) and functional (biological effects). Krieger et al. (2022) show that these two properties largely coincide (similar molecules have similar effects), although important exceptions exist. For antibiotics, the focus tends to be more on functional novelty, with expert reviews (e.g., WHO (2021)) paying attention to biological targets and modes of action. I define my novelty measure in line with this practice and base it on the technology used by the antibiotic.

To define the drug's technology, I follow Krieger et al. (2021) and Cunningham et al.

[1]Following lists of priority pathogens created by the ECDC and FDA

Figure 2: *Novelty definition*



Each horizontal line represents a project progressing through development phases over time (horizontal axis). All projects active to the left of the vertical dashed line (first launch in technology) are considered novel. Projects that start to the right of the dashed line are not novel.

(2021) and use the string variable *mechanism of action*. This variable which summarizes the biological target of the molecule and how it affects it, e.g., *protein synthesis inhibitors*, which stop or slow cell growth.

I define a project as novel or not by comparing the mechanism of action (technology) it uses to the stock of technologies used in successful (launched on the market) projects at the time I first observe the project. If, when the project starts, its technology has been used in already launched drugs, it is not novel. Novel projects use technologies that may have been used in earlier projects but not in those that successfully made it to the market. Figure 2 illustrates this definition.

Comparing with the evaluation of the novelty of the main new antibiotics approved since 2017 in WHO (2021), I correctly classify as novel the two products that represent new chemical classes, vaborbactam and lefamulin.

3.4 Summary statistics

3.4.1 Sample composition

Table 1 presents some descriptive statistics of the project-phase level sample used for estimation, and the subsamples of antibiotics and antibiotics receiving incentives. Pharmaceutical innovation is known for its high degree of attrition throughout the development process, and the table shows that this pattern is true for antibiotics regardless of the in-

Table 1: *Sample composition*

	Number of observations							
	Overall		Antibiotics		Pull incentive		Push incentive	
	N	% novel	*N*	% novel	*N*	% novel	*N*	% novel
Preclinical	27 840	70.42	1 147	52.40	98	55.10	149	75.17
Phase 1	4 655	74.59	168	58.93	40	60.00	39	79.49
Phase 2	3 387	76.14	138	55.07	38	60.53	36	80.56
Phase 3	1 910	66.65	94	52.13	34	47.06	14	78.57
Total	37 792	71.25	1547	53.33	210	55.71	238	76.90

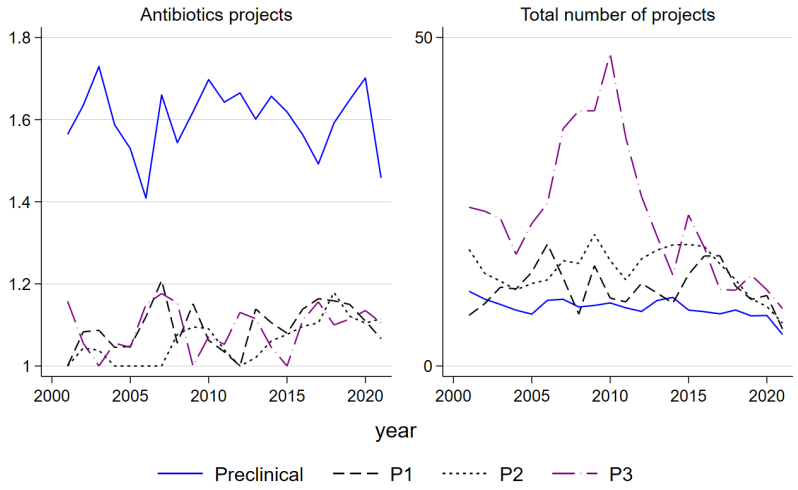
centives the projects receive, and projects in Phase 3 constitute a small fraction of the number of projects active in the Preclinical phase. On average, antibiotics are less novel than drugs from other classes, and the mean novelty decreases further for antibiotics further in clinical trials. The mean novelty of projects eligible for the pull incentive is close to the class average, while the push incentive seems to be given to more novel projects.

3.4.2 Antibiotic innovators and their portfolios

Figure 3 overviews the portfolios of firms active in antibiotic development. The left side of the graph shows the average number of antibiotic projects per phase across firms active in the given phase. On average, a firm innovating in antibiotics has approximately 1.6 antibiotics at the early stage (i.e., preclinical level) and fewer, approximately 1.1 projects, at any of the three clinical trial stages. This result appears constant across time. The right side of the graph shows the average firm size (in terms of all projects, in all drug classes) across firms active in each development phase in antibiotics. The mean firm size is the smallest among firms active in preclinical development, consistent with the idea of small biotech firms inventing new molecules but not being able to finance clinical trials. In preclinical trials, the average firm size has also remained stable in the last 20 years, while in clinical trials, especially in Phase 3 and since 2012, the average firm size has steadily decreased, approaching that of firms active in preclinical trials, consistent with the exit of big pharma from antibiotic development.

Figure 4 compares size and portfolio advancement among all antibiotic innovators and

Figure 3: *Antibiotic pipeline and firm size*

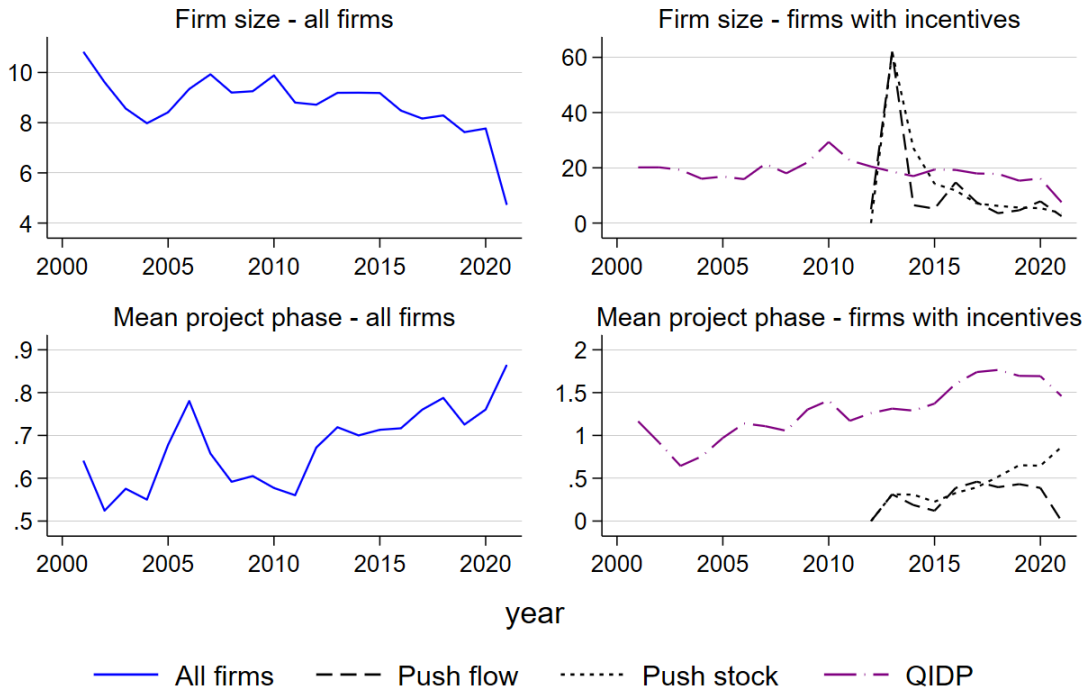


Graph on the left: mean number of projects active by development phase by firm, conditional on being active in the given phase in antibiotic development. Graph on the right: mean total number of projects active in all drug classes by firm, conditional on being active at the in the given phase in antibiotic development.

those that have been exposed to the push and pull incentives. The exit of large firms indicated in Figure 3 is once again visible in the top left part of the figure, although not as pronounced, given that early-stage projects constitute a large share of the antibiotic projects. Focusing on firms with incentives reveals that firms with the pull incentive (QIDP) are more than twice as large as the average firm in the market, while firms receiving the push incentive are closer to the average (with the exception of a large grant given to GlaxoSmithKline, a big pharma firm, in 2013). This result suggests that the push funding focuses on small firms that have limited other options for financing their projects. In the bottom panel of Figure 4, I plot the mean advancement of the firm portfolio, or the mean project phase, calculated by assigning the value of 0 to projects in the preclinical phase, 1 to projects in Phase 1, 2 to projects in Phase 2, and 3 to projects in Phase 3. While the mean project phase was rather stable between 2000 and 2010, a substantial upward trend can be observed since 2011. This result could be driven by either more projects advancing into later phases or fewer new projects entering development. The bottom right part of the graph shows that the portfolio advancement has slowly increased, particularly for firms with the pull incentive but also for firms that have received push funding at

any time in the past (push stock in the figure), suggesting that the incentives help firms advance their projects. Finally, the mean project phase does not change over time for firms receiving the push incentive at the time of the payment (push flow), confirming that the profile of the firm receiving the push incentive has not changed over time.

Figure 4: *Portfolios of firms active in antibiotic development by incentive type*

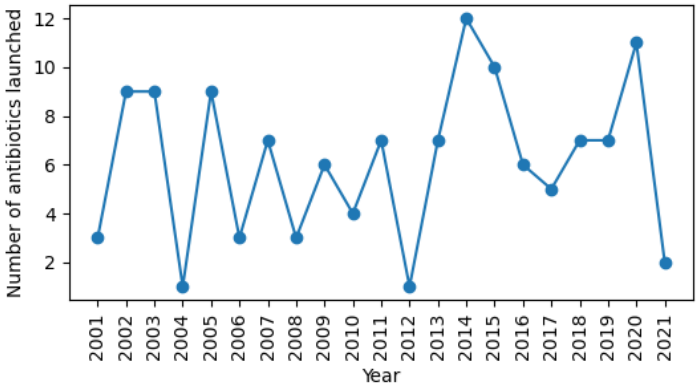


3.4.3 Launches

Figure 5 plots the number of antibiotic launches per year. Between 2000 and 2011, the number of launches annually remained stable, and the market saw 2-9 antibiotic launches per year. Striking are the drop to 1 in 2012 and the subsequent jumps to 7 and 12 in 2013 and 2014, respectively. This result could be partially explained by the firms anticipating the passing of the GAIN Act in 2012 and delaying the launch of their molecules to benefit from the QIDP designation. Interestingly, the number of launches remains at a higher level between 2014 and 2020, ranging from 5 to 11.

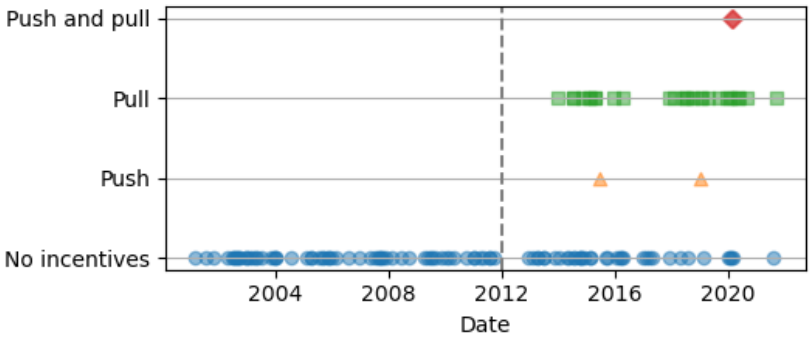
Figure 6 plots launches by their exact date and the incentive category. The jump in launches in 2013-14 did not translate into QIDP launches, as only one product with the

Figure 5: *Total antibiotic launches by year*



designation was launched in 2013 and 3 in 2014. More than half of the QIDP launches occurred in 2018 and later, more than 6 years after the GAIN Act was implemented. Only 3 projects that were exposed to a push incentive during their development have been launched thus far.

Figure 6: *Antibiotic launches by incentive type.*



Dashed line indicates the year the GAIN Act was passed.

4 Analysis

4.1 Overview

Due to the length of the pharmaceutical R&D process, the effects of innovation incentives could materialize decades after their introduction. In this section, I propose 2 approaches to evaluate the effectiveness of innovation incentives in antibiotics in the relatively short term, that is, the 10 years that passed between their introduction and the end of my dataset.

First, using a reduced form approach, I estimate the incentive effect on a short-term outcome, the clinical trial success rate. Second, by modeling the innovation process and including the incentives in the model, I estimate the model primitives using the available data and simulate additional time periods to see the long-term effect on market entries.

Both parts of the empirical analysis are built around a simple decision rule of a pharmaceutical innovator, illustrated in Equation (1):

$$p_{success} \times \Pi > c \tag{1}$$

The firm will invest in a drug candidate only if the expected profit (discounted lifetime profit Π weighted by probability of completing clinical trials with success and gaining approval $p_{success}$) is higher than development cost c. I assume that firms rank drug candidates by the difference, $p_{success} \times \Pi - c$, and invest first in those with the highest difference values. The pull incentives affect the LHS of Equation (1), as they increase profit, while the push incentives decrease the cost of innovation.

