EXPECTED PROFITS AND THE SCIENTIFIC NOVELTY OF INNOVATION

DAVID DRANOVE
NORTHWESTERN UNIVERSITY

CRAIG GARTHWAITE
NORTHWESTERN UNIVERSITY AND NBER

MANUEL HERMOSILLA
JOHNS HOPKINS UNIVERSITY

March 2020

Abstract

Innovation policy involves trading off monopoly output and pricing today in exchange for incentives for firms to develop new products. While existing research demonstrates that expected profits fuel R&D investments, little is known about the novelty of funded projects. We expand this literature by examining the scientific novelty of pharmaceutical R&D following the creation of Medicare Part D. We find little evidence that the implied positive demand shock prompted firms to undertake scientifically novel R&D, as measured by whether scientific approaches had been used before. However, some evidence suggests that firms invested in products involving novel combinations of scientific approaches.

1. Introduction

Pharmaceutical pricing attracts large amounts of economic, policy, and political attention. This is perhaps not surprising. The five, six, and even seven figure price tags for many drug treatments not only bear little relationship to marginal costs, they make potentially life-saving treatments unaffordable for nearly anyone who lacks generous insurance coverage, while making insurance itself increasingly expensive for everyone. Moreover, many of these high pharmaceutical prices are the result of deliberate policy choices. Not only do drug makers receive patent protection, the United States government deliberately leaves drug pricing largely to market forces. This stands in sharp contrast to most other developed nations, where drug prices are constrained through the exercise of monopsony power by purchasing entities. Nevertheless, patent protection and market-based pricing remain a pivotal component of the existing U.S. innovative system, as they encourage the development of new products by solving a fundamental economic "hold up" problem; i.e. biopharmaceutical firms may be unwilling to make value-creating investments in new products without a reasonable belief they will be able to appropriate a large enough share of the created value.

In accordance with the theory underlying these policies, there is broad consensus in the academic literature that increases in the expected profitability of new drugs elicits a supply response (e.g. Ward and Dranove, 1997; Acemoglu and Linn, 2004; Finkelstein, 2004; Cerda, 2007; Kyle and McGahan, 2012; Blume-Kohout and Sood, 2013; Dubois et al., 2015). Opponents of policies intended to reduce drug prices cite this body of empirical evidence demonstrating *dynamic innovation effects*, and argue that price regulation would diminish the future supply of potentially life-enhancing and life-saving *breakthrough* drugs that offer new treatment pathways.[1] Supporters of policies to limit drug prices counter that the overall supply response masks considerable heterogeneity in the types of drugs

[1] For example, Heartland Institute Senior Fellow Joseph Bast (2004) writes: "Increasing importation means cutting off the stream of investment that makes this system sustainable. It means fewer new lifesaving drugs."

developed in pursuit of new market opportunities. Specifically, they argue that dynamic innovation effects are largely limited to *me-too* or *follow-on* drugs, whereby drug makers "play it safe" and primarily duplicate the science of established products.[2] The largely unstated assumption in this argument is that the welfare created by such products is insufficient to justify the deadweight loss caused by today's higher prices and the associated rent seeking.

This debate highlights the fact that regulations aimed at reducing drug prices may not only impact consumer welfare through changes in the *amount* of R&D investment undertaken by firms, but also through the *nature* of funded projects. While the existing literature conclusively shows that the prospect of lowering expected profits will reduce the total amount of innovation, it provides little evidence about the composition of this response.

In this paper, we expand the existing literature by revisiting the industry response to the increase in expected profits created by the passage of Medicare Part D.[3] We specifically focus on the novelty (rather than the number) of clinical trial activity initiated in response to the positive demand shock implied by the program. There are many ways to classify novelty.[4] In this paper, we consider whether a drug represents the first application of a new molecular-targeting design and would therefore be considered novel from a *scientific perspective*. This definition is intended to capture information about how the inherent risk and/or cost of development to the innovative firm varies across potential products.

---

[2] For example, Marcia Angell, former editor of the *New England Journal of Medicine*, states that "[i]n fact, the big drug companies now concentrate mainly on … producing variations of top-selling drugs already on the market—called '*me-too*' drugs (emphasis added). There is very little innovative research in the modern pharmaceutical industry, despite its claims to the contrary" (Angell, 2010). Wikipedia defines "me-too drugs" as a drug product that contains an active pharmaceutical ingredient that is chemically related, and usually very structurally similar, to a known active pharmaceutical ingredient…The term follows from the phrase "me too" and is usually used in a negative way, the idea being the me-too drug simply rode the coattails of the research and development done to develop the prototype API."

[3] In a recent paper that informs our methodology, Blume-Kohout and Sood (2013) find that after the 2003 passage of Medicare Part D, which increased the demand for drugs by elderly Americans, drug makers ramped up their research into drugs targeting seniors. However, these authors do not examine the novelty of the response, which is the central question of our paper.

[4] For example, in a prior working paper, we considered a measure of whether a drug offered therapeutic innovation, i.e. whether a drug under development had the potential to be the first to treat a particular condition (Dranove et al, 2014). One might also examine sales of drugs that reach the market, though this is a mixture of both innovation and potential rent seeking.

While it is clear potential new drug products vary in their scientific novelty, there is an obvious question as to why firm investment responses to demand shocks may systematically differ based on the potential product's novelty. A simple way to think about this would be to consider that firms make investment decisions in new products by selecting projects from a set of opportunities, each of which has an expected discounted net present value (NPV). The elasticity of investments in new products with respect to a demand shock ultimately depends on the hazard rate of the associated density function of these NPVs. It is not hard to imagine that the distribution of NPVs differs based on the novelty of the potential product and that there are situations and market conditions where the associated hazard rate is smaller for more innovative drugs.[5] Lacking knowledge about hazard rates by the innovativeness of the project, the magnitude of the elasticity of innovation to expected market profits at any given point in time, and whether or not that impact varies by the novelty of the project, is an open empirical question that we attempt to address in our setting.

In answering the question of whether and how the investment decisions of firms vary based on scientific novelty, we focus on clinical trial activity rather than the number of released products. We do this for several reasons. First, very few experimental drugs actually reach the market, greatly limiting the statistical power of any study focusing solely on the introduction of new products. In addition, our focus in this paper is on the way firms' R&D investments respond to changes in expected demand, which is better proxied by variation in the number of executed clinical trials rather than that of successful products. Our measure of scientific novelty captures an important element of the uncertainty facing innovative firms at the time they make investment decisions, and thereby informs

[5] It can be shown that for some underlying distributions of NPVs and some range of threshold NPVs above which research is undertaken, taking a mean preserving spread over NPVs – which could represent an increase in innovativeness – will necessarily reduce the elasticity.

the debate on whether drug makers target new market opportunities by engaging largely in safer "me-too" research, or riskier "breakthrough" research.

Indeed, as we show below, our measure of scientific novelty is correlated with other potentially welfare relevant metrics such as sales revenue and federal designations of therapeutic efficacy. Information from scientifically novel activities could even be beneficial by fostering future attempts at innovation, even if the original activities did not themselves directly result in a new product. Accordingly, our results show that patents underlying the projects that our metric qualifies as scientifically novel tend to garner a larger amount of forward patent citations that others deemed less novel. This suggests that scientifically novel drug development projects may affect welfare by influencing the path of science even if they do not directly result in new products.

To characterize the scientific novelty of drug development activities (i.e., clinical trials), we exploit the fact that the modern drug development process often relies on efforts to increase therapeutic value through the deployment of a specific "target-based action" (TBA). Each TBA corresponds to a pair of both a targeted biologic entity and the mechanism used to modify its function. For example, a drug associated with the TBA "p38 MAP kinase inhibitor" acts by inhibiting the function of p38 mitogen-activated protein kinases, while the TBA described as a "thromboxane A2 antagonist" works by antagonizing (i.e., blocking the response of) the thromboxane A2 lipid. As such, TBAs provide precise descriptions of the biological approach used by experimental products to produce a pharmacological effect. Given the comprehensiveness of our dataset and the large the number TBAs used by products therein (over seven thousand), we can confidently score the scientific novelty of R&D activities by assessing the extent to which a given TBA (or combination thereof) has been deployed in the past.