4.2 Short-term effects: reduced-form evidence

With only 10 years of data after incentives were introduced, a reduced form analysis must rely on a short-term outcome. A useful feature of the long process of developing a new drug is that it is divided into distinct, standard phases, at the end of which the innovator must decide whether to invest further in the drug candidate or to terminate it. Defining phase success as the drug candidate advancing into a subsequent phase, I obtain an intermediary, project-level outcome.

Typically, the success of a drug candidate in a given phase (i.e., the project advancing to a subsequent phase), is the product of two factors interacting. First is the objective outcome of the phase (e.g., the result of a clinical trial), which informs the innovator about the project quality (safety or efficacy). Second, having learned the outcome, the innovator makes a decision following the rule in Equation (1).

4.2.1 Estimating the pull incentive effect

The GAIN Act is a policy change, and I estimate its impact in a difference-in-difference setting. First, I distinguish the treated drug candidates from the controls. The treated

group consists of QIDP-eligible drug candidates. The QIDP designation gives a project privileges throughout the clinical trials (see Section 2.2.2 for details), and typically clinical trials are designed for a specific indication (which is an important determinant of QIDP eligibility). Therefore, I assume that after 2012, for projects in clinical trials, the QIDP eligibility is equivalent to the revealed QIDP status, as the firm has no reason to delay applying for the incentive. However, the eligibility of preclinical projects, as well as those that ended before 2012, is not known. Given the clear eligibility criterion and the breadth of my data, I attempt to predict QIDP eligibility for this subset of projects.

Using the text descriptions of the projects in Pharmaprojects in fields *summary*, *drug disease*, *drug country*, *Phase 1*, *Phase 2*, and *Phase 3*, I construct a set of variables related to the QIDP eligibility criterion: *"antibacterial or antifungal drug for human use intended to treat serious or life-threatening infections"*, i.e., bacteria targeted, indications, and infection severity (details in Table 6). With this variable set, I train a random forest classifier on the sample of antibiotic projects that were in clinical trials after 2012, that is, the projects for which the QIDP status is known. Next, I use the model to predict the out-of-sample QIDP eligibility for the antibiotic projects that ended before 2012 or were active after 2012 but did not reach clinical trials. Figure 7 visualizes how the in-sample and out-of-sample groups are constructed.

The random forest classifier performs quite well for the in-sample groups while it still predicts that some out-of-sample projects are QIDP-eligible. Of the 108 QIDP projects active in clinical trials after 2012, the model correctly predicts 80 as eligible and predicts only 6 false positives (projects predicted to be eligible but not QIDP in the data).

Out-of-sample, I find 124 more projects that are eligible, including 69 that ended before 2012. While this number is low, compared to the total of almost 1300 projects in my antibiotic sample, it is in line with the reasoning that the incentives were introduced in response to the lack of innovation in the area.

The treated group consists of the QIDP projects and the projects predicted to be eligible. However, some of the projects that obtained the QIDP status started before 2012. As the fact that they obtained the QIDP status relies on them being active after

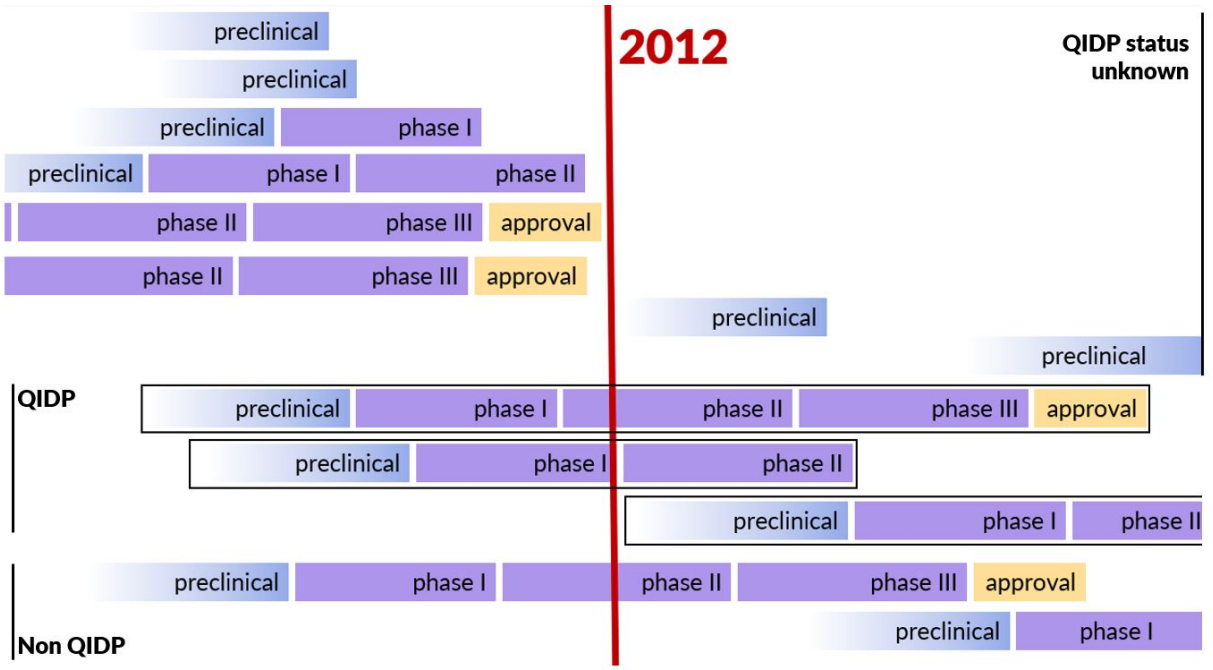
2012 and hence successful before that year, I drop all the observations related to these projects for the development phases that started before 2012.

The credibility of the estimate relies on the validity of the parallel trends assumption. In the current context, given the low numbers of observations in the years before the GAIN Act, the statistical tests that could support parallel trends would have very low power. Figure 8 plots the success rates of the QIDP projects and other antibiotics over time. While the outcomes of the QIDP projects are more volatile, they do not seem to exhibit a trend diverging from the other antibiotics.

4.2.2 Estimating the push incentive effect

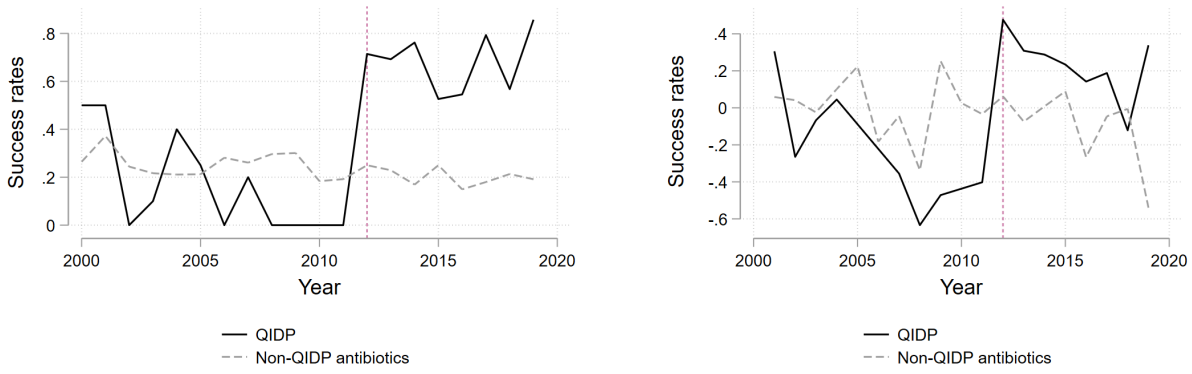
The push incentive, is a series of payments made to the project's developer, making its effect more difficult to capture. First, funding is not assigned randomly; it could be that firms with the best, most promising projects receive it or, conversely, that firms innovating in high-priority but also high-risk areas receive it. The received funding could be invested in very early stage research, and its effects might only become apparent in the future. For

Figure 7: *Projects used for predictions.*



Each horizontal line represents a project progressing through development phases over time (horizontal axis). The top projects, marked as 'QIDP status unknown' are the out-of-sample projects. 'QIDP' and 'Non QIDP' projects in the bottom half of the graph are the sample used to generate the predictions.

Figure 8: *Trends in phase success rates*



(a) *Raw values*

(b) *Net of the effect of controls (novelty, firm size, ownership change, bacteria targeted, phase fixed effects, project age) and the effect of the push incentive*

projects in clinical trials, i.e., those that have already been selected by the firm, additional financial flows can have an effect only through alleviating the firm's budget constraint.

At the end of a clinical trial, the firm receives a signal about the project's quality. If the signal is bad, the project is terminated regardless of the incentives it received, as it will not obtain regulatory approval. If the signal is good, the firm can decide to continue development or terminate. Assuming that the funders cannot predict the trial outcome, the funding flows received by a firm during a clinical trial do not suffer from the selection issue suggested above, and the incentive effect would work through affecting the decision rule in Equation (1), for example, by alleviating financial constraints the firm might be facing.

I focus on funding flows that the project's originator received while the project in a development phase and their effect on the outcome in this phase. As the funding received during this period, as well as the project's quality, could be affected by past funding, I control for the stock of push funding the firm obtained previously. I also add a dummy controlling for the characteristics that could differentiate firms that have ever received push funding from other firms and that thus affect their projects' success rates.

The push payments are matched at the firm level and only to the antibiotic projects of the firm. I do not match the funding to specific projects, as the funds are often disbursed for broader programs and it would not be possible to identify which molecules exactly are

part of those. Moreover, to the extent that the push incentive relaxes the firm's budget constraint, it can also have spillover effects on projects to which it is not directly aimed.

4.2.3 Framework

The descriptive analysis suggests that the introduction of the innovation incentives coincides with changes in the outcomes of antibiotic R&D projects. To obtain more precise and richer insight into these changes, I turn to a regression analysis using a linear probability model.

My dataset is at the project i - phase d level, meaning that I observe each project only once per development phase ($d \in \{0, 1, 2, 3, 4\}$: preclinical and the three phases of clinical trials and approval), and only in the development phases they enter. Success, Y_{id}, is equal to one each time a project is observed at a later phase at some point in the future and zero for projects discontinued after Phase d.

I measure the effect of the pull incentive in a difference-in-difference type of framework. Given that my prediction model finds pre-2012 projects that would have qualified for the QIDP designation, I can control for the characteristics of the QIDP projects that potentially affect their success rates, ($Pull_i$, equal to 1 for projects with the QIDP designation in the data and projects predicted to be QIDP eligible), and compare their outcomes with and without the policy or before and after 2012 ($Post2012_{id}$).

To control for the differences between the firms that receive push funding and those that do not, I introduce the dummy $Push_{id}^{any}$, which is equal to one for project i belonging to a firm that received any push funding by the time project i enters Phase d. To account for the possible long-term effects of past push incentives, I use the stock of push funding accumulated by the developer firm until the time project i entered Phase d, $Push_{id}^{stock}$. Finally, $Push_{id}^{flow}$ is the sum of the push funding received by the firm during the period project i was in Phase d.

My estimating equation takes the form:

$$\begin{aligned} Y_{id} =& \alpha_1 Post2012_{id} \times Pull_i + \alpha_2 Pull_i + \alpha_3 Post2012_{id} \\ &+ \beta_1 Push_{id}^{flow} + \beta_2 Push_{id}^{stock} + \beta_3 Push_{id}^{any} \\ &+ \gamma X_{id} + \delta_d + \varepsilon_{id} \end{aligned} \tag{2}$$

where the coefficient of interest for the pull and push incentives are α_1 and β_1, respectively.

The estimating equation also contains a set of controls, X_{id}: vintage (project start year) fixed effects and project age, whether the project has been exposed to the merger and acquisition activity of the firm developing it during Phase d, fixed effects for bacteria targeted by the project, novelty and total number of projects in the same technology prior to project d's starting year. δ_d is a vector of the phase fixed effects, and ε_{id} is the error term.

As incentivizing investment in novel projects is particularly important in the case of antibiotics, I estimate a version of Equation (2) augmented with the interaction of the two terms of interest with novelty, using a measure based on market launches in the project's technological class, as described in Section 3.3.

4.2.4 Results

Table 2 presents the estimates of selected coefficients from the estimation of Equation (2) separately for projects in preclinical development and in clinical trials, with and without the push dummy and with and without interactions of the main variables of interest: QIDP projects after 2012 and push flows.

The results suggest that the pull incentive had a positive effect on the eligible projects, both at the preclinical stage and in clinical trials. The preclinical success rate of QIDP-eligible projects after the introduction of the GAIN Act in 2012 increased by 31.5 percentage points overall, and this change was mainly driven by novel projects advancing into clinical development. At the clinical stage, trials of QIDP projects that started after the introduction of the pull incentive are 21 percentage points more successful than are the trials of the projects that were predicted to be eligible for the incentive but ended before 2012. In contrast to the preclinical stage, the pull incentive appears to have no effect for

novel projects in clinical trials.

The estimates of the effect of the push incentive are close to zero and mostly insignificant in the baseline specifications in Columns (3)-(4) of Table 2. The negative effect in Column (3) disappears after the inclusion of the push dummy in Column (4), suggesting that firms that received push funding engage in less successful or more risky projects on average.

Adding the interaction of novelty with the push flow, we can see in Column (8) that \$1M of financing received by the originator of a project in clinical trials increases the probability of success of this project by 1.3 percentage points, with no significantly different effect for novel projects, and the point estimate pointing at no effect at all for novel projects. Indeed, in Table 11, which presents the estimates on a sample restricted to decisions made until 2018, the effect on the interaction of push flows with novelty almost exactly cancels out the baseline effect of the push incentive.

Table 2: *Phase success probability*

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Preclinical	Preclinical	Trials	Trials	Preclinical	Trials
Pull	-0.0741	-0.0742	0.0031	-0.0004	-0.0049	-0.1335
	(0.0540)	(0.0540)	(0.1128)	(0.1127)	(0.0808)	(0.1257)
Pull × Post 2012	0.3146***	0.3151***	0.2021	0.2119*	0.0669	0.2322
	(0.1007)	(0.1000)	(0.1232)	(0.1230)	(0.1338)	(0.1507)
Pull × Post 2012 × New technology					0.5015***	-0.1353
					(0.1806)	(0.2322)
Push flow (\$M)	0.0023	0.0024	-0.0059**	0.0018	0.0123**	0.0133**
	(0.0031)	(0.0030)	(0.0027)	(0.0056)	(0.0062)	(0.0068)
Push flow × New technology					-0.0110	-0.0135
					(0.0069)	(0.0088)
Push stock (\$M)			-0.0022**	-0.0007		-0.0007
			(0.0009)	(0.0006)		(0.0006)
Push (dummy)		-0.0054		-0.1245***	-0.0058	-0.1270***
		(0.0418)		(0.0323)	(0.0442)	(0.0297)
Observations	27,794	27,794	9,894	9,894	27,794	9,894
R-squared	0.0448	0.0448	0.0860	0.0866	0.0455	0.0873

Standard errors in parentheses, clustered at the originator level. Linear probability model of phase success (advancement to a subsequent phase), including as controls: novelty, antibiotics dummy, new chemical entity dummy, firm size (number of projects in development), acquisition dummy, number of projects in technology, whether the project targets any of the bacteria listed by FDA, or ECDC, or WHO priority pathogens, phase fixed effects, vintage fixed effects, phase start year fixed effect, age of the project.

*** p<0.01, ** p<0.05, * p<0.1

As the push flows are measured in millions of dollars, comparing the coefficients on the push and pull incentive can provide an estimate of the value of the pull incentive to

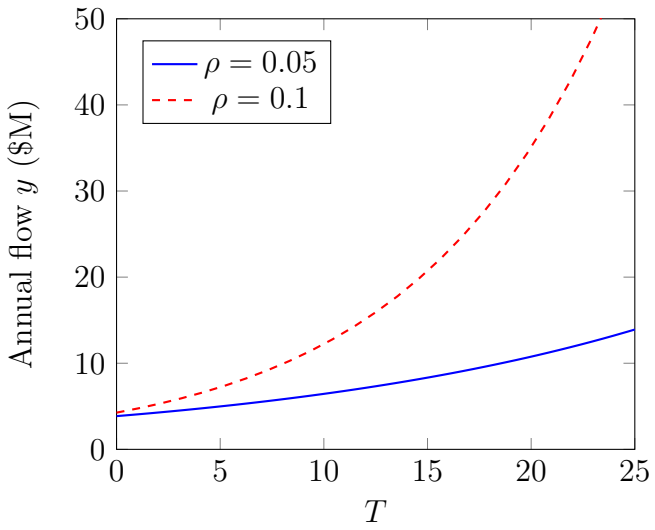
the innovators. The point estimate of the effect of the pull incentive is 17.46 times higher than the point estimate on the push incentive in clinical trials, suggesting that for the firms, on average, the present value of the 5 years exclusivity extension is approximately \$17.46M for projects in clinical trials.

Depending on the discount factor and on when the firm expects to benefit from the pull incentive, this net present value of \$17.46M can correspond to different annual flows. Figure 9 plots the estimates of these flows according to the function:

$$y = \frac{17.46}{\sum_{t=1}^{5}(1-\rho)^{t-1} \times (1-\rho)^{T}}$$

Where y is the annual profit flow over the 5 years of the QIDP exclusivity extension, ρ is the discount factor and T is the time frame at which the firm expects to start accruing the benefits from the pull incentive. For example, $T = 10$ could mean that the project is expected to be launched on the market in 5 years and will benefit for 5 years from the New Chemical Exclusivity before the QIDP exclusivity extension. In the $T = 10$ example, the NPV calculation suggests that the annual flow from the pull incentive would be approximately \$6.45M, or \$32.25M in total, assuming 5% discount factor; or \$12.23M (\$61.15M) with a 10% discount factor.

Figure 9: *Expected annual flows from the pull incentive*



Value of the pull incentive implied by the regression results as a function of the time horizon T when the firm is expecting to start perceiving the flows associated with the incentive, and the discount factor ρ

4.3 Dynamic model of pharmaceutical innovation

The reduced form evidence is informative on the changes in dynamics in the antibiotic pipeline after the introduction of the innovation incentives. However, it is only suggestive of how these changes will reflect on new market entries. Modeling the innovation process allows me to simulate a longer time period (then the period observed in the data) and provide a fuller picture of the effect of the incentives, including the most relevant outcome – market entries of new antibiotics. Including the effect of the incentives in the model will allow me to consider different incentive schemes in counterfactual simulations. Comparing the current incentive schemes to a scenario without incentives will provide an estimate of their effect. Moreover, experimenting with more generous incentives can indicate how much more should be invested to achieve an effect of a given magnitude.

The intuition behind the model is as follows: Through clinical trials, the originator learns about the value of molecule i by observing its outcomes and by observing the outcomes of other molecules using similar technologies that are at least as advanced in the trials as i. Table 12 provides some suggestive evidence for this process. When other similar molecules are observed to be discontinued by projects that are at least as advanced projects using the same technology, the molecule has a higher probability of being discontinued in the subsequent year.

In the case of antibiotics targeting resistant bacteria, competition on the product market can be assumed a less important factor, as effective management of antimicrobial resistance requires that multiple products be available and used. Therefore, the outcomes of similar molecules contribute more to the understanding of molecule i's type than to its expected market size.

At each point in time in the trials, i's originator observes the molecule's outcomes and decides whether to continue development or discontinue i. Continuing the trials is associated with the flow cost c_{id} which depends on the trial phase (later phases require more trial participants and higher costs).

The push and pull innovation incentives affect this process differently. The pull incentives simply increase the value of the molecule (conditional on successfully completing

development), while the push incentives effectively decrease the cost of conducting research at both the preclinical and clinical stages.

4.3.1 Setup

I use a continuous time dynamic discrete choice model developed in Arcidiacono et al. (2016) and Blevins (2014) and adapted to the pharmaceutical industry in Khmelnitskaya (2021) to model innovation in antibiotics. This choice has a number of particularly attractive features for this setting. First, the intervals between decisions are not fixed: the length of clinical trials may differ between drugs, and a particularly bad outcome can end a trial at any point in time. Second, the decisions of the different players interact, either because the expectations on the competitive environment on the product market change (a smaller concern for antibiotics) or because the outcomes of similar molecules affect the firm's estimate of its product's success probability. The continuous time setting allows me to incorporate the complex dynamics of the multi-agent setting in a computationally feasible way.

I consider an infinite horizon game in continuous time. Firms decide whether to invest in the development of molecules indexed by $i = 1, ..., N$. Each firm can own multiple molecules, but the decisions are independent across molecules once they enter development. I model only decisions for molecules already in development, and the entry into clinical trials, while allowed to affect the firm decision-making is kept exogenous in the model.

Throughout development, firms receive information about their molecules through the clinical trials outcomes (unobserved by the econometrician) and through the outcomes of other molecules that are technologically close. Firms can decide to either advance the project into the next phase ($j = 1$) or discontinue the molecule ($j = 0$). There are three phases of clinical trials indexed by d, and profits are realized only after completing Phase 3 and obtaining regulatory approval. State variable s_k summarizes the state of the R&D pipeline in the i's technological class in a vector of size 4 containing the number of molecules at each development phase, i.e. the clinical trials $d = \{1, 2, 3\}$ and approval stage $d = 4$: $s_k = (s_k^1, s_k^2, s_k^3, s_k^4)$ for $k \in 1, 2, ..., K$. Assuming a maximum number of

molecules at a given development stage at the same time (limited, e.g., by the finite number of hospitals and amount of human capital available to organize clinical trials), the state space is finite.

The definition of the state vector is important, as next to the phase and incentives, it will define the molecule's value. The current definition of the state vector is restrictive, but it captures the fact that more projects being developed concurrently in a technological class means more information about this technology for the innovators. Additionally, it allows the quality of this information to vary according to the distribution of the projects across development phases.

In each clinical trial phase, the decisions for molecule i in state k occur at random times according to a Poisson process with rate λ_d (phase-specific move arrival rate). Function $l(i,j,k)$ gives the new state conditional on the owner of project i taking action j in state k. σ_{idjk} denotes the conditional choice probability of project i in Phase d making a decision j in state k.

The value function of molecule i in Phase d in state k, for a small increment of time, h, for $d = \{1,2,3\}$ (clinical trials) follows:

$$V_{idk} = \frac{1}{1+\rho h}\left[-c_{id}h + \underbrace{\sum_{d'=1,2,3} q_{d'} V_{idk'_{d'}}}_{\substack{\text{nature}\\ \text{arrival}}} + \underbrace{\sum_{m\neq i}\lambda_{d(m)}h\sum_{j}\sigma_{md(m)jk}V_{id,l(m,j,k)}}_{m\neq i \text{ moves}} + \right. \tag{3}$$

$$\left. \underbrace{\lambda_d h \text{Emax}(V_{i,d+1,l(i,1,k)} + \varepsilon_{idk}, 0)}_{i \text{ moves}} + \underbrace{\left(1 - \sum_{d'=1,2,3} q_{d'}h - \sum_{m\neq i}\lambda_{d(m)}h - \lambda_d h\right)V_{idk}}_{\text{no change}} + o(h)\right]$$

Analyzing the value function term by term:

1. At each point in time, to continue development, the firm must pay the phase-specific flow cost c_{id}. Flow cost c_{id} is a function of molecule i's current Phase d, and current push incentive flows $push_{id}$:

$$c_{id} = c_d - \beta_d^{push} push_{id}$$

2. With probability $q_{d'}$, there is an exogenous arrival of a new project into development in phase d', pushing the industry into state $k'_{d'}$ (state k with one more molecule in Phase d')

3. Each one of the other projects, $m \neq i$, can move with probability $\lambda_{d(m)}$, where $d(m)$ denotes the phase of project m

4. With probability λ_d, i will move. The firm can either advance i into the next Phase $d+1$ or discontinue the project. A payoff is realized only after approval ($d = 4$) and depends on the pull incentive $QIDP_i$.

5. With probability $\left(1 - \sum_{d'=1,2,3} q_{d'}h - \sum_{m \neq i} \lambda_{d(m)}h - \lambda_d h\right)$, there is no change and the industry remains in state k

As $h \to 0$:

$$V_{idk} = \frac{-c_{id} + \sum_{d'=1,2,3} q_{d'} V_{idk'_{d'}} + \sum_{m \neq i} \lambda_{d(m)} \sum_j \sigma_{md(m)jk} V_{id,l(m,j,k)} + \lambda_d \text{Emax}(V_{i,d+1,l(i,1,k)} + \varepsilon_{idk}, 0)}{\rho + \sum_{d'=1,2,3} q_{d'} + \sum_{m \neq i} \lambda_{d(m)} + \lambda_d} \tag{4}$$

For $d = 4$ (approval):

$$V_{i4} = \frac{-c_4 + p\pi_i}{\rho + 1} \tag{5}$$

As approval probability p is exogenous, and if given approval, the firm definitely enters the market and gains profit π_i. The expected profit depends on whether the project benefits from the QIDP designation $QIDP_i = 1$:

$$\pi_i = \begin{cases} \pi^{base} & \text{if } QIDP_i = 0 \\ \pi^{base} + \pi^{QIDP} & \text{if } QIDP_i = 1 \end{cases}$$

My model deviates from that of Khmelnitskaya (2021) in a number of ways. I abstract from product-market competition, which is less of a concern in this market, and do not model strategic considerations in development decisions (although, if relevant, they will be captured by the changes across states k), in favor of including the innovation incentives in the model. I focus on one drug class, antibiotics, and use a more granular, technology-specific, state variable, which contains more information about the project's probability of success.

4.3.2 Estimation

The estimation procedure follows in two steps. First, I estimate the move arrival rates λ_d (collected in vector $\lambda = [\lambda_1, \lambda_2, \lambda_3]$); exogenous arrival probabilities, $q = [q_1, q_2, q_3]$; and conditional choice probabilities, σ_{idjk} from the data. In the second step, fixing λ, q and σ_{idjk}, I estimate the model primitives $\theta = \{\beta_2^{push}, \beta_3^{push}, \pi^{base}, \pi^{QIDP}\}$.