An important aspect of our analysis is the recognition that there are two ways to consider novelty with respect to TBAs. The first aligns with the usual notion of scientific novelty discussed in

the literature, namely, whether an experimental drug is the first to use or "translate" a *single given scientific approach* (i.e., a TBA) for therapeutic purposes. We track this form of novelty through a "translational novelty score," which is based simply on the number of previously tested products using the same approach. The underlying assumption behind this measure is that each application of a TBA in a clinical trial increases scientific knowledge about that particular TBA, thereby decreasing the riskiness of its subsequent applications. In other words, the first firms to deploy a given TBA are making riskier investments that have the potential for true scientific breakthroughs but also the greater potential for failure. This paves the ways for later firms to deploy the same TBAs with less risk. In addition to making smaller scientific contributions, subsequent deployments more closely correspond to the kinds of "me-too" investments that tend to be derided by industry critics.

The second way to measure drug novelty is rooted in the observation that many experimental drugs expand treatment options by utilizing more than a single TBA. Scientific novelty can therefore also stem from the deployment of new *combinations* of TBAs. Our interest in this form of novelty spurs from the increasingly wide recognition in the medical community that drugs modifying the function of a single biological target may be limited in their ability to treat diseases with a complex systems biology, like asthma, type 2 diabetes mellitus, HIV, and bacterial infections, among others. As the medical community attempts to tackle these more complex conditions, drug combinations or "cocktails" are increasingly becoming the standard of care (Fishman and Porter, 2005; Zimmerman et al., 2007). Such combinations are purported to enable pharmacological effects that are either more robust, complementary or reinforcing (Zimmerman et al., 2007). Accordingly, approximately 24 percent of the candidates in our data employ multiple TBAs.

Despite these facts, existing definitions of novelty in the literature largely ignore the potential innovation that can come from combinations of existing scientific techniques. Our second measure, which we call the "recombinant novelty" score, recognizes that the recombination of scientific

approaches previously deployed in isolation is itself risky (albeit potentially not as risky as the first use of a TBA in isolation), but can also create new therapeutic value. This is true even in situations where the underlying TBAs of the bundle have themselves been independently deployed. Such previous deployments provide some information about the TBAs in isolation but do not necessarily decrease the risk of a novel combination of TBAs.

By classifying both the target and the mechanism, TBAs provide more accurate information about scientific novelty than simply considering the broader concept of a mechanism of action. This level of detail allows us to more accurately classify whether products represent decisions by firms to undertake novel research. This is also true compared to other measures used in previous work where novelty was defined based on the number of available treatment options at the condition level (Dranove et al., 2014).

To demonstrate the importance of the detail provided by our two novelty metrics, consider, the set of newly developed products to treat hepatitis-C (HCV). In 2008, Vertex initiated a phase III clinical trial for telapravir (Incivek). This product was a combination of existing TBAs and under our translational novelty metric—which gives zero for the least novel and one for the most novel therapies—earned a score of 0.33.[6] Given there were existing (but ineffective) hepatitis-C treatments available at the time, telapravir would not be scored as particularly novel under a metric based on new treatment options at the condition level such as those used in Dranove et al. (2014).

Approximately 3 years after the release of telapravir, Gilead released its first blockbuster HCV cure sofosbuvir (Sovaldi). Again, given sofosbuvir was not the first product to treat HCV and the fact that telapravir was a fairly effective treatment, it would not be considered particularly novel under a metric based on new treatment options. However, sofosbuvir is classified in the most novel category

[6] The product is a combination of the TBA "hepatitis c virus ns3 protease inhibitor" with had zero previous phase 3 deployments and "p-glyco protein inhibitor" which had 4 previous phase 3 deployments

using our TBA-based metric of translational novelty (i.e. a product using a TBA that has never been deployed before in that phase of clinical development). This illustrates the clarity of information provided by our TBA-based measure compared to disease-based novelty metrics.

While sofosbuvir was a dramatic and innovative step forward in HCV, it still was an incomplete treatment for many individuals with this condition. Soon after its release, Gilead released a product that combined sofosbuvir and ledipasvir into a single product (Harvoni). Under our translational novelty metric this product receives a score of 0.29.[7] However, this was the first time these TBAs were combined and therefore it is classified in the most novel category for recombinant novelty. This novelty was rewarded in the market as this combination product quickly surpassed the sales of sofosvubir, while also becoming the standard of care for HCV.

This standard of care, however, still required 12 weeks of treatment and was not available for all genotypes. This led to another combination product of glecaprevir and pibrentasvir (Mavyret). As the field had meaningfully advanced by this point, this product had a translational novelty score of only 0.1.[8] However, this was only the second time this combination of TBAs had been deployed in a phase 3 trial, so the recombinant novelty score was 0.5. This was again reflected in the data where both Mavyret and Harvoni compete heavily for market share in the remaining HCV population (Liu, 2019). Taken together, the example of HCV products demonstrates how the detail available in our TBA based metric allows us to carefully classify the scientific novelty of potential new products.

To identify the causal effect of Medicare Part D we follow the previous literature and exploit the age-specific nature of the beneficiaries of the expansion (e.g., Blume-Kohout and Sood, 2013). Older individuals have a different disease profile than their younger counterparts and we calculate the

[7] This product was a combination of the TBA "hepatitis c virus ns5b polymerase inhibitor" which had 3 previous phase 3 deployments and "hepatitis c virus protein ns5a inhibitor" which had 2 previous deployments. (2 deployments)

[8] This product had the same combination of TBAs as Harvoni, i.e. a "hepatitis c virus ns3 protease inhibitor" and a "hepatitis c virus protein ns5a inhibitor."

share of each clinical indication exposed to patients covered by Medicare. We then examine whether changes in this clinical trial activity after the passage of Part D varies systematically with this measure of Medicare market share. Our assumption is that if the change in expected returns resulting from Part D is driving a change in clinical trial activity, products targeting indications with greater exposure to Medicare would experience a larger increase in such activity compared to those with less exposure.

Overall, our findings suggest the strongest responses to the change in demand caused by the increase in Part D come from clinical trials that represent less scientific novelty. For example, we find a clear increase in clinical trials for those TBAs most frequently used in prior drug development efforts. This increase begins relatively quickly and stabilizes in 2009, suggesting some drugs were likely "pulled off the shelf" and entered into clinical trials almost immediately after Part D increased their expected profitability. To provide perspective on the overall magnitude of the effect, from 2012-2018 there was a 106 percent larger number of clinical trials for these least novel drugs.

In contrast, the change in clinical trials for the most novel category (i.e. TBAs deployed for the first time) was more muted and not evident until 2015. In particular, for the 2012-2018 period, we estimate only a 14 percent increase in trials for the most novel drugs. Furthermore, the estimated increase in the most novel category was almost entirely driven by small molecule products. Biological products saw no increase in clinical trials for previously undeployed TBAs; however, we do estimate an increase in the development of biological products for TBAs in all other novelty categories. As we discuss below, it is unclear to what degree the use of TBAs in small molecule products is a measure of an *ex ante* decision to undertake a more novel or risker project compared to an *ex post* rationalization of the investment. Looking at novelty among combinations of TBAs, we find increases in clinical trials across all levels of novelty. For the least novel recombinant category, we show an 84 percent increase in clinical trials compared to a far more modest 14 percent increase in clinical trial activity for the most novel combinations. It is possible that this more muted effect for the most novel products reflects a

longer development timeline for these products. That said, we do have nearly 15 years of data since the passage of Part D so there is a meaningful length of time to observe effects.