Step 1

λ and q_d are estimated directly from the R&D data. The move arrival rates are the inverse of the average phase duration in years. The exogenous arrival rate is the mean ratio of the exogenous arrivals per year to the number of active technologies.

The conditional choice probabilities σ_{idjk} are estimated with a logit of a polynomial function of state variables using the decisions observed in the data, yielding $\tilde{\sigma}_{idjk}$.

To obtain point estimates of the effect of the push (pull) incentive on cost (profit), I set the Phase 2 and 3 constants in the cost function, c_2 and c_3, and the approval cost c_4, to the values estimated by Sertkaya et al. (2014) for antimicrobials, with 30% overhead, following Outterson (2021).

Step 2

The second step of the estimation procedure relies on rewriting the V terms as functions of the conditional choice probabilities following Hotz and Miller (1993).

The firm's decision at the end of a clinical trial is the solution to the term $\text{Emax}(V_{i,d+1,l(i,1,k)} + \varepsilon_{idk}, 0)$ in the value function. Section B.2 of the Appendix presents the linear representation of the value function, showing that for each phase, the vector of values V_{id} across all states k can be expressed in terms of the model parameters and the $\text{Emax}(V_{i,d+1,l(i,1,k)} + \varepsilon_{idk}, 0)$ terms. Assume that ε_{idk} is i.i.d. Type I extreme value. The scaling parameter of ε_{idk}'s distribution will affect the degree to which the random shock affects the decisions and can be estimated with the other primitives. Normalizing the scaling parameter to 1, I can replace the Emax term by its closed form expression:

$$\text{Emax}(V_{i,d+1,l(i,1,k)} + \varepsilon_{idk}, 0) = \ln(1 + e^{V_{i,d+1,l(i,1,k)}}) + \nu \tag{6}$$

where ν is the Euler constant.

The probability of advancing from Phase d in state k ($j = 1$) is then:

$$\begin{aligned} \sigma_{id1k} &= Pr(V_{i,d+1,l(i,1,k)} + \varepsilon_{idk} > 0) \\ &= \frac{e^{V_{i,d+1,l(i,1,k)}}}{1 + e^{V_{i,d+1,l(i,1,k)}}} \end{aligned} \tag{7}$$

Inverting, I obtain:

$$V_{i,d+1,l(i,1,k)} = \ln(\frac{\sigma_{id1k}}{1 - \sigma_{id1k}}) \tag{8}$$

Obtaining:

$$\text{Emax}(V_{i,d+1,l(i,1,k)} + \varepsilon_{idk}, 0) = \ln(1 + \frac{1}{1 - \sigma_{id1k}}) + \nu \tag{9}$$

Using Equation (8) and the $\tilde{\sigma}_{idjk}$ estimates, I can replace the V_{idk}'s on the RHS of Equation (4) with appropriate values, rewriting V_{idk} as a function of the model primitives and parameters estimated in Step 1. Then, I can recover the structural conditional choice probabilities (CCP), $\tilde{\sigma}_{idjk}(\theta, q, \lambda, \tilde{\sigma})$ using Equation (7):

$$\tilde{\sigma}_{id1k}(\theta, q, \lambda, \tilde{\sigma}) = \frac{e^{V_{i,d+1,l(i,1,k)}}}{1 + e^{V_{i,d+1,l(i,1,k)}}}$$

For the set of decisions observed in the data, I can write then the likelihood

$$l_{id}(\theta, q, \lambda, \tilde{\sigma}) = \frac{\mathbb{1}\{j(i,d,k) = 0\} + \mathbb{1}\{j(i,d,k) = 1\} \times e^{V_{i,d+1,l(i,1,k)}}}{1 + e^{V_{i,d+1,l(i,1,k)}}}$$

where $j(i,d,k)$ is the decision observed in the data made for project i in phase d in state k. Aggregating over decisions observed in the data, I can write the pseudo log-likelihood function:

$$L(\theta, q, \lambda, \tilde{\sigma}) = \sum_{i,d} \ln\left(l_{id}(\theta, q, \lambda, \tilde{\sigma})\right)$$

Section B.4 presents details on the construction of the value functions used in estimation.

4.3.3 Results

Table 3 presents the estimation results. With the standard errors calculated from 400 bootstrap samples, the estimates are precisely estimated.

The effects of the push incentive are very close in magnitude to the average annual costs of clinical trials captured by the constants in the cost function ($8.5M vs the $10.94M constant in Phase 2, c_2, and $18.5M vs $27.26M in Phase 3, c_3), suggesting that the benefits of the push incentive are greater than the value of the transfer (on average approximately $3.5M) and significantly reduce the role of cost considerations in the firm's decision problem, especially in Phase 1.

The profit estimates confirm that antibiotics have a very low expected profit, with just under $300M in discounted lifetime profits at the baseline. The results suggest that a QIDP designation gives the firm very little additional benefit, consistent with the critiques expressed in the literature (Darrow and Kesselheim, 2020). The estimate of π^{QIDP} is however a lower bound on the incentive effect if the QIDP-eligible projects are less profitable than the rest of the antibiotics.

The estimation results suggest a large effect of the push incentives and a modest effect of the pull incentive. In the next section, I simulate a longer time period and different counterfactual scenarios to obtain a more nuanced view of the incentive effects on the progress of projects throughout the development phases.

Table 3: *Results from Maximum Likelihood estimation of the structural model (in $10M)*

	Estimate	Standard Error
Push, Phase 2 (β_2^{push})	0.850090	0.0829929
Push, Phase 3 (β_3^{push})	1.850127	0.0918365
Base profit (π^{base})	29.813637	0.0100535
Profit QIDP (π^{QIDP})	1.138357	0.0481879

Standard errors obtained by bootstrap of the ML estimate of the dynamic model with 400 replications

4.3.4 Counterfactuals

The counterfactual simulations are guided by Equation (3). Starting with a sample of projects from the beginning of 2012, for each technological class, I allow one change to

occur daily and simulate 20 years of data.

In the simulations, I use the parameters used in the estimation, the estimated coefficients and the values of projects at different states obtained by solving the model through value function iteration. The firm's decision is guided by these values (and a random shock), which depend on the project incentives.

Table 4 compares outcomes observed in the data to those from 100 simulations of the first 10 years at the equilibrium (baseline). The current version of the model underestimates the success rates and numbers of projects in the pipeline. The low success rates can be explained by the fact that in the model, firms are making decisions mostly following cost and profit considerations (and the low estimated profit explains their unwillingness to continue development), while in reality some firms might intrinsically value having projects in clinical development. This mismatch motivates including more observed firm heterogeneity in future versions of the model, and better accounting for unobserved heterogeneity by estimating the scaling parameter of the ε_{idk}. The underestimated number of active projects is on the one hand the direct effect of low success rates (especially for projects at higher phases of development), but it also suggests that using the mean rate of the observed exogenous entry might not be enough and it could be possible to allow it to evolve over time albeit at substantial computational cost.

Even if the simulation results do not exactly match the data, it is still useful to compare different counterfactual scenarios to see how, in broad terms, the incentives affect firm decision making and the antibiotic pipeline. In the simulations I can manipulate three (sets of) parameters that affect both the incentives and the outcomes observed: the pull incentive, the probability of receiving push incentive, and exogenous entries.

While entry into development is not modeled, the number and type of projects present in the pipeline affect the outcomes and might be affected by the incentives. To this end, I separately model the entry of the QIDP and other projects using the mean values from the period after 2012 (i.e., with the incentive present) and attribute the entry rates of projects exposed to the two incentives to the incentive or not, depending on the scenario. To estimate and solve the model, I could rely only on the technological classes already

Table 4: *Counterfactual results compared to data*

	Data	Model
	Phase 1	
N	259	141.56
Success rate	0.598	0.098
	Phase 2	
N	202	108.92
Success rate	0.495	0.052
	Phase 3	
N	116	49.03
Success rate	0.560	0.637

Numbers of decisions observed and success rates (conditional on taking a decision in the given phase). In the data – column Data; and means from 100 simulations of a 10 year period at the equilibrium – column Model.

present on the market. However, bringing the simulation results for the whole class of antibiotics closer to reality requires modeling the entry of new technologies. I follow the same approach as for entries into the development of projects in existing classes and use the observed rates for QIDP and non-QIDP projects.

The push incentive is decided when the firm makes a decision. All projects have the same probability of receiving the push incentive, following the observed values for the first 10 years of the simulation and fixed at the 2021 level after that period.

I simulate the data at the baseline (incentives as observed) and counterfactual scenarios by removing the incentives and reducing exogenous entry as well as in one of the counterfactuals without incentives.

Table 5 presents the mean numbers of decisions made and the mean success rates (share of projects advancing to the next phase) per phase from 100 simulations of each of the scenarios. It also shows the number of QIDP projects arriving in approval over the simulated 20 years. Column (1) is the baseline scenario where the parameters associated with the incentives are kept at the levels observed between 2012 and 2021. Figure 10 plots relative changes in the number of market entries over the 20 simulated years by counterfactual scenario.

The first comparison can be made between the baseline and a scenario with no incentives at all, to estimate the effect of current incentive scheme. To the extent that incentives can affect entry, the comparison with no incentives and no entry of incentivized projects at all (Column (2)) will give the upper bound of the effect, and the comparison with a scenario where the incentives are removed but the entry rates remain unchanged (Column (3)), the lower bound of the effect. The effect on success rates in Phases 1 and 2 is almost the same in both cases, and they drop to values very close to 0. Success rates in Phase 3 are also similar in columns (2) and (3) and close to the baseline value, suggesting the current incentives do not affect decision making at the later stages of development. Due to the lower success rates in earlier phases and, to a lesser extent, less entry in Column (2), there are much fewer projects in Phase 3, which will automatically translate into fewer projects entering the market. The model predicts that the current incentive scheme will increase the number of market entries by 20-25% overall, and, at the upper bound, double the number of QIDP products entering the market.

The model allows me also to isolate the effect of the incentives, by removing them one by one while holding everything else constant. In Column (4) the pull incentive is removed and in Column (5), the push incentive. In both cases the entry rates remain unchanged, the estimates providing the lower bound of the effect of the relevant incentive. Comparing Columns (1) and (4), the effect of the pull incentive turns out to be marginal, with a slightly lower number of projects in Phase 3 and a 2 pp. lower success rate when the incentive is removed. Removing the push incentive, on the other hand, has much stronger effect across all the outcomes except the Phase 3 success rate. Figure 10 suggests that the impact of the push funding in the current incentive scheme is twice as large as the impact of the pull incentive.

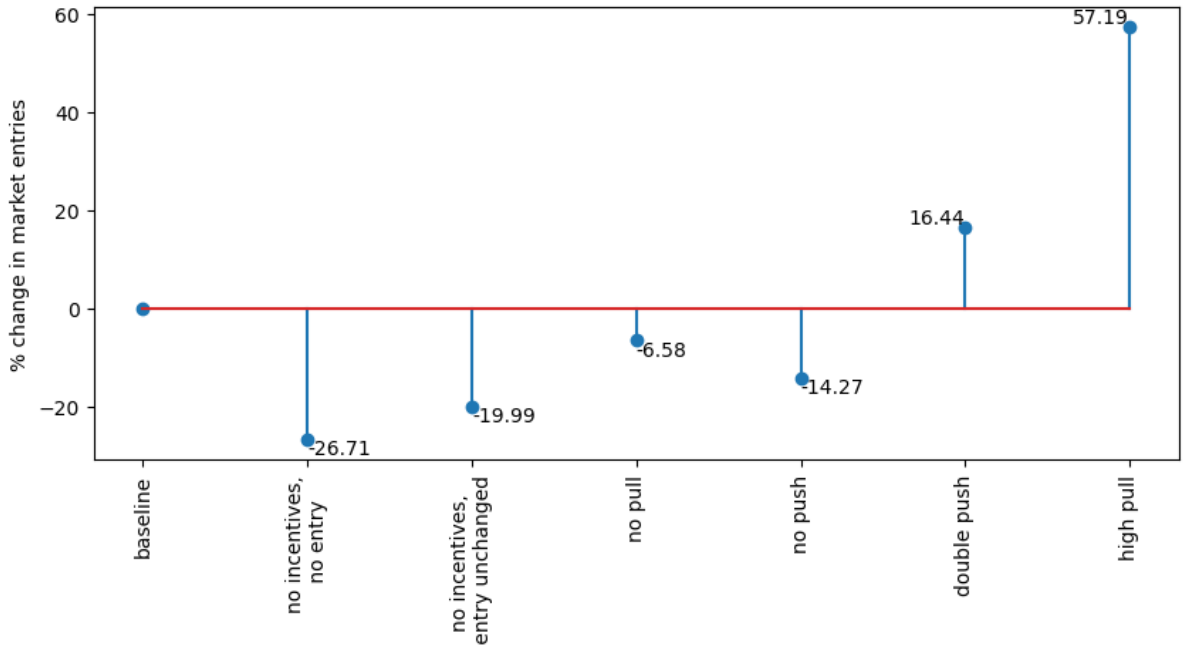
The last two columns of Table 5 consider scenarios with much higher incentives. Column (6) presents the results from a simulation in which the push incentive was doubled (at the extensive margin, by doubling the probabilities of receiving the subsidy). The resulting Phase 1 and 2 success rates are also doubled, and the number of projects arriving at the decision point in the last phase increases by approximately 13%.