Taken together, our results suggest that the dynamic innovation effects from the passage of Part D involves clinical trials that do not represent the largest amounts of scientific novelty. The most translationally novel products exhibit little response to the marginal demand shocks. There is, however, more novelty in the deployment of combinations of TBAs in response to potential returns. As we consider these results, it is important to remember that the demand shock from Part D represents a relatively modest shock to a pharmaceutical industry that includes over approximately $400 billion in U.S. revenue and over $1 trillion in global revenue per year. As we discuss in more detail below, it remains unclear whether a larger shock would change this pattern of novelty, perhaps by changing the returns to truly novel products. Before presenting our measures of novelty and main results, we next discuss the existing literature, describe our data, and provide a more thorough discussion of TBAs and the drug development process.

2. Background

2.1 Prior Literature

The connection between expected profits and investments in innovation has been widely examined. Several studies document the link between expected profitability and pharmaceutical R&D. Ward and Dranove (1997) and Acemoglu and Linn (2004) examine long-term shifts in U.S. demand, based on epidemiological data and demographic trends, respectively. Cerda (2007) extends these methods to account for the endogeneity of market size to prior drug discovery, while Dubois et al. (2015) extend the prior work to cover global demand. Finkelstein (2004) finds that government subsidies to vaccine research generated increased production of new vaccines. Kyle and McGahan

(2012) show that increases in patent protection encourage more R&D. None of these papers examine whether the response varies with the novelty of the drug.

The closest paper to ours in this literature is Blume-Kohout and Sood (2013), who also study the pharmaceutical innovative response to Part D. Their paper, however, differs from ours in several ways. The fundamental difference is that these authors, like the rest of the existing literature, restrict their attention to documenting an overall supply response and ignore aspects related to the *type* of drugs whose development is catalyzed. Thus, by construction, they cannot comment on the degree of novelty in the innovative response that they identify. Beyond this core difference, we are also able to examine a considerably longer sample and separate out the development of small molecules and biological drug products, products where the economics of the production process might reasonably lead to differential responses to the same demand shock.[9] Our results suggest that all of these considerations (novel versus non-novel innovation, short versus long-term impacts, small versus large molecules) are first order aspects in characterizing the novelty of Part D's dynamic innovation effects.

Another related study is Krieger et al. (2018). This study appears to be even more similar to our paper in that the authors also consider questions regarding the scientific novelty of new products in response to Medicare Part D. However, our two papers address different underlying economic questions. Krieger et al. (2018) is interested in whether firms that gain more access to internal capital, as a result of having a stock of products positively affected by the creation of Part D, change the riskiness of their R&D investments. They are primarily examining an important corporate finance question about potential credit market inefficiencies rather than our separate question about the response to a change in expected market size. In addition, these authors rely on a different underlying measure of novelty based on the chemical structure similarity of experimental small molecules. While

[9] The period covered by sample used by Blume-Kohout and Sood (2013) is 1998-2010 period. The period covered by our sample is 1997-2018.

our "translational" novelty measure is quite similar in spirit to theirs, their approach in measuring novelty is limited in the sense that it is only available for chemically synthesized drugs (i.e., small molecules) and can't be readily applied to biological products.[10] In recent years, these biological products have represented a growing share of high cost drugs and therefore are quite relevant to the question about the innovation effects of drug spending.

2.2 Medicare Part D

Medicare is a social insurance program in the U.S. that primarily covers individuals over the age of 65. First created in 1965, this program originally covered some portion of the costs for physician and hospital services, but offered very limited pharmaceutical coverage. As pharmaceutical spending grew so did political pressure to expand Medicare to cover these products. This resulted in the passage of Medicare Part D as part of the Medicare Modernization Act of 2003. Prior to the Act's passage, it was unclear whether there would be a prescription drug benefit added to Medicare and certainly little information about its eventual form. Part D became effective in 2006. In our analysis, we consider 2004 as the date where firms would first change their investment decisions in response to the law. However, we are also cognizant that there was marked uncertainty about the impact of Part D on the industry prior to its final regulations were established. Therefore, we allow our estimates of the change in the firm's investments to change over the post-reform time period.

The implementation of Part D caused an immediate increase in pharmaceutical insurance coverage for seniors. In 2006, nearly 26 million elderly individuals were covered by the expansion. This number grew to over 30 million by 2011, the end of our sample period. Many of these individuals already had coverage and therefore the program is estimated to have provided *new coverage* to

[10] As we explain below in Section 3.2, our treatment of the Part D shock (which follows from Duggan and Scott-Morton, 2010) also differs significantly from Krieger et al. (2018). Because they are interested in a firm's response based on additional internal capital, by necessity firms must have at least one product on the market. In contrast, we are interested in the "pull" response from the change in market size so we can also include It is also important to note that the many smaller firms without approved drugs might be the source of the most novel innovations.

approximately 5 million individuals (Gruber and Engelhart, 2011). Therefore, Part D represents a change in expected returns that is smaller than might be expected by simply observing the size of the program. Despite this relatively high degree of crowd-out, studies have found that the broader Part D coverage caused an increase in elderly pharmaceutical use (Ketcham and Simon, 2008; Yin et al., 2008).

Pharmaceutical companies might be concerned that an expanded Medicare would use its monopsony power to reduce drug prices in the same way it does on payments to hospitals and physicians. This would decrease the potential increase in revenues resulting from a more broadly insured population. However, the structure of Part D made this unlikely. Unlike other government programs such as the Veterans Administration, Part D is run by a series of private insurance programs (similar to the health insurance exchanges under the Affordable Care Act). In addition, the law explicitly prohibits the Center for Medicare and Medicaid Services (CMS) from directly bargaining with pharmaceutical firms. This suggests that Part D represented a substantial and positive financial shock for pharmaceuticals products targeting conditions with a large number of elderly patients. This shock may have been even more apparent for biologic products in our database since the profits of small molecule products are generally limited by the eventual prospect of generic entry. For biologicals, in contrast, the prospect of generic (biosimilar) competition was a remote possibility over the covered period.[11] As a result, firms introducing new biologic products could expect near monopoly status on their specific product for a longer time horizon than the length of their patent.[12] This could be particularly important for products that have been "sitting on the shelf" for a period of time resulting in less patent protection for firms making new investments to bring the product to market.

3. Data

3.1 Pharmaceutical R&D data

[11] At the time period of the passage of Part D (and for over a decade after) there was no regulatory process for a firm to introduce biosimilars (broadly equivalent to generic biologic products).
[12] Competition for these biological products could still emerge from therapeutic substitutes.

We utilize pharmaceutical R&D data from Clarivate Analytics' *Cortellis* pharmaceutical intelligence subscription service, which is widely regarded as the most comprehensive and up-to-date repository of records for international pharmaceutical R&D activities.[13] Cortellis uses information from a variety of sources, updated and curated daily by more than 500 experts.[14] Our April 2019 data include information on over 70,000 molecules developed worldwide by over 4,300 companies since the early 1970s. These data provide a broad picture of the development efforts of the global pharmaceutical industry. We will focus on drug development activities surrounding the implementation of Part D, from 1997-2018.

Cortellis data describe product development histories at the indication level (i.e., product/targeted disease level). The data track development from the preclinical stage, when molecules are formulated, optimized, and tested with animal and simulation models, through the initiation of clinical trials on humans (i.e., Phases 1, 2, and 3), and, finally, through regulatory assessment activities (review and eventual approval). Cortellis assigns each record a single date indicating when testing begins.

Table 1 illustrates how we construct our analytic sample. Throughout the Table, we differentiate between the two main types of products: (1) small molecules and (2) biologicals. Panel A describes the number of products and trials available from the full dataset. Although companies initiate the lion's share of development activities within our sample period, a significant share was initiated before it. We also note that there are about twice as many trials as there are products. This accounts for the fact that products often have more than one indication (i.e., targeted condition), each of which

[13] The comprehensiveness of the Cortellis dataset has made it a popular data source for recent research in the area (e.g., Chandra et al., 2019; Krieger et al., 2018; Gaessler and Wagner, 2018; Hermosilla, 2019).
[14] Information sources include company reports, clinical trial registries, academic articles and conferences, regulatory agencies, specialized media, among others.

requires some amount of independent development activity. Panels B and C describe the subsets of the full dataset that have sufficient information for our analysis.

As we describe in the next section, we base our measures of technological novelty on associated target-based actions (TBAs), which broadly correspond to the mechanism by which products produce a pharmacological effect. TBA information is available for about 60 percent of products in the full dataset (70 percent of trials) and these are the products that we include in our analysis.[15] As shown in Panel B, products with TBA information are deployed in about 96,000 trials across sample and non-sample periods. We utilize this entire set of records to construct the novelty metrics.

While Cortellis data describe the targeted condition by each indication, our analysis is conducted using slightly more aggregated cross-sectional units, 3-digit ICD-9 disease category codes.[16] As detailed below, this aggregation is required to link to our proxy for the magnitude of the Part D demand shock across targeted conditions. Panel C of Table 1 reports statistics for the subset of the data for which the proxy was successfully linked. This subsample includes over 76,000 trials initiated during the sample period (corresponding to about 36,000 unique products).

Figure 1 presents the distribution of trials across therapeutic areas, showing a marked domination of cancer-targeted research. This distribution is broadly consistent with the patterns reported by Ernst & Young (2012), which aims to describe the overall state of the industry. Figure 2 presents the number of trials initiated each year, by stage. The bulk of trials (about 60 percent)

[15] The Cortellis editorial team collects drugs' TBA information from the same types of sources from which other drug information is obtained (company websites, press releases, journal articles, clinical trial registries, conferences, etc.). According to our communication with the company, "if that field [i.e., TBA] is empty it means either that information has not been publicly released, or is not relevant to that particular drug" (quote from email exchange with Cortellis' support team, bracketed clarification added). Whereas missing TBA information for small molecules may primarily stem from cases in which these have not been ascertained, for biologicals it is likely due to firms' secrecy concerns. There could be a concern that firms developing truly novel products may be more likely to obscure this information from the market. In Appendix A we examine whether the creation of Part D resulted in products with more exposure to Medicare being less likely to provide TBA information. We find no evidence of such an effect.

[16] ICD stands for International Classification of Diseases. The ICD-9 ontology (now ICD-10) is the standard categorization used for medical billing.

correspond to preclinical development. Phase 1 and Phase 2 each account for about 15 percent and Phase 3 accounts for about 5 percent. Whereas the number of Phase 1-3 trials display a slight increasing trend, the number of preclinical trials experiences a sharp increase in 2008-2010.

3.2 Measuring the Part D demand shock across targeted conditions

To measure the impact of Part D on the demand for drugs across targeted conditions, we adopt the strategy of Duggan and Scott-Morton (2010) and Blume-Kohout and Sood (2013). This strategy relies on a variable called "Medicare Market Share" ("MMS"), which reflects the varying importance of drug consumption by Medicare enrollees across treatment areas. We identify the effect of Part D by examining changes in the number of clinical trials for conditions that are more or less exposed to the expansion as measured by MMS.

We compute MMS using the Medical Expenditure Panel Survey (MEPS). This is a large and representative sample for the utilization of prescription drugs, medical services, and insurance in the United States. MMS is constructed from the yearly "prescription" and "insurance" MEPS files of years prior to the enactment of Part D (1997-2003). Prescription files list the drugs consumed by respondents. The associated medical conditions are reflected by a 3-digit ICD-9 treatment area code for each record. Insurance files report the type of insurance coverage held by the respondent each year. Combining the information of these two files, we compute MMS as share of prescriptions issued to Medicare enrollees within each 3-digit ICD-9 disease category.[17] A high MMS value will reflect a larger impact of the Part D shock on the demand for drugs within a certain ICD-9 disease category.

We link MMS to our main dataset by classifying targeted diseases in the Cortellis dataset into their respective ICD-9 categories.[18] The matched dataset covers 337 different ICD-9 categories, each

[17] MEPS representativeness weights are used to construct MMS. While our definition MMS is similar in spirit to Duggan and Scott-Morton (2010), it is not exactly the same. In particular, because their analysis is focused on prices at the brand name drug level, their MMS variable is computed as the percentage of Medicare patients using the drug prior to the passage of Part D.

[18] Two expert medical coders independently assigned 3-digit ICD-9 codes to the targeted conditions listed in the Cortellis dataset. Coding differences were then resolved by one of the authors in consultation with the coders.

of which includes, on average, 3.8 conditions of the Cortellis disease ontology. In line with this aggregation procedure, we will henceforth use "condition" to refer to each 3-digit ICD-9 category. Figure 3 describes the resulting variation of MMS across conditions. The distribution makes intuitive sense: conditions commonly associated with younger people tend to receive lower MMS values; those associated with older populations, larger ones. For example, MMS equals or approaches zero for conditions such as infertility, acne, and contraception. Conditions such as ischemic stroke, bipolar disorder, and headaches are associated with MMS values around the distribution's median. Macular edema, Parkinson's disease, heart failure, and dementia have MMS values in the upper decile.

4. Measuring translational and recombinant novelty using target-based actions (TBAs)

In this section we discuss the science of target-based drug discovery, how we measure scientific novelty, and how our measure of novelty compares to the existing literature.

4.1 Target-based drug discovery

Until the 1980s, drug discovery was primarily an empirical endeavor where firms screened thousands of chemically-synthesized molecules ("small molecules") in search of symptom reduction (Boyd, 1999). From a pharmacological point of view, such effects could only arise if the screened molecule bound to a disease-modifying biological target such as a protein (Zheng et al., 2006). During this period, however, it was rarely known before screening whether and which target a molecule would bind to. *Ex-post* studies could help to ascertain this mechanistic knowledge, but techniques available to do so were still incipient. As a result, most drugs went through early stages of development without a clear understanding of how they modified the biology of the targeted disease (Drews, 2003; Swinney and Anthony, 2011).

The so-called "biotechnology revolution" ignited by the scientific breakthroughs of the 1970s and 1980s fundamentally changed this drug discovery paradigm. These advances allowed researchers

to better understand the biological mechanisms deployed by small molecule drugs and implement a process of rational drug design. More importantly, they also made it possible to "engineer" biological drugs (also referred to as "large molecule" drugs). As opposed to small molecules, biological products are living organisms (e.g., antibodies) purposefully designed to bind to specific targets.

Fueled by the realization that mechanistic knowledge of the drug target could help develop more effective and safe drugs (for example, through improved dosing or sharper characterization of response and toxicity), target-related information became an important input for drug discovery. Besides enabling the design of biologicals, target-related knowledge also catalyzed the innovation of small molecules by helping to "direct" randomized screenings towards chemical structures associated with favored mechanisms. This "hypothesis-based" or "target-based" approach has become a leading way by which the modern pharmaceutical industry identifies new compounds (Swinney, 2004).

To interpret TBA data and how it relates to the strategic decisions of firms, it is important to highlight the differences in the discovery process for small molecules versus biologicals. By definition, biologicals are always formulated through target-based drug discovery. Thus, the respective novelty of TBAs can be seen as the *fundamental technological choice* made by the developer. For small molecules on the other hand, TBAs reflect a technological choice only if knowledge of the target actually guided the screening process. In cases where this information was absent or ignored at the discovery stage, TBAs should be interpreted as an incidental property of the molecule rather than a strategic decision of the firm. While pipeline data (such as Cortellis) do not tell us whether TBA information guided the discovery of each small molecule product, an in-depth review of the protocols used to discover drugs approved by the FDA between 1999 and 2008 found that about one-third of approved small molecules had been discovered without reliance on mechanistic knowledge of the target (Swinney and Anthony, 2011). This suggests that, over the period covered by our sample, non-targeted discovery

was still extensively utilized. Therefore, as we interpret our results, TBA deployment among small molecules should be taken as a less deliberate choice than for biologicals.