Table 5: *Counterfactual results*

	Scenario						
	(1) Baseline	(2) None (1)	(3) None (2)	(4) No pull	(5) No push	(6) Double push	(7) High pull
Push	✓			✓		×2	✓
Pull	✓				✓	✓	$100M
Entry	✓		✓	✓	✓	✓	✓
Phase 1							
N	287.42	210.92	284.60	290.10	285.83	295.39	297.40
Success rate	0.128	0.029	0.033	0.124	0.033	0.214	0.142
Phase 2							
N	186.32	147.68	165.25	186.06	164.70	207.42	193.46
Success rate	0.077	0.008	0.007	0.073	0.007	0.157	0.237
Phase 3							
N	90.36	70.23	75.88	87.70	77.21	102.71	118.33
Success rate	0.614	0.579	0.585	0.591	0.616	0.629	0.737
QIDP in approval	11.12	5.11	7.26	7.59	9.19	12.63	42.47

Table presents the mean number of decisions observed and mean success rates in each development phase from 100 simulations (per scenario) of 20 years of daily events in the antibiotic R&D pipeline, starting from the set of projects active on January 1st 2012 and allowing for entry of new projects into active technological classes and for entry of new classes into development. In column (2) the entry of QIDP projects and projects that received the push incentive in the past is suppressed. In column (6) the number of projects receiving the push incentive is doubled and in column (7) the value of the pull incentive is increased to $100M (from the estimated $11.4M)

Figure 10: *Relative change in market entries by counterfactual scenario*



In Column (7), the pull incentive is increased to $100M, a 9-fold increase from the estimated effect. With such a substantial increase, there is an effect on early stage success rates, as well as an increase in Phase 3 success rate, which translates into a 57% increase

in market entries, mostly driven by QIDP projects.

4.3.5 Discussion

The counterfactual results suggest that over the course of 20 years from its beginning, the current incentive scheme can be expected to be successful at increasing the number of market entries of new antibiotics. The pull incentive in its current form has almost no impact on the earlier stages other than potentially attracting more projects into development. This fact suggests that the bulk of its effect might work through firms' repurposing of projects already in development to fit the eligibility criteria. Depending on the advancement of these projects, the pull incentive might affect very few novel projects. To develop new and novel antibiotics that need to progress through all the clinical trial phases, push incentives are indispensable, and pull incentive can play a role only if it is substantially increased.

While the effect of the pull incentive in the counterfactuals appears much smaller than the point estimate obtained in the reduced form analysis, it should be noted that the reduced form estimates are noisy. Additionally, as shown in Table 7, decomposing the difference-in-differences results by development phase shows that the effect of the pull incentive is the most prominent in Phase 3 and much more ambiguous for the earlier phases.

5 Conclusion

This paper analyzes the performance of innovation incentives in an underserved pharmaceutical class, antibiotics. The high clinical need for new antibiotics coincides with a small market size because of a biological process – antimicrobial resistance. New therapies are needed to treat resistant infections, but to limit the emergence of resistant bacteria, new antibiotic have to be reserved only for the cases where they are strictly necessary.

Two incentive types have been introduced since 2012 to make innovation in antibiotics more attractive: research subsidies, or push incentives, and an extension of the market exclusivity period, constituting a form of a pull incentive.

Estimating the effect of these incentives on the main outcome of interest, new prod-

uct launches, is challenging due to the relatively short time that has passed since their introduction relative to the long innovation process in this industry. This paper takes two approaches to offer the broadest possible view on this issue.

First, in a reduced-form setting, I show that the incentives have been successful at invigorating the antibiotic pipeline by focusing on an intermediary outcome: clinical trial success. The pull incentive appears to have encouraged firms to invest more in the eligible projects, and the push incentives have reduced the financial constraints of the innovators. On the other hand, neither incentive has had any effect on novel projects.

Second, I set up a structural model to simulate the effects of the incentives that will be realized in the future and consider counterfactual incentive schemes. I find that the push incentive plays an important role in helping projects advance through Phases 1 and 2 of clinical trials, while the effect of the pull incentive is only pronounced in Phase 3. In comparison to a counterfactual of no incentive, the current situation has produced large effects at every development stage, especially if the incentives attract new projects into development.

The results, while informative, show the importance of incorporating firm heterogeneity in the model in the future versions of the paper. The analysis could also be made richer by including novelty explicitly as a characteristic of a technological class. This paper contributes to the nascent literature on the supply-side economics of antibiotics with many important avenues for future research, such as understanding the drivers of entries and exists of innovators in antibiotics, treated as exogenous in this paper.

References

Acemoglu, D. and J. Linn (2004). Market size in innovation: theory and evidence from the pharmaceutical industry. *The Quarterly Journal of Economics 119*(3), 1049–1090.

Adda, J. (2020). Preventing the spread of antibiotic resistance. *AEA Papers and Proceedings 110*, 255–59.

Ambrose, P. G. (2011). Antibiotic bill doesn't GAIN enough ground. *Nature Medicine 17*(7), 772.

Arcidiacono, P., P. Bayer, J. R. Blevins, and P. B. Ellickson (2016). Estimation of dynamic discrete choice models in continuous time with an application to retail competition. *Review of Economic Studies 83*(3), 889–931.

Blevins, J. R. (2014). Identification and Estimation of Continuous Time Dynamic Games. (April 2014), 1–35.

Cassini, A., L. D. Högberg, D. Plachouras, A. Quattrocchi, A. Hoxha, G. S. Simonsen, M. Colomb-Cotinat, M. E. Kretzschmar, B. Devleesschauwer, M. Cecchini, D. A. Ouakrim, T. C. Oliveira, M. J. Struelens, C. Suetens, D. L. Monnet, R. Strauss, K. Mertens, T. Struyf, B. Catry, K. Latour, I. N. Ivanov, E. G. Dobreva, A. Tambic Andraševic, S. Soprek, A. Budimir, N. Paphitou, H. Žemlicková, S. Schytte Olsen, U. Wolff Sönksen, P. Märtin, M. Ivanova, O. Lyytikäinen, J. Jalava, B. Coignard, T. Eckmanns, M. Abu Sin, S. Haller, G. L. Daikos, A. Gikas, S. Tsiodras, F. Kontopidou, Á. Tóth, Á. Hajdu, Ó. Guólaugsson, K. G. Kristinsson, S. Murchan, K. Burns, P. Pezzotti, C. Gagliotti, U. Dumpis, A. Liuimiene, M. Perrin, M. A. Borg, S. C. de Greeff, J. C. Monen, M. B. Koek, P. Elstrøm, D. Zabicka, A. Deptula, W. Hryniewicz, M. Caniça, P. J. Nogueira, P. A. Fernandes, V. Manageiro, G. A. Popescu, R. I. Serban, E. Schréterová, S. Litvová, M. Štefkovicová, J. Kolman, I. Klavs, A. Korošec, B. Aracil, A. Asensio, M. Pérez-Vázquez, H. Billström, S. Larsson, J. S. Reilly, A. Johnson, and S. Hopkins (2019). Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: a population-level modelling analysis. *The Lancet Infectious Diseases 19*(1), 56–66.

Cunningham, C., F. Ederer, and S. Ma (2021). Killer acquisitions. *Journal of Political Economy 129*(3), 649–702.

Darrow, J. J. and A. S. Kesselheim (2020). Incentivizing antibiotic development: Why isn't the generating antibiotic incentives now (gain) act working? *Open Forum Infectious Diseases 7*(1), 10–12.

Dubois, P., O. De Mouzon, F. Scott-Morton, and P. Seabright (2015). Market size and pharmaceutical innovation. *The RAND Journal of Economics 46*(4), 844–871.

Grabowski, H. and M. Kyle (2008). 11. mergers and alliances in pharmaceuticals: effects on innovation and r&d productivity. *The economics of corporate governance and mergers 262*.

Grace, C. and M. Kyle (2009). Comparative advantages of push and pull incentives for technology development: lessons for neglected disease technology development. *Global Forum Update on Research for Health 6*, 147–151.

Haucap, J., A. Rasch, and J. Stiebale (2019). How mergers affect innovation: Theory and evidence. *International Journal of Industrial Organization 63*, 283–325.

Howell, S. T. (2017). Financing Innovation: Evidence from R&D Grants. *The American Economic Review 107*(4), 967–1004.

Huang, S., M. A. Ribers, and H. Ullrich (2022). Assessing the value of data for prediction policies: The case of antibiotic prescribing. *Economics Letters 213*, 110360.

Huang, S. and H. Ullrich (2021). Physician effects in antibiotic prescribing: Evidence from physician exits.

Khmelnitskaya, E. (2021). Competition and Attrition in Drug Development. *Working Paper*.

Krieger, J., D. Li, and D. Papanikolaou (2022). Missing Novelty in Drug Development. *The Review of Financial Studies 35*(2), 636–679.

Krieger, J., X. Li, and R. T. Thakor (2021). Find and Replace: R&D Investment Following the Erosion of Existing Products. *Harvard Business School Entrepreneurial Management Working Paper 19-058*.

McCutchen Jr, W. W. (1993). Estimating the impact of the r&d tax credit on strategic groups in the pharmaceutical industry. *Research Policy 22*(4), 337–351.

Murray, C. J., K. S. Ikuta, F. Sharara, L. Swetschinski, G. Robles Aguilar, A. Gray, C. Han, C. Bisignano, P. Rao, E. Wool, S. C. Johnson, A. J. Browne, M. G. Chipeta, F. Fell, S. Hackett, G. Haines-Woodhouse, B. H. Kashef Hamadani, E. A. Kumaran, B. McManigal, R. Agarwal, S. Akech, S. Albertson, J. Amuasi, J. Andrews, A. Aravkin, E. Ashley, F. Bailey, S. Baker, B. Basnyat, A. Bekker, R. Bender, A. Bethou, J. Bielicki, S. Boonkasidecha, J. Bukosia, C. Carvalheiro, C. Castañeda-Orjuela, V. Chansamouth, S. Chaurasia, S. Chiurchiù, F. Chowdhury, A. J. Cook, B. Cooper, T. R. Cressey, E. Criollo-Mora, M. Cunningham, S. Darboe, N. P. Day, M. De Luca, K. Dokova, A. Dramowski, S. J. Dunachie, T. Eckmanns, D. Eibach, A. Emami, N. Feasey, N. Fisher-Pearson, K. Forrest, D. Garrett, P. Gastmeier, A. Z. Giref, R. C. Greer, V. Gupta, S. Haller, A. Haselbeck, S. I. Hay, M. Holm, S. Hopkins, K. C. Iregbu, J. Jacobs, D. Jarovsky, F. Javanmardi, M. Khorana, N. Kissoon, E. Kobeissi, T. Kostyanev, F. Krapp, R. Krumkamp, A. Kumar, H. H. Kyu, C. Lim, D. Limmathurotsakul, M. J. Loftus, M. Lunn, J. Ma, N. Mturi, T. Munera-Huertas, P. Musicha, M. M. Mussi-Pinhata, T. Nakamura, R. Nanavati, S. Nangia, P. Newton, C. Ngoun, A. Novotney, D. Nwakanma, C. W. Obiero, A. Olivas-Martinez, P. Olliaro, E. Ooko, E. Ortiz-Brizuela, A. Y. Peleg, C. Perrone, N. Plakkal, A. Ponce-de Leon, M. Raad, T. Ramdin, A. Riddell, T. Roberts, J. V. Robotham, A. Roca, K. E. Rudd, N. Russell, J. Schnall, J. A. G. Scott, M. Shivamallappa, J. Sifuentes-Osornio, N. Steenkeste, A. J. Stewardson, T. Stoeva, N. Tasak, A. Thaiprakong, G. Thwaites, C. Turner, P. Turner, H. R. van Doorn, S. Velaphi, A. Vongpradith, H. Vu, T. Walsh, S. Waner, T. Wangrangsimakul, T. Wozniak, P. Zheng, B. Sartorius, A. D. Lopez, A. Stergachis, C. Moore, C. Dolecek, and M. Naghavi (2022). Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet 399*(10325), 629–655.

Ornaghi, C. (2009). Mergers and innovation in big pharma. *International Journal of Industrial Organization 27*(1), 70–79.

Outterson, K. (2013). Newsletter Vol . 30 No . 1 All Pain , No GAIN : Need for Prudent Antimicrobial Use Provisions to Complement the GAIN Act. *30*(1), 2012–2014.

Outterson, K. (2021). Estimating The Appropriate Size Of Global Pull Incentives For Antibacterial Medicines. *Health Affairs (Project Hope) 40*(11), 1758–1765.

Outterson, K., J. H. Powers, G. W. Daniel, and M. B. McClellan (2015). Repairing the broken market for antibiotic innovation. *Health Affairs 34*(2), 277–285.

Rao, A. (2020). Strategic research and development investment decisions in the pharmaceutical industry. *Marketing Science 39*(3), 564–586.

Ribers, M. A. and H. Ullrich (2022). Machine learning and physician prescribing: a path to reduced antibiotic use.