4.2 Patterns of TBA deployment

We base our novelty metrics on the idea that each time a product is tested in a clinical trial, information that characterizes the performance of its underlying scientific approach is produced and dispersed among the scientific community. This information likely reduces the risk of subsequent deployments of the particular scientific approach. Thus, for a product with a given TBA, our measure of novelty is based on the number of other products with the same TBA that have been previously tested. We refer to this number as the number of prior TBA "deployments." Here we review the three main patterns of TBA deployment that we observe in our sample.

The first of these patterns corresponds to the continuous growth in the stock of "known" TBAs (i.e., the set of unique deployed TBAs). At the beginning of our sample (1997), products listed by Cortellis referenced 1,064 unique TBAs. Over time, this number grew at a roughly stable rate of about 350 TBAs a year. By 2018, the number of known TBAs reached 7,627. This trend suggests the industry has continuously engaged in attempts to translate novel science for therapeutic use.

Over this period, each known TBA also accumulated deployments. For example, at the beginning of our sample, the average TBA had been deployed in approximately 4 different experimental products. By 2018, this number was about 7.2. Throughout the sample period, the distribution of deployments exhibits a long tail. The vast majority of known TBAs are rarely deployed (five or fewer times), while a small minority accumulates a large number of deployments.

Finally, we note the frequent joint deployment of more than one TBA by a single experimental product. As discussed above, Zimmerman et al. (2007) notes combination products like these can be beneficial in the treatment of complex diseases that affect multiple tissues or cells (such as diabetes), or for multigenic diseases with sparse genetic triggers (as many forms of cancer). One example of this

type of multi-TBA products is the lung cancer drug Lometrexol (Eli Lilly). This product is an antimetabolite, acting by down-regulating the function of metabolites (chemicals that partake in the process of metabolism). Specifically, Lometrexol interferes with the use of folic acid, which is why it is associated to the TBA "folate receptor antagonist." Since Lometrexol also exhibits antineoplastic properties, it is also associated to the TBA "gar transformylase inhibitor."[19] About 30 percent of the trials (24 percent of products) in our data are for multi-TBA drugs. The majority of these are associated with two TBAs.[20] As with individual TBAs, the number of known TBA bundles (i.e., unique TBA combinations including more than one TBA) growths at a roughly constant pace throughout the sample, going from about 504 in early 1997 to 6,153 by late 2018.

4.3 Translational novelty

Our measure of translational novelty quantifies the intensity with which the TBA (or set of TBAs) used by a drug entering a clinical trial has been previously deployed. To describe the metric, formally consider a trial *i* which deploys a single TBA *k*, and let $D_{k(i)}$ represent the number of previous deployments of TBA *k* in stages that are *at least as advanced* as that of trial *i*. For example, if *i* was a Phase 2 trial, $D_{k(i)}$ would represent the count of other products associated with *k* that have been tested in Phase 2 and/or Phase 3 prior to the initiation of trial *i*.[21] However, $D_{k(i)}$ would not count any deployments of the TBA in a Phase 1 trial. We introduce this restriction because testing at more advanced stages produces results of higher evidentiary value (e.g., Phase 3 results will improve the ability to predict a TBA's performance in preclinical testing, but not *vice versa*). With this, the novelty of trial *i* is computed as:

[19] See https://pubchem.ncbi.nlm.nih.gov/compound/Lometrexol.
[20] In particular, 71.4 percent of trials are associated to a single TBA, 20.3 percent to two TBAs, 4.72 percent to three, and the remainder (3.5 percent) to four or more.
[21] A product that has been tested in more than one phase counts as single deployment.

$$NOVELTY_i = \frac{1}{1 + D_{k(i)}} \in (0,1] \quad (1)$$

Thus, the score equals 1 only if trial *i* deploys *k* for the first time ($D_{k(i)} = 0$). If $D_{k(i)} = 1$, the score is 0.5; if $D_{k(i)} = 2$, the score is 0.33; and so on. That is, the score decreases convexly as the number of *k*'s previous deployments increases.

For our measure of translational novelty, trials for products with TBA bundles require us to choose a method to aggregate the individual scores of each included TBA (i.e., the right-hand side of the above expression) into single translational novelty trial score.[22] Our primary approach is to simply average the number of deployments across all included TBAs. Thus, for these "multi-TBA trials," the score equals one only if all included TBAs are deployed for the first time. For trials that combine TBAs that have been previously deployed with others that are being deployed for the first time, resulting scores will be lower than one, outside the 1/n sequence, and often within the (0.5,1) interval. Because the specific values taken by novelty scores are largely irrelevant for our analysis (i.e. we rely on the ordering of these scores and not their particular magnitude), these aggregation issues are not meaningful to our results, and we show in section 5.2 that they do not drive our conclusions.

Figure 4 describes the resulting dispersion of translational NOVELTY scores. The distribution shows a marked bimodal shape, with about 13 percent of trials scoring 1 (completely novel TBAs) and 57 percent scoring below 0.1 (which corresponds to a TBA previously deployed more than 10 times). Thus, a relatively small share of trials can be deemed as translationally novel while the bulk of trials leverage existing scientific knowledge from previous deployments.

We previously noted that TBAs accumulate deployments throughout the sample. This implies that NOVELTY scores will, on average, decrease in later years. Since the implementation of Part D

[22] This is not a concern for our measure of recombinant novelty where we are concerned with the novelty of the bundle and not its individual components.

occurs in the latter portion of our sample, this accumulation of TBA deployments could bias our estimates towards finding stronger responses among less novel drugs. In Appendix B, we present an alternative metric formulation that addresses this issue by quantifying novelty through each TBA's percentile in up-to-date deployment distributions. The two formulations produce highly correlated metrics and thus the corresponding results are quite similar.

4.4 Recombinant novelty

As described earlier, approximately 30 percent of the trials in our data are for multi-TBA drugs. We compute a second recombinant score for this set, measuring deployments at the TBA bundle level. For example, a trial for the multi-TBA Lometrexol referenced earlier would count as single deployment of the two-TBA bundle that includes both the "folate receptor antagonist" and the "gar transformylase inhibitor" TBAs. Deployment counts are then inputted into equation (1) to score the novelty of the combination. Measured in this way, about a third of multi-TBA trials receive scores lower than one (i.e., at least one previous deployment of that exact combination of TBAs). The remainder of trials (all of which deploy a TBA bundle for the first time) can be split in two categories: (i) those of which all included TBAs have been previously deployed in the context of either a different bundle or in isolation (39 percent overall) and (ii) those which are novel because they include a TBA that is itself being deployed for the first time (28 percent overall) – either on its own or in combination with another TBA. For ease of reference in separating these two groups of novel recombinations, we call the first type of bundle, where novelty comes solely from the new combination of TBAs and not the individual components, "pure recombination" trials.

4.5 Caveats and relationship to other approaches to measure drug novelty

Our approach to measuring novelty differs from the bulk of existing literature in two main ways. First, whereas the literature has primarily focused on characterizing novelty through counts of regulatory approvals, our approach fits into the relatively small subset of studies that assess novelty by

looking "under the hood" of drug products, through the lens of a fine-grained categorization of underlying technologies.[23] Two of these studies are the descriptive works and of Kneller (2010) and Lanthier et al. (2013). Kneller (2010) classifies the 252 New Molecular/Biological Entities approved by the FDA between 1998 and 2007 into two novelty categories (follow-ons and scientifically novel) according to whether the associated mechanism of action had been previously deployed by an approved drug. Lanthier et al. (2013) also focus on New Entities approved by the FDA (1987-2011), classifying them into "first-in-class," "advance-in-class," and "addition-to-class" or "me-too".[24] Although their partition by pharmacological classes is coarser than it would be if based on TBAs, it is difficult to assess by how much. Shih et al. (2018) examine pharmaceutical pipelines (1996-2016), focusing on the deployment of novel mechanisms of action at the targeted disease level. Their analysis addresses translational novelty by partitioning projects into those deploying "validated" mechanisms (i.e., products deploying them have previously reached regulatory approval or near-approval), those deploying "un-validated" mechanisms (i.e., products deploying them have failed development), and those deploying "emerging" mechanisms (i.e., products deploying them could get approval but have still not been submitted). Shih et al. find that the deployment of novel mechanisms has been relatively more successful for rare/orphan diseases.

Another study in the same realm is Krieger et al. (2018, "KLP"), who assess translational novelty based on a measure of chemical structure similarity. KLP define a molecule's novelty as the inverse similarity score with respect to the most similar other (known) molecule. Their approach is premised on the "Similarity Property Principle," which states that structurally similar molecules are likely to have similar functional properties (Johnson and Maggiora, 1990). Since TBAs correspond to precise characterizations of the mechanisms of action of a drug product, the KLP approach targets

[23] This literature is reviewed by Kesselheim et al. (2013). Other approaches described in this review rely on counts regulatory designations and measures of novelty self-reported by developers.
[24] Lanthier et al. (2013) define classes based on FDA-established pharmacologic class designations.

the same notion of translational novelty of our framework.[25,26]

Beyond these differences, we offer three important caveats to our approach to measuring novelty. First, Krieger (2019) shows that products targeting the same TBA may vary in terms of their chemical structures, suggesting additional TBA deployments may propose a measure of therapeutic novelty that our methodology does not fully capture. Second, while Cortellis is regarded as comprehensive and up-to-date, there is wide recognition that information for a non-negligible share of all pharmaceutical R&D does not enter the public domain and may thus be "missing" from all datasets (e.g., Doshi et al., 2013). This consideration suggests that, if anything, our metrics may at times overstate "true" novelty – however this would only be a concern to which this differed systematically with exposure to Medicare. Finally, there could be a concern that previous deployments that resulted in successful trials may provide more information to subsequent innovators than those which didn't provide such information. In this sense, these deployments might be more "novel." Therefore, in our robustness results below we will also present evidence on whether firms differentially target TBAs which have been "validated" (i.e. resulted in successful products) compared to those that have not.

5. Part D impact on translational novelty

5.1 Main effects

[25] KLP note that a molecule's chemical structure may not fully characterize its pharmacological effect---"similar molecules may have divergent properties." This suggests that, by relying on TBAs instead of structural similarity, we are reducing the amount of noise embedded in our measure of technological novelty.

[26] There are two additional points of differentiation between the KLP approach and ours. First, since chemical structures are available for small molecules only, the KLP approach cannot be directly used to score the novelty of biological products. This is an important caveat since biologicals account for a large percentage of R&D (about one third of trials in our data) and are often perceived as the more active source of therapeutic breakthroughs. Second, the variation of novelty scores resulting from their and our method has a different interpretation, particularly at lower realms of the novelty distribution. To see this recall that our approach awards a low translational novelty score if the tested TBA has a relatively large number of previous deployments. In contrast, low KLP scores follow from the existence of at least one other molecule of high structural similarity. This means that: (i) whereas our scores are determined by global patterns of TBA deployment, KLP scores are primarily driven by local patterns (i.e., similar innovation), and (ii) low KLP novelty scores do not necessarily reflect the repeated deployment of a specific scientific approach (in their case, chemical structure).

To estimate potential changes in the novelty of R&D investments in response to the Part D demand shock, we group our trials into four groups based on their NOVELTY score. The first three groupings involve drugs with NOVELTY < 1; i.e., drugs with TBAs that were used at least once before. We divide these into three equal-sized groups: Categories T1 (least novel), T2, and T3.[27] On average, trials in Category T1 have NOVELTY = 0.02 (65 previous deployments), whereas those in Category T3 have NOVELTY = 0.33 (2.6 previous deployments). Category T4 includes trials with NOVELTY = 1 (i.e. those for which all associated TBAs are deployed for the first time). Thus, our categorization primarily relies on the ordering of novelty scores, not their specific values.

Figure 5 shows the total number of trials initiated each year. Each panel presents the trials for products in each category of novelty, split by whether the target indication has an above- or below-median MMS value. Prior to the passage of Part D in 2003, the number of below- and above-median MMS trials displayed broadly similar patterns, suggesting the absence of confounding pre-trends. This stability continued through 2004 and 2005 while the specifics of the regulatory framework for Part D were being debated. After Part D went into effect in 2006, there was a gradual increase in the number of trials for products deploying less novel TBAs (those in categories T1 and T2) and targeting conditions with an above-median MMS, compared to similarly novel trials targeted below-median MMS conditions. Consistent with the results of Blume-Kohout and Sood (2013), this response magnifies through 2009 and appears to stabilize around 2012. The change is a far less pronounced for trials in Category T3 and not noticeable at all for the most novel trials of Category T4. Together, these patterns suggest that the Part D demand shock primarily fueled clinical trial activity for the less translationally novel drugs.

[27] Category T3 includes the vast majority of multi-TBA trials that include a TBA being deployed for the first time.

To quantify these graphical relationships, we turn to regression analysis. We define N_{istc} as the total number of stage *s* trials initiated during year *t*, which target condition *i* (3-digit ICD-9) and belong to translational novelty category *c*. We estimate the following negative-binomial specification:

$$N_{istc} = f\left(\alpha_c + \beta_{0c} \cdot MMS_i + \sum_p \left[\beta_{pc} \cdot MMS_i \cdot 1[t \in p]\right] + \eta_{a(i)} + \delta_s + \lambda_t\right) \quad (2)$$

The summation index *p* represents a series of five consecutive time periods (2004-5, 2006-8, 2009-11, 2012-14, 2015-18), which we introduce as means to understand the potentially gradual unfolding of the impact of Part D on trial activity. These are captured by the category-specific parameters β, each of which represents the period-specific gradient of trial activity to shock exposure (MMS). Parameters α correspond to category-specific intercepts, and η, δ, and λ respectively represent fixed effects for therapeutic area, stage, and trial initiation year.

Under the assumption that there was no relationship between MMS and development activity prior to 2003, positive $\hat{\beta}$ estimates indicate that the Part D demand shock led to increased trials initiated in a given time period/novelty category. Since there is a large percentage of *istc* observations with zero initiated trials (about 75 percent), *f* corresponds to the zero-inflated version of the negative binomial count model. Throughout our analysis, we estimate robust standard errors that allow for arbitrary correlation among observations.

Table 2 presents the estimates obtained from the full sample of trials. The estimated effect of Part D on clinical trial activity varies systematically across novelty categories. The change in clinical trials is most immediate and largest in magnitude for the trials that exhibit the least translational novelty. For Category T1 trials (least novel), there is a nearly immediate effect after the program's passage. This effect grows over time. For the Category T2, the effect of Part D is milder and delayed, with the first noticeable increase beginning in 2009. For Categories T3 and T4, the effect of Part D

only appears in the last period (2015-18) and the effect is meaningfully smaller than for the less novel categories. The long delay in this effect could reflect the fact that truly innovative activity takes a longer time to respond. We are not aware of any confounding factors in the market affecting drugs for seniors that occurred after the creation of Part D.

To illustrate the magnitude of the implied effects in Table 2, in Figure 6 we compare the actual pattern of trials (solid lines) against the prediction in the counterfactual where Part D was not implemented (dashed lines).[28] For graphical clarity, we only present the least and most novel categories (respectively shown by red and blue lines). The effect of Part D on Category T1 trials emerges promptly in 2004 and stabilizes around 2012. For the latter period (2012-18), estimates imply that the Part D shock lead to a 106 percent increase in the number of initiated trials in this category. In contrast, for Category T4, the actual and counterfactual trends are relatively close throughout the sample. Over the 2012-18 period, our estimates indicate that Part D led to a 14 percent increase in the number of initiated trials in this category. For the second and third categories that we omitted from the figure, the estimated 2012-18 impacts are 32 percent and 11 percent respectively.