Sarpatwari, A., R. F. Beall, A. Abdurrob, M. He, and A. S. Kesselheim (2018). Evaluating the impact of the Orphan Drug Act's seven-year market exclusivity period. *Health Affairs 37*(5), 732–737.

Sertkaya, A., J. Eyraud, A. Birkenbach, C. Franz, N. Ackerley, V. Overton, and K. Outterson (2014). Analytical Framework for Examining the Value of Antibacterial Products. *11*.

Simpkin, V. L., M. J. Renwick, R. Kelly, and E. Mossialos (2017). Incentivising innovation in antibiotic drug discovery and development: Progress, challenges and next steps. *Journal of Antibiotics 70*(12), 1087–1096.

WHO (2021). *2020 Antibacterial Agents in Clinical and Preclinical Development*.

Yin, W. (2008). Market incentives and pharmaceutical innovation. *Journal of Health Economics 27*(4), 1060–1077.

A Reduced form

A.1 Details on estimation

Table 6: Variables used in the QIDP eligibility prediction (variable name is the keyword unless otherwise stated)

Variable	Description	Field
Bacteria		
escherichiacoli		
acinetobacter		
aspergillus		
streptococcuspyogenes		
clostridiumdifficile		
enterobacter		
campylobacter	Bacteria names with spelling variations	Summary
gonorrhoea		
candida		
cryptococcus		
enterococcus		
helicobacterpylori		
coccidioides		
eskape	Targets any of the ESKAPE pathogens (Enterococcus, MRSA, K. pneumoniae, Acinetobacter, Pseudomonas aeruginosa, Enterobacter)	
anyecdc	Targets any of the bacteria surveilled by ECDC	
Indications		
pneumonia		
cystic fibrosis		
systemic		
surgery	Includes spelling variations	
catheter		Summary, Drug Disease
resistant		
bacterial vaginosis		
tb	Tuberculosis, includes spelling variations	
bloodstream		
skin		Summary
intra-abdominal		Summary
urinary		Summary
Severity		
uncomplicated		Summary
complicated		Summary, Drug Disease
severe		Summary, Drug Disease
hospital		Summary, Drug Disease
threat		Summary
Other		
formulation		Summary
phage		Summary
combination		Summary
probiotic		Summary
prevent	*Prevent* or *Prophylaxis* with spelling variations	Summary
usa		Drug Country
ctdotgov	Includes a link to clinicaltrials.gov	Phase I, Phase II, Phase III
topical		Summary
derivative		Summary
peptide		Summary

A.2 Additional results

Table 7: Results (by development phase)

VARIABLES	(1) Phase I	(2) Phase II	(3) Phase III	(4) Phase I	(5) Phase II	(6) Phase III
Pull	0.1586	0.0362	-0.5028***	0.1557	-0.2481**	-0.4399***
	(0.2184)	(0.1616)	(0.1167)	(0.2638)	(0.1038)	(0.1226)
Pull × Post 2012	-0.0208	0.3101*	0.6074***	-0.1613	0.7935***	0.3286*
	(0.2346)	(0.1790)	(0.1494)	(0.3118)	(0.0981)	(0.1975)
Pull × Post 2012 × New technology				0.0734	-0.7396***	0.5359**
				(0.3695)	(0.2642)	(0.2603)
Push flow ($M)	0.0243***	-0.0042	-0.0100	0.0338***	-0.0078	-0.0556***
	(0.0045)	(0.0026)	(0.0072)	(0.0057)	(0.0067)	(0.0100)
Push flow × New technology				-0.0151**	0.0036	0.1064***
				(0.0071)	(0.0078)	(0.0128)
Push stock ($M)	-0.0007	-0.0008	0.0011	-0.0007	-0.0009	0.0080
	(0.0010)	(0.0020)	(0.0105)	(0.0010)	(0.0021)	(0.0105)
Push (dummy)	-0.0924**	-0.1972***	-0.1456***	-0.0866**	-0.1869***	-0.1652***
	(0.0368)	(0.0318)	(0.0379)	(0.0373)	(0.0357)	(0.0387)
Observations	4,634	3,363	1,897	4,634	3,363	1,897
R-squared	0.0591	0.0447	0.1118	0.0612	0.0458	0.1143

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 8: Results (antibiotics subsample)

VARIABLES	(1) Preclinical	(2) Preclinical	(3) Trials	(4) Trials	(5) Preclinical	(6) Trials
Pull	-0.0722	-0.0722	-0.1188	-0.1120	0.0087	-0.0299
	(0.0532)	(0.0532)	(0.1400)	(0.1408)	(0.0803)	(0.2210)
Pull × Post 2012	0.3078***	0.3077***	0.3810**	0.3661**	0.0635	0.2856
	(0.0995)	(0.0991)	(0.1673)	(0.1698)	(0.1296)	(0.2662)
Pull × Post 2012 × New technology					0.5205***	0.1634
					(0.1820)	(0.3116)
Push flow ($M)	0.0005	0.0009	0.0328**	0.0206	0.0149***	0.0254
	(0.0020)	(0.0019)	(0.0131)	(0.0139)	(0.0035)	(0.0163)
Push flow × New technology					-0.0153***	-0.0120
					(0.0033)	(0.0196)
Push stock ($M)			-0.0005	-0.0022*		-0.0023*
			(0.0012)	(0.0013)		(0.0013)
Push (dummy)		-0.0342		0.2276*	-0.0463	0.2475**
		(0.0777)		(0.1192)	(0.0728)	(0.1220)
Observations	1,101	1,101	342	342	1,101	342
R-squared	0.1535	0.1536	0.3135	0.3178	0.1630	0.3207

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 9: Excluding completely QIDP projects started before 2012

VARIABLES	(1) Preclinical	(2) Preclinical	(3) Trials	(4) Trials	(5) Preclinical	(6) Trials
Pull	-0.0741	-0.0742	-0.0422	-0.0465	-0.0049	-0.2303**
	(0.0540)	(0.0540)	(0.1128)	(0.1125)	(0.0808)	(0.1083)
Pull × Post 2012	0.3146***	0.3151***	0.2707*	0.2930**	0.0669	0.1964
	(0.1007)	(0.1000)	(0.1381)	(0.1376)	(0.1338)	(0.1899)
Pull × Post 2012 × New technology					0.5015***	0.0008
					(0.1806)	(0.2597)
Push flow ($M)	0.0023	0.0024	-0.0061**	0.0016	0.0123**	0.0159*
	(0.0031)	(0.0030)	(0.0027)	(0.0054)	(0.0062)	(0.0090)
Push flow × New technology					-0.0110	-0.0162
					(0.0069)	(0.0108)
Push stock ($M)			-0.0023***	-0.0008		-0.0008
			(0.0009)	(0.0006)		(0.0005)
Push (dummy)		-0.0054		-0.1270***	-0.0058	-0.1315***
		(0.0418)		(0.0310)	(0.0442)	(0.0275)
Observations	27,794	27,794	9,861	9,861	27,794	9,861
R-squared	0.0448	0.0448	0.0855	0.0861	0.0455	0.0871

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10: Results from logit estimation

VARIABLES	(1) Preclinical	(2) Trials	(3) Preclinical	(4) Trials
Pull	-0.5078	0.0127	-0.0228	-0.6403
	(0.4564)	(0.5075)	(0.5233)	(0.7023)
Pull × Post 2012	1.7048***	1.0207*	0.4574	1.0863
	(0.5689)	(0.5911)	(0.7251)	(0.7956)
Pull × Post 2012 × New technology			2.9776**	-0.4743
			(1.3663)	(1.1471)
Push flow ($M)	0.0135	0.0054	0.0607*	0.0439
	(0.0131)	(0.0225)	(0.0322)	(0.0391)
Push flow × New technology			-0.0529	-0.0435
			(0.0362)	(0.0478)
Push stock ($M)		-0.0045		-0.0043
		(0.0031)		(0.0028)
Push (dummy)	-0.0171	-0.5768***	-0.0183	-0.5836***
	(0.2315)	(0.1389)	(0.2421)	(0.1308)
Observations	27,785	9,888	27,785	9,888

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 11: Results for a subsample of projects that started before 2018

VARIABLES	(1) Preclinical	(2) Preclinical	(3) Trials	(4) Trials	(5) Preclinical	(6) Trials
Pull	-0.0716	-0.0716	-0.0121	-0.0158	-0.0041	-0.1426
	(0.0533)	(0.0533)	(0.1121)	(0.1121)	(0.0796)	(0.1267)
Pull × Post 2012	0.2667***	0.2669***	0.2382*	0.2440**	0.0300	0.2953*
	(0.1019)	(0.1019)	(0.1215)	(0.1215)	(0.1323)	(0.1535)
Pull × Post 2012 × New technology					0.4513**	-0.1709
					(0.1975)	(0.2357)
Push flow ($M)	-0.0126***	-0.0113***	-0.0070***	-0.0019	-0.0298***	0.0522***
	(0.0048)	(0.0029)	(0.0025)	(0.0058)	(0.0035)	(0.0047)
Push flow × New technology					0.0242***	-0.0564***
					(0.0042)	(0.0047)
Push stock ($M)			-0.0067***	-0.0056***		-0.0054***
			(0.0008)	(0.0008)		(0.0008)
Push (dummy)		-0.0099		-0.1084***	-0.0069	-0.1165***
		(0.0332)		(0.0303)	(0.0342)	(0.0299)
Observations	25,523	25,523	8,642	8,642	25,523	8,642
R-squared	0.0445	0.0445	0.0924	0.0928	0.0454	0.0937

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

B Structural model

B.1 Motivation

Table 12: Probability of discontinuation (antibiotics, 2000-2020)

VARIABLES	(1)	(2)	(3)	(4)
Same technology, same or later phase discontinuations at t-1	0.000977**		0.001078***	
	(0.000387)		(0.000231)	
Same technology, same or later phase discontinuations at t (other projects)		0.000431		0.000898***
		(0.000345)		(0.000224)
Observations	3,393	3,393	3,699	3,699
R-squared	0.320207	0.318940	0.139971	0.138583
project FE	✓	✓		
year FE	✓	✓	✓	✓
technology FE			✓	✓

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

B.2 Linear representation

For a project in phase 3 we have:

$$V_{i3k} = \frac{\gamma_{i3} + \sum_{d'} q_{d'} V_{idk'_{d'}} + \sum_{m \neq i} \lambda_d(m) \sum_j \sigma_{md(m)jk} V_{i3,l(m,j,k)} + \lambda_3 \text{Emax}(V_{i4,l(i,1,k)} + \varepsilon_{i3k}, 0)}{\rho + \sum_{d'} q_{d'} + \sum_{m \neq i} \lambda_{d(m)} + \lambda_3} \tag{10}$$

Let V_d be a $K \times 1$ vector of V_{idk}

Let Q_0 denote a $K \times K$ vector, such that its (k, k') element is equal to $q_{d'}$ for all k' that are the result of an entry into development of a new molecule in state k, and equal to 0 otherwise.

Let Σ_m denote the $K \times K$ transition matrix implied by the choice probabilities $\sigma_{md(m)}$.

Let $\text{Emax}(V_d)$ denote a $K \times 1$ vector which k^{th} element corresponds to $\text{Emax}(V_{i,d+1,l(i,1,k)} + \varepsilon_{idk}, 0)$

Let D_c be the $K \times K$ diagonal matrix, which (k, k) element contains the denominator of V_{idk}, $\dfrac{1}{\rho + \sum_d q_d + \sum_{m \neq i} \lambda_{d(m)} + \lambda_d}$

Then, stacking expressions from Equation 10 for all the states k, we have the linear representation of the value function:

$$V_3 = D_3 \left[\gamma_{i3} + Q_0 V_3 + \sum_{m \neq i} \Sigma_m V_3 + \lambda_3 \text{Emax}(V_4) \right] \tag{11}$$

Solving for V_3:

$$D_3^{-1}V_3 = \gamma_{i3} + Q_0V_3 + \sum_{m\neq i}\Sigma_m V_3 + \lambda_3 \text{Emax}(V_4)$$

$$D_3^{-1}V_3 - Q_0V_3 - \sum_{m\neq i}\Sigma_m V_3 = \gamma_{i3} + \lambda_3 \text{Emax}(V_4)$$

$$\left[D_3^{-1} - Q_0 - \sum_{m\neq i}\Sigma_m\right] V_3 = \gamma_{i3} + \lambda_3 \text{Emax}(V_4)$$

$$V_3 = \left[D_3^{-1} - Q_0 - \sum_{m\neq i}\Sigma_m\right]^{-1} (\gamma_{i3} + \lambda_3 \text{Emax}(V_4))$$

And in general, for any phase $d \in \{1, 2, 3\}$:

$$V_d = \left[D_d^{-1} - Q_0 - \sum_{m\neq i}\Sigma_m\right]^{-1} (\gamma_{id} + \lambda_d \text{Emax}(V_{d+1})) \tag{12}$$

B.3 Estimation: Step 1

Table 13: Parameter values used in the model

Parameter	Value
Move arrival rates	
λ_2	0.273
λ_3	0.303
Exogenous entry rates	
q_1	0.072
q_2	0.028
q_3	0.017
Cost	
c_2	1.094
c_3	2.726
c_4	10.8
Approval probability	
p	0.84
Discount factor	
ρ	0.05