5.2 Robustness of our main results

As we mentioned before, Appendix B describes an alternative specification for the novelty metric that accounts for the fact that average novelty declines in time due to TBAs' deployment accumulation. Key results are qualitatively replicated. Appendix C presents a falsification test as well a series of additional robustness checks for the results presented in Table 2. For falsification, we deploy our model on 1997-2003 data assuming that Part D was enacted in 2001. Key interactions are not statistically significant.

In terms of robustness, we first focus on potential concerns posed by the large share of Oncology trials in our sample. It may be that (i) cancer-specific patterns in the translation of new

[28] These counterfactual predictions are obtained by setting $\hat{\beta}$ estimates to zero.

science drive the results in the broader sample, and (ii) the impact of Part D on cancer-targeted innovation may be confounded by the availability of Medicare Part B insurance for cancer treatment. The primary pattern of our results remains largely unchanged when we exclude cancer-targeted trials. We also address the aggregation of TBA-specific translational novelty scores for multi-TBA trials. We re-estimate our model under three protocols: (i) multi-TBA trials excluded from the estimation sample, (ii) aggregation of TBA-specific scores NOVELTY chooses the maximum score among included TBAs, and (ii) aggregation chooses the minimum score. In every case, results quite similar to our main estimates of Table 2.

We also address whether the firm investments may differ based on whether a TBA had achieved validation in the form of an approved product. We base our analysis on the above-referenced work of Shih et al. (2018), who classify drug development projects based on whether the deployed mechanism has received regulatory approval. To gauge where our novelty measure should account for the possibility that firms differ in their selection of products based on whether a TBA had been validated, Figure 7 shows the percentage of trials deploying validated TBAs, differencing by whether the targeted condition has an above- or below-median MMS score.[29] About half of all trials deploy validated TBAs, a figure that is largely stable over time. More importantly, there do not appear to be meaningful differences based on Medicare orientation—the two trends are highly correlated, both before and after the passage of Part D. That is, whereas regulatory validation may release information that reduces the technological uncertainty of TBAs in question, it does not appear that regulatory validation was a first-order dimension guiding the firm's selection of R&D investments following the passage of the Part D program.

5.3 Small molecules versus biologicals

[29] We use the same definition of Shih et al. (2018). This is, a trial is said to deploy a validated TBA if at the time the trial is initiated, the TBA has reach the pre-registration stage or above (i.e., registration or market launch).

There are fundamental differences in the drug development process for small molecule products compared to biologicals. Recall that many small molecule products include TBAs as an *ex post* feature of drug development while for nearly all biological products it is an *ex ante* feature of the development decision. To more fully understand the riskiness of firm investment decisions, and not just the novelty of the projects they pursue, we separately analyze trials for small and large molecules.

Table 3 reproduces our main estimates for small molecules (Panel A) and biologicals (Panel B). While both set of results retain the broad pattern of the full-sample estimates, they also suggest that the impact of Part D was generally smaller for small molecules than for biologicals. In addition, the impact of Part D for biologicals is stronger for less translationally novel categories. This invites questions as to whether the late-period increase in Category T4 trial activity, which is exclusively attributable to small molecules, was the result of a fully premeditated strategy or was more of an incidental outcome given TBAs are ascertained ex-post for many small molecules. Future research should further consider the strategic decisions of firms with respect to novel small molecule products.

5.4 Early versus late-stage development

The drug development process involves multiple stages of clinical trials. After each completed trial, firms receive information about the efficacy and potential commercial success of the underlying product. They then make decisions about whether to move forward based on this information. Assuming a positive trial outcome in a phase, the decision to continue investing in a particular drug is a result of the firm's expectations of commercial success. Therefore, it is conceivable that at any one time firms have a stock of products that have made their way through some stages of development but have stalled do to a lack of commercial potential. A demand shock could justify moving these marginal products back "off the shelf" and into additional trials.

Table 4 decomposes the increase in R&D activity into early and late stage trials. Early stage activities are given by preclinical trials in Panel A while late stage activities are comprised of Phases 1-

3 of clinical trials in Panel B. Whereas results for late-stage activities closely resemble those in the full sample, for early stage development we observe a relatively more uniform impact across novelty categories. This suggests that the Part D shock may have prompted a degree of experimentation with novel approaches in the context of the relatively less expensive pre-clinical testing that was not fully followed up by more expensive later-stage development. We also find prompt and strong impacts on late-stage Category T1 trials (estimated 195 increase for 2012-18). This suggests that some of the early responses to the Part D shock may have resulted from the re-activation of the development of previously shelved products (likely those considered marginally unprofitable prior to the shock).

6. Part D impacts on recombinant novelty

We previously described how novelty stemming from specific TBA combinations used by multi-TBA products may entail a distinct type of therapeutic value. By relying on metrics that average the individual novelty of TBAs in bundles, our previous analysis has neglected this source of novelty stemming from the bundle itself. Here we investigate whether the Part D shock fueled novel trial in the sense that the tested molecule comprises a previously unexploited *combination* of TBAs.

Figure 8 plots the fraction of all trials initiated each year for multi-TBA products, separating them according to whether the targeted disease has a below- or above-median MMS score. Both series exhibit a marked increase around 2003, which could be driven by the boost that genomic science received around that time due to the completion of the Human Genome Project (whose first draft and final versions were respectively made available in 2000 and 2003). However, these trends are broadly parallel over most of the sample period, which suggests this increase results from broad changes in science rather than Part D increasing clinical trials for products combining TBAs.

To evaluate whether Part D may have fueled pharmaceutical R&D of different degrees of recombinant novelty we restrict our attention to multi-TBA trials for which this notion of novelty is pertinent (21,780 trials) and again partition them into categories. Recall that category R1 represents

the least novel combination, as these drugs involve bundles of TBAs that have previously been tested in the same combination. On average, products in this first category had 13 previous deployments. The remaining trials are split into Categories R2 and R3. Category R2 involves "pure recombination" trials, for which the novelty is only based on the new combination of products. Category R3 trials in which at least one TBA has never been previously tested individually or in a bundle. The drugs are largely even split across these categories: 32 percent of multi-TBA trials are placed into Category R1, 40 percent into Category R2, and 28 percent into Category R3.[30]

Figure 9 shows the number of trials initiated each year within each category based on whether they target a condition with an above or below median MMS. Consistent with a positive Part D impact, the gap between the number of above- and below-median trials in Categories R1 and R2 is roughly constant prior to the passage of the program but widens in the years after. This serves as suggestive evidence of a positive Part D impact that starts to manifest in 2009 for Category R2 and in 2012 for Category R1. Part D effects are largely imperceptible for Category R3.

Econometric estimates (from the same specification in the previous section) are largely consistent with these graphical findings. Results in Table 5 imply that the Part D shock led to an increment of 84 and 55 percent in the number of trials initiated in Categories R1 and R2 over the 2012-18 period, respectively. Consistent with our results for Category T4 in the previous section, the estimates of Table 5 suggest that the impact on Category R3 (TBA bundles using a translationally novel TBA) was a much more modest, at about 14 percent.

Tables 6 and 7 report effects by molecule size and development stage respectively. Part D prompted a relatively larger increase of R2 trials for biologicals (93 percent) than small molecules (58 percent) over 2012-18. In line with earlier results, there is no measurable impact for biologicals. Similar to our analysis for translational novelty, the effect of Part D unfolds more promptly for late stage than

[30] While this split is fairly even, there is no reason by construction this needed to be true.

preclinical activities, being first perceptible during 2006-8 (category R2) for the former but only during 2012-14 for the latter. Over 2012-18, the impacts on categories R1 and R2 are 70 and 40 percent for early stage activities, and 86 and 66 for late-stage activities.