Table 14: CCP estimation

	Estimate	**Standard Error**	**z**	**P> |z|**
Constant	0.7575	0.131	5.788	0.000
s_k^1	-0.9815	0.148	-6.634	0.000
s_k^2	0.1634	0.122	1.343	0.179
s_k^3	0.4298	0.144	2.995	0.003
s_k^4	0.3610	0.098	3.681	0.000
Phase 2	-0.9232	0.176	-5.231	0.000
Phase 3	-0.9218	0.212	-4.357	0.000
$(s_k^1)^2$	0.1443	0.032	4.558	0.000
$(s_k^2)^2$	0.0137	0.024	0.571	0.568
$(s_k^3)^2$	-0.0537	0.025	-2.147	0.032
$(s_k^4)^2$	-0.0513	0.013	-4.039	0.000
$(s_k^1)^3$	-0.0079	0.002	-4.335	0.000
$(s_k^2)^3$	-0.0005	0.001	-0.378	0.705
$(s_k^3)^3$	0.0017	0.001	1.388	0.165
$(s_k^4)^3$	0.0016	0.000	3.349	0.001
$s_k^1\times$ Phase 2	0.3083	0.101	3.059	0.002
$s_k^2\times$ Phase 2	-0.5183	0.077	-6.728	0.000
$s_k^3\times$ Phase 2	0.3888	0.093	4.158	0.000
$s_k^4\times$ Phase 2	0.0282	0.056	0.508	0.612
$s_k^1\times$ Phase 3	0.3725	0.114	3.256	0.001
$s_k^2\times$ Phase 3	-0.2315	0.100	-2.323	0.020
$s_k^3\times$ Phase 3	-0.3755	0.106	-3.542	0.000
$s_k^4\times$ Phase 3	0.3063	0.074	4.142	0.000
Observations:	1524	**Pseudo R-squ.:**	0.1266	
Df Residuals:	1501	**Log-Likelihood:**	-879.88	
Df Model:	22	**LL-Null:**	-1007.5	

B.4 Estimation: Step 2 details

There are 3 types of decisions observed, advancing from phase 1 to phase 2, from 2 to 3 and from 3 to approval ($d = 4$). It's easier to start from the last decision:

$3 \to 4$ For $l_{3k_0}(\theta, q, \lambda, \tilde{\sigma})$ I need $\tilde{\sigma}_{i31k_0}(\theta, q, \lambda, \tilde{\sigma})$, so I need $V_{i,4,l(i,1,k_0)}$. As shown in Equation 5, V_{i4} is independent of the state k:

$$V_{i4} = \frac{-c_4 + \lambda_4 p \pi_i}{\rho + 1}$$

$2 \to 3$ For $l_{2k_0}(\theta, q, \lambda, \tilde{\sigma})$ I need $\tilde{\sigma}_{i21k_0}(\theta, q, \lambda, \tilde{\sigma})$, so I need $V_{i,3,l(i,1,k_0)}$. $l(i, 1, k_0) = k'$ is the state adjacent to k_0 with one less project in phase 2 and one more in phase 3 (due to i's advancement from 2 to 3). For a $k_0 = (s_{1k}, s_{2k}, s_{3k}, s_{4k})$, $k' = (s_{1k'}, s_{2k'}, s_{3k'}, s_{4k'}) = (s_{1k}, s_{2k} - 1, s_{3k} + 1, s_{4k})$

$$V_{i3k'} = \frac{-c_{i3} + \overbrace{\sum_{d'} q_{d'} V_{i3k'_{d'}}}^{1.} + \overbrace{\sum_{m\neq i} \lambda_{d(m)} \sum_j \sigma_{md(m)jk'} V_{i3,l(m,j,k')}}^{2.} + \overbrace{\lambda_2 \text{Emax}(V_{i,4,l(i,1,k')} + \varepsilon_{i3k'}, 0)}^{3.}}{\rho + \sum_{d'} q_{d'} + \sum_{m\neq i} \lambda_{d(m)} + \lambda_3}$$

Now I need to disaggregate $\sum_{m\neq i} \lambda_{d(m)} \sum_j \sigma_{md(m)jk'} V_{i3,l(m,j,k')}$ depending on m's phase $d(m)$.

$$\begin{aligned}\sum_{m\neq i} \lambda_{d(m)} \sum_j \sigma_{md(m)jk'} V_{i3,l(m,j,k')} = &\sum_{m_1:d(m_1)=1} \lambda_1 \sum_j \sigma_{m_1 1jk'} V_{i3,l(m_1,j,k')} \\ &+ \sum_{m_2:d(m_2)=2} \lambda_2 \sum_j \sigma_{m_2 2jk'} V_{i3,l(m_2,j,k')} \\ &+ \sum_{m_3:d(m_3)=3} \lambda_3 \sum_j \sigma_{m_3 3jk'} V_{i3,l(m_3,j,k')}\end{aligned}$$

Then, using Equation 8 I can write each $V_{i3k'_x}$ in expressions 1-3 as a function of $\sigma_{i21k''_x}$. For example, $V_{i3k'_{11}}$ in terms of CCPs will be equal to $\ln(\frac{\sigma_{i21k''_{11}}}{1-\sigma_{i21k''_{11}}})$, where k''_{11} is a state that results in k'_{11} after a project in phase 2 advances. So if $k'_{11} = (s_{1k} - 1, s_{2k} - 1, s_{3k} + 1, s_{4k})$, then $k''_{11} = (s_{1k} - 1, s_{2k}, s_{3k}, s_{4k})$.

from expression 1.:

$$\begin{aligned} k'_1 &= (s_{1k} + 1, s_{2k} - 1, s_{3k} + 1, s_{4k}) \quad & k''_1 &= (s_{1k} + 1, s_{2k}, s_{3k}, s_{4k}) \\ k'_2 &= (s_{1k}, s_{2k}, s_{3k} + 1, s_{4k}) \quad & k''_2 &= (s_{1k}, s_{2k} + 1, s_{3k}, s_{4k}) \\ k'_3 &= (s_{1k}, s_{2k} - 1, s_{3k} + 2, s_{4k}) \quad & k''_3 &= (s_{1k}, s_{2k}, s_{3k} + 1, s_{4k}) \end{aligned}$$

from expression 2.:

$$\begin{aligned} k'_{21} = l(m_1, 1, k') &= (s_{1k} - 1, s_{2k}, s_{3k} + 1, s_{4k}) \quad & k''_{21} &= (s_{1k} - 1, s_{2k} + 1, s_{3k}, s_{4k}) \\ k'_{22} = l(m_1, 0, k') &= (s_{1k} - 1, s_{2k} - 1, s_{3k} + 1, s_{4k}) \quad & k''_{22} &= (s_{1k} - 1, s_{2k}, s_{3k}, s_{4k}) \\ k'_{23} = l(m_2, 1, k') &= (s_{1k}, s_{2k} - 2, s_{3k} + 2, s_{4k}) \quad & k''_{23} &= (s_{1k}, s_{2k} - 1, s_{3k} + 1, s_{4k}) \\ k'_{24} = l(m_2, 0, k') &= (s_{1k}, s_{2k} - 2, s_{3k} + 1, s_{4k}) \quad & k''_{24} &= (s_{1k}, s_{2k} - 1, s_{3k}, s_{4k}) \\ k'_{25} = l(m_3, 1, k') &= (s_{1k}, s_{2k} - 1, s_{3k}, s_{4k} + 1) \quad & k''_{25} &= (s_{1k}, s_{2k}, s_{3k} - 1, s_{4k} + 1) \\ k'_{26} = l(m_3, 0, k') &= (s_{1k}, s_{2k} - 1, s_{3k}, s_{4k}) \quad & k''_{26} &= (s_{1k}, s_{2k}, s_{3k} - 1, s_{4k}) \end{aligned}$$

Then I can rewrite $V_{i3k'}$ as:

$$\begin{aligned} V_{i3k'} = \frac{1}{\rho + \sum_{d'} q_{d'} + \sum_{m\neq i} \lambda_{d(m)} + \lambda_3} \times \Bigg[&-c_{i3} + \overbrace{q_1 V_{i3k'_1} + q_2 V_{i3k'_2} + q_3 V_{i3k'_3}}^{1.} \\ &+ \overbrace{\sum_{m_1:d(m_1)=1, m_1\neq i} \lambda_1 \left(\sigma_{m11k'} V_{i3,k'_{21}} + (1 - \sigma_{m11k'}) V_{i3,k'_{22}}\right)}^{2.1.} \\ &+ \overbrace{\sum_{m_2:d(m_2)=2:m_2\neq i} \lambda_2 \left(\sigma_{m21k'} V_{i3,k'_{23}} + (1 - \sigma_{m21k'}) V_{i3,k'_{24}}\right)}^{2.2.} \\ &+ \overbrace{\sum_{m_3:d(m_3)=3:m_3\neq i} \lambda_3 \left(\sigma_{m31k'} V_{i3,k'_{25}} + (1 - \sigma_{m31k'}) V_{i3,k'_{26}}\right)}^{2.3.} + \overbrace{\lambda_3 \left(\ln(1 + e^{V_{i4}}) + \nu\right)}^{3.} \Bigg] \end{aligned}$$

plugging in the CCPs:

$$= \frac{1}{\rho + \sum_{d'} q_{d'} + \sum_{m \neq i} \lambda_{d(m)} + \lambda_3} \times \Bigg[-c_{i3}$$

$$+ \overbrace{q_1 \ln(\frac{\sigma_{i21k_1''}}{1-\sigma_{i21k_1''}}) + q_2 \ln(\frac{\sigma_{i21k_2''}}{1-\sigma_{i21k_2''}}) + q_3 \ln(\frac{\sigma_{i21k_3''}}{1-\sigma_{i21k_3''}})}^{1.}$$

$$+ \overbrace{\sum_{m_1 : d(m_1)=1, m_1 \neq i} \lambda_1 \left(\sigma_{m11k'} \ln(\frac{\sigma_{i21k_{21}''}}{1-\sigma_{i21k_{21}''}}) + (1-\sigma_{m11k'}) \ln(\frac{\sigma_{i21k_{22}''}}{1-\sigma_{i21k_{22}''}}) \right)}^{2.1.}$$

$$+ \overbrace{\sum_{m_2 : d(m_2)=2 : m_2 \neq i} \lambda_2 \left(\sigma_{m21k'} \ln(\frac{\sigma_{i21k_{23}''}}{1-\sigma_{i21k_{23}''}}) + (1-\sigma_{m21k'}) \ln(\frac{\sigma_{i21k_{24}''}}{1-\sigma_{i21k_{24}''}}) \right)}^{2.2.}$$

$$+ \overbrace{\sum_{m_3 : d(m_3)=3 : m_3 \neq i} \lambda_3 \left(\sigma_{m31k'} \ln(\frac{\sigma_{i21k_{25}''}}{1-\sigma_{i21k_{25}''}}) + (1-\sigma_{m31k'}) \ln(\frac{\sigma_{i21k_{26}''}}{1-\sigma_{i21k_{26}''}}) \right)}^{2.3.}$$

$$+ \overbrace{\lambda_3 \left(\ln(1 + e^{\frac{-c_{i4} + p_4 \pi_i}{\rho+1}}) + \nu \right)}^{3.} \Bigg]$$

Finally, assuming that the strategies are symmetric ($\sigma_{mdjk} = \sigma_{djk}$), I can simplify:

$$V_{i3k'} = \frac{1}{\rho + \sum_{d'} q_{d'} + \sum_{m \neq i} \lambda_{d(m)} + \lambda_3} \times \Bigg[-c_{i3}$$

$$+ \overbrace{q_1 \ln(\frac{\sigma_{i21k_1''}}{1-\sigma_{i21k_1''}}) + q_2 \ln(\frac{\sigma_{i21k_2''}}{1-\sigma_{i21k_2''}}) + q_3 \ln(\frac{\sigma_{i21k_3''}}{1-\sigma_{i21k_3''}})}^{1.}$$

$$+ \overbrace{s_{k'}^1 \lambda_1 \left(\sigma_{m11k'} \ln(\frac{\sigma_{i21k_{21}''}}{1-\sigma_{i21k_{21}''}}) + (1-\sigma_{m11k'}) \ln(\frac{\sigma_{i21k_{22}''}}{1-\sigma_{i21k_{22}''}}) \right)}^{2.1.}$$

$$+ \overbrace{s_{k'}^2 \lambda_2 \left(\sigma_{m21k'} \ln(\frac{\sigma_{i21k_{23}''}}{1-\sigma_{i21k_{23}''}}) + (1-\sigma_{m21k'}) \ln(\frac{\sigma_{i21k_{24}''}}{1-\sigma_{i21k_{24}''}}) \right)}^{2.2.}$$

$$+ \overbrace{(s_{k'}^3 - 1) \lambda_3 \left(\sigma_{m31k'} \ln(\frac{\sigma_{i21k_{25}''}}{1-\sigma_{i21k_{25}''}}) + (1-\sigma_{m31k'}) \ln(\frac{\sigma_{i21k_{26}''}}{1-\sigma_{i21k_{26}''}}) \right)}^{2.3.}$$

$$+ \overbrace{\lambda_3 \left(\ln(1 + e^{\frac{-c_4 + p\pi_i}{\rho+1}}) \right)}^{3.} \Bigg]$$