7. Relationship between novelty and welfare

Our estimates demonstrate the clinical trials initiated in response to higher expected profits caused by the creation of Medicare Part D varied markedly in their scientific novelty. An important caveat for economists and policymakers is that far more work is needed to understand how this novelty translates to welfare. It is not clear there is a one-to-one relationship between scientific novelty and the economic value created by products. Several studies offer evidence to the idea that increased flows of follow-on pharmaceutical innovation may yield welfare gains by widening choice sets and/or intensifying price competition (Arcidiacono et al., 2013; Bokhari and Fournier, 2013; Branstetter et al., 2016; Chaudhuri et al., 2006; Dutta, 2011; Granlund, 2010). Therefore, it is possible even the least novel products (i.e. those where we observe large innovation investments) could increase welfare.

That said, these benefits likely pale in comparison to those provided by truly novel innovations, whose impact on quality, price competition, and as means to facilitate follow-up innovation should be larger. To support this point, we note our novelty measures are correlated with other characteristics that might be more directly welfare relevant. As a starting point, drug revenue likely has some relation to welfare as it is bounded from above by the willingness to pay of the consumer. If we compare the products listed in the *MedAd News* list of the top 200 highest selling drugs in 2017 against a benchmark sample of products launched between 2008 and 2017,[31] we find that products included on the list had an average translational novelty score for their Phase 3 trials of 0.39 compared to a score of 0.29 for those not included on the list. Similarly, if we look at recombinant

[31] Out of the 200 products listed in the *MedAd News* list, we identified 196 in our data. TBA information was available for 167 of these. The benchmark sample contains other 858 products for which we also have TBA information.

novelty, the Phase 3 novelty score for products included on the list was 0.79 compared to 0.71 for those not included on the list. Both of these differences are statistically significant at a p-value < 0.001.

This result appears to extend beyond the novelty of phase III trials, as shown by the estimates in Table 8. To produce these estimates, we first summarize a product's scientific novelty (which varies across the product's different trials) through a "novelty fixed effect."[32] We then use these fixed effects to explain the probability that each product is listed on the *MedAd News* ranking. The estimate in the first column of Table 8 addresses translational novelty. The positive and precisely estimated coefficient suggests that more translationally novel products are associated with higher listing probabilities. Specifically, the coefficient implies that a one standard deviation larger novelty fixed effect is associated with a 0.024 higher listing probability (30 percent increase over the baseline probability). When we repeat the exercise to study recombinant novelty (second column), we obtain a similar result, although the parameter is imprecisely estimated. In this case, a one standard deviation larger (recombinant) novelty fixed effect is associated to 0.019 higher probability (17 percent increase) of making it to the *MedAd News* list. The estimate of the third column provides support for the concept that multi-TBA products could provide specific additional value. According to the model estimate, these products appear in the ranking with an about 0.06 higher probability (105 percent increase).

Regulatory designations are another potentially relevant welfare metric. As part of the approval process, regulators award different types of designations to products that, if approved, would expand the assortment of therapeutic alternatives in specific ways. Broadly, these designations fall into two categories: (i) those rewarding products that target unmet needs or might constitute real advances over

[32] These fixed effects are obtained from a regression that uses the novelty score as dependent variable, and which also includes month, therapeutic area, and stage fixed effects. Fit statistics suggest that the product-level fixed effects explain the majority of novelty score variation: the R-squared falls from 0.93 to 0.06 when they are excluded from the specification. Fixed effects are entered into the Probit regression after normalizing them to the unit interval. The Probit specification also includes fixed effects for each product's market launch year and for whether each product is a biologic.

existing alternatives, and (ii) those targeting rare diseases.[33] Relying again on "novelty fixed effects," Table 9 shows Probit estimates that relate scientific novelty with the probability of at least one designation of each category.[34] The positive coefficients indicate that products of higher translational and recombinant novelty levels are associated with higher probabilities of receiving designations, with one standard deviation marginal effects ranging between 35 and 55 percent of the baseline probabilities. As shown by the estimates of the last two columns, move novel multi-TBA products are also significantly more likely to receive these awards.

Lastly, we investigate whether our measures of novelty provide some evidence about the way that clinical trial activity could influence future scientific efforts. As we note above, one reason to examine changes in clinical trial activity (rather than simply products that are released) is that such scientific efforts could serve as the basis of future products. At a minimum, they could reduce the risk of future drug development efforts using the same TBAs. To that end, we investigate the correlation of our scientific novelty measures with forward citations of each product's associated US patent. Forward citations can be taken as a proxy for how much a product has fostered follow-on innovation, but also, as a measure of the created economic value (Trajtenberg, 1990; Hall et al., 2005). Table 10 presents the coefficients that we obtain when we regress the log number of forward citations (measured over three- and five-year horizons) on our product-level measures of scientific novelty.[35] Results indicate that the patents of more translationally novel products tend to receive a larger number

[33] In the first category we include "fast track," "accelerated approval," "breakthrough," "priority review" and "promising medicine" designations. The second category includes "orphan drug" and "rare disease" designations.

[34] Designations are aggregated into two categories because they are relatively infrequent events: the percentage of product receiving designations is 8 and 15 for the first and second designation type, respectively. In addition, whereas our data reports whether each product has been awarded a designation, it does not indicate when designations have been awarded. To control for potential censoring effects, the Probit model also includes fixed effects for the year in which each product initiates its first clinical trial.

[35] For a subset of products listed in the Cortellis dataset, we observe the patent number of each product "key patent." The dataset analyzed here considers only products in our sample for which we observe a key patent that has been granted by the USPTO. Forward citations were retrieved during 2018, based on the number of other USPTO patents that cite the former. Patent statistics are available for a total of 6,597 products in our sample. However, to avoid censoring issues, our analysis of 3-year citations uses the data of products whose patents were granted no later than 2014 (N=5,521); the analysis of 5-year citations, those whose patents were granted no later than 2012 (N=4,709).

of forward citations, with estimates pointing to a one-standard-deviation effect of between 4% (five-year horizon) and 5% (three-year horizon). We also find products scoring higher on recombinant novelty also tend to produce a larger number of citations, although the effect is quite imprecise. While this effect may stem from the reduced number of observations entering the regression, it could also be expected that patents that simply recombine ideas have a smaller impact on follow-on innovation.[36]

8. Discussion and Conclusion

Although there exists robust evidence documenting that changes in expected market profitability impact the amount of pharmaceutical R&D (e.g., Acemoglu and Linn, 2004, Blume-Kohout and Sood, 2013, Dubois et al., 2015), these findings almost universally refer to aggregate effects. Policymakers interested in determining the optimal tradeoff between pharmaceutical profits and the incentives for innovation also require some notion of how changes in market size affect the *nature* of R&D. Since much of the debate fueled by potential regulation on drug prices hinges on whether it may hinder the innovation of potential breakthrough drugs (e.g., Danzon, 2000; Bast, 2004), it is important to understand how these regulations might impact the future supply of the most innovative drugs. We fill a gap in the literature by implementing a large-scale analysis of the pharmaceutical industry's innovative response to an exogenous demand shock (Medicare Part D) that differentiates between innovative activities based on their scientific novelty.

Our main result is obtained by categorizing development activities (i.e., clinical trials) based on the novelty of the underlying science of the tested product, which in our data is reflected by descriptions of the targeted biological entities and method used to modify their functions. As such, our measure of novelty speaks directly to firms' willingness to respond to a marginal change in expected demand by engaging in the novel therapeutic translation of science. We find that whereas

[36] This finding mirrors that of Krieger et al. (2018), who also find that key patents of more novel drugs tend to be associated with a larger number of forward citations.

the Part D demand shock led to a prompt and economically significant increase in the amount of development activities oriented at the development of scientific follow-ons, it had a meaningfully more modest and delayed positive impact on those for potential scientific breakthroughs, i