$1 \to 2$ For $l_{1k_0}(\theta, q, \lambda, \tilde{\sigma})$ I need $\tilde{\sigma}_{i11k_0}(\theta, q, \lambda, \tilde{\sigma})$, so I need $V_{i,2,l(i,1,k_0)}$. $l(i, 1, k_0) = k'$ is the state adjacent to k_0 with one less project in phase 1 and one more in phase 2 (due to

i's advancement from 1 to 2). For a $k = (s_{1k}, s_{2k}, s_{3k}, s_{4k})$, $k' = (s_{1k'}, s_{2k'}, s_{3k'}, s_{4k'}) = (s_{1k} - 1, s_{2k} + 1, s_{3k}, s_{4k})$

$$V_{i2k'} = \frac{-c_{i2} + \overbrace{\sum_{d'} q_{d'} V_{i2k'_d}}^{1.} + \overbrace{\sum_{m \neq i} \lambda_{d(m)} \sum_{j} \sigma_{md(m)jk'} V_{i2,l(m,j,k')}}^{2.} + \overbrace{\lambda_2 \text{Emax}(V_{i,3,l(i,1,k')} + \varepsilon_{i2k'}, 0)}^{3.}}{\rho + \sum_{d'} q_{d'} + \sum_{m \neq i} \lambda_{d(m)} + \lambda_2}$$

Now I need to disaggregate $\sum_{m \neq i} \lambda_{d(m)} \sum_{j} \sigma_{md(m)jk'} V_{i2,l(m,j,k')}$ depending on m's phase:

$$\begin{aligned}
\sum_{m \neq i} \lambda_{c(m)} \sum_{j} \sigma_{md(m)jk'} V_{i2,l(m,j,k')} = & \sum_{m_1 : d(m_1)=1, m_1 \neq i} \lambda_1 \sum_{j} \sigma_{m1jk'} V_{i2,l(m_1,j,k')} \\
& + \sum_{m_2 : d(m_2)=2 : m_2 \neq i} \lambda_2 \sum_{j} \sigma_{m2jk'} V_{i2,l(m_2,j,k')} \\
& + \sum_{m_3 : d(m_3)=3 : m_3 \neq i} \lambda_3 \sum_{j} \sigma_{m3jk'} V_{i2,l(m_3,j,k')}
\end{aligned}$$

Then, using Equation 8 I can write each $V_{i2k'_x}$ in expressions 1-3 as a function of $\sigma_{i11k''_x}$. For example, $V_{i2k'_{11}}$ in terms of CCPs will be equal to $\ln(\frac{\sigma_{i11k''_{11}}}{1-\sigma_{i11k''_{11}}})$, where k''_{11} is a state that results in k'_{11} after a project in phase 1 advances. So if $k'_{11} = (s_{1k} - 1, s_{2k} - 1, s_{3k} + 1, s_{4k})$, then $k''_{11} = (s_{1k} - 1, s_{2k}, s_{3k}, s_{4k})$.

from expression 1.:

$$\begin{aligned}
k'_1 &= (s_{1k}, s_{2k} + 1, s_{3k}, s_{4k}) & k''_1 &= (s_{1k} + 1, s_{2k}, s_{3k}, s_{4k}) \\
k'_2 &= (s_{1k} - 1, s_{2k} + 2, s_{3k}, s_{4k}) & k''_2 &= (s_{1k}, s_{2k} + 1, s_{3k}, s_{4k}) \\
k'_3 &= (s_{1k} - 1, s_{2k} + 1, s_{3k} + 1, s_{4k}) & k''_3 &= (s_{1k}, s_{2k}, s_{3k} + 1, s_{4k})
\end{aligned}$$

from expression 2.:

$$\begin{aligned}
k'_{21} &= l(m_1, 1, k') = (s_{1k} - 2, s_{2k} + 2, s_{3k}, s_{4k}) & k''_{21} &= (s_{1k} - 1, s_{2k} + 1, s_{3k}, s_{4k}) = k' \\
k'_{22} &= l(m_1, 0, k') = (s_{1k} - 2, s_{2k} + 1, s_{3k}, s_{4k}) & k''_{22} &= (s_{1k} - 1, s_{2k}, s_{3k}, s_{4k}) \\
k'_{23} &= l(m_2, 1, k') = (s_{1k} - 1, s_{2k}, s_{3k} + 1, s_{4k}) & k''_{23} &= (s_{1k}, s_{2k} - 1, s_{3k} + 1, s_{4k}) \\
k'_{24} &= l(m_2, 0, k') = (s_{1k} - 1, s_{2k}, s_{3k}, s_{4k}) & k''_{24} &= (s_{1k}, s_{2k} - 1, s_{3k}, s_{4k}) \\
k'_{25} &= l(m_3, 1, k') = (s_{1k} - 1, s_{2k} + 1, s_{3k} - 1, s_{4k} + 1) & k''_{25} &= (s_{1k}, s_{2k}, s_{3k} - 1, s_{4k} + 1) \\
k'_{26} &= l(m_3, 0, k') = (s_{1k} - 1, s_{2k} + 1, s_{3k} - 1, s_{4k}) & k''_{26} &= (s_{1k}, s_{2k}, s_{3k} - 1, s_{4k})
\end{aligned}$$

from expression 3.:

$$k'_3 = l(i, 1, k') = (s_{1k} - 1, s_{2k}, s_{3k} + 1, s_{4k}) \quad k''_3 = (s_{1k} - 1, s_{2k} + 1, s_{3k}, s_{4k}) = k'$$

Then I can rewrite $V_{i2k'}$ as:

$$V_{i2k'} = \frac{1}{\rho + \sum_{d'} q_{d'} + \sum_{m \neq i} \lambda_{d(m)} + \lambda_2} \times \Bigg[\gamma_{i2} + \overbrace{q_1 V_{i2k'_1} + q_2 V_{i2k'_2} + q_3 V_{i2k'_3}}^{1.}$$
$$+ \overbrace{\sum_{m_1 : d(m_1)=1, m_1 \neq i} \lambda_1 \left(\sigma_{m11k'} V_{i2,k'_{21}} + (1 - \sigma_{m11k'}) V_{i2,k'_{22}} \right)}^{2.1.}$$
$$+ \overbrace{\sum_{m_2 : d(m_2)=2 : m_2 \neq i} \lambda_2 \left(\sigma_{m21k'} V_{i2,k'_{23}} + (1 - \sigma_{m21k'}) V_{i2,k'_{24}} \right)}^{2.2.}$$
$$+ \overbrace{\sum_{m_3 : d(m_3)=3 : m_3 \neq i} \lambda_3 \left(\sigma_{m31k'} V_{i2,k'_{25}} + (1 - \sigma_{m31k'}) V_{i2,k'_{26}} \right)}^{2.3.} + \overbrace{\lambda_2 \text{Emax} \left(V_{i,3,k'_3} + \epsilon_{i2k}, 0 \right)}^{3.} \Bigg]$$

plugging in the CCPs:

$$= \frac{1}{\rho + \sum_{d'} q_{d'} + \sum_{m \neq i} \lambda_{d(m)} + \lambda_2} \times \Bigg[-c_{i2}$$
$$+ \overbrace{q_1 \ln(\frac{\sigma_{i11k''_1}}{1 - \sigma_{i11k''_1}}) + q_2 \ln(\frac{\sigma_{i11k''_2}}{1 - \sigma_{i11k''_2}}) + q_3 \ln(\frac{\sigma_{i11k''_3}}{1 - \sigma_{i11k''_3}})}^{1.}$$
$$+ \overbrace{\sum_{m_1 : d(m_1)=1, m_1 \neq i} \lambda_1 \left(\sigma_{m11k'} \ln(\frac{\sigma_{i11k''_{21}}}{1 - \sigma_{i11k''_{21}}}) + (1 - \sigma_{m11k'}) \ln(\frac{\sigma_{i11k''_{22}}}{1 - \sigma_{i11k''_{22}}}) \right)}^{2.1.}$$
$$+ \overbrace{\sum_{m_2 : d(m_2)=2 : m_2 \neq i} \lambda_2 \left(\sigma_{m21k'} \ln(\frac{\sigma_{i11k''_{23}}}{1 - \sigma_{i11k''_{23}}}) + (1 - \sigma_{m21k'}) \ln(\frac{\sigma_{i11k''_{24}}}{1 - \sigma_{i11k''_{24}}}) \right)}^{2.2.}$$
$$+ \overbrace{\sum_{m_3 : d(m_3)=3 : m_3 \neq i} \lambda_3 \left(\sigma_{m31k'} \ln(\frac{\sigma_{i11k''_{25}}}{1 - \sigma_{i11k''_{25}}}) + (1 - \sigma_{m31k'}) \ln(\frac{\sigma_{i11k''_{26}}}{1 - \sigma_{i11k''_{26}}}) \right)}^{2.3.}$$
$$+ \overbrace{\lambda_2 \left(\ln(1 + \frac{1}{1 - \sigma_{i21k''_3}}) \right)}^{3.} \Bigg]$$

Finally, assuming that the strategies are symmetric ($\sigma_{mdjk} = \sigma_{djk}$), I can simplify:

$$V_{i3k'} = \frac{1}{\rho + \sum_{d'} q_{d'} + \sum_{m \neq i} \lambda_{d(m)} + \lambda_2} \times \Bigg[-c_{i2}$$

$$+ \overbrace{q_1 \ln(\frac{\sigma_{i11k''_1}}{1-\sigma_{i11k''_1}}) + q_2 \ln(\frac{\sigma_{i11k''_2}}{1-\sigma_{i11k''_2}}) + q_3 \ln(\frac{\sigma_{i11k''_3}}{1-\sigma_{i11k''_3}})}^{1.}$$

$$+ \overbrace{s^1_{k'} \lambda_1 \left(\sigma_{m11k'} \ln(\frac{\sigma_{i11k''_{21}}}{1-\sigma_{i11k''_{21}}}) + (1-\sigma_{m11k'}) \ln(\frac{\sigma_{i11k''_{22}}}{1-\sigma_{i11k''_{22}}}) \right)}^{2.1.}$$

$$+ \overbrace{(s^2_{k'} - 1) \lambda_2 \left(\sigma_{m21k'} \ln(\frac{\sigma_{i11k''_{23}}}{1-\sigma_{i11k''_{23}}}) + (1-\sigma_{m21k'}) \ln(\frac{\sigma_{i11k''_{24}}}{1-\sigma_{i11k''_{24}}}) \right)}^{2.2.}$$

$$+ \overbrace{s^3_{k'} \lambda_3 \left(\sigma_{m31k'} \ln(\frac{\sigma_{i11k''_{25}}}{1-\sigma_{i11k''_{25}}}) + (1-\sigma_{m31k'}) \ln(\frac{\sigma_{i11k''_{26}}}{1-\sigma_{i11k''_{26}}}) \right)}^{2.3.}$$

$$+ \overbrace{\lambda_3 \left(\ln(1 + e^{\frac{-c_4 + p\pi_i}{\rho+1}}) \right)}^{3.} \Bigg]$$

B.5 Simulations

Basing off Equation (3), at any point in time, one of three types of events will occur: exogenous arrival of a project in development, project move due to a firm's decision, or no change.

In particular:

1. With probabilities $q_{d'h}$, there is an exogenous arrival of a new project into development at phase c
2. With probabilities $s^d_k \lambda_d h$, one of the projects move due to a firm's decision
3. With the remaining probability there is no change.

In the simulations, I set h to one day ($\frac{1}{365}$, as the other parameters are defined at the annual basis), and take an initial set of projects. In each iteration I draw a uniform random e which defines the type of the event. The probability of a decision taking place in a given technological class can vary in each period, as it depends on the number of projects in the different phases of development (s^d_k). The different events have the following consequences:

1. **Arrival:**
   - Adjustment of the state variable
   - Draws of the project's characteristics (QIDP eligibilty) using pre-determined probabilities
2. **Decisions in phases 1 and 2:**
   (a) Draw a project that gets to take a decision
   (b) Determine whether the project receives the push incentive (draw a random uniform and compare to the probability of push)

(c) Take a draw of ϵ_{ick}

- If $V_{i,c+1,l(i,1,k)} + \epsilon_{ick} > 0$ the project moves forward
- Otherwise, the project is discontinued

(d) Adjustment of the state variable

3. **Decisions in phase 3:**

(a) Draw a project that gets to take a decision

(b) Take a draw of ϵ_{i3k}

- If $V_{i4} + \epsilon_{i3k} > 0$ the project moves to approval (dropped from the set of active projects and saved in a list of projects in approval)
- Otherwise, the project is discontinued

(c) Adjustment of the state variable

In parallel, the process of an exogenous arrival of new technologies into the pipeline takes place, expanding the number of technologies active in development according to rates observed in the data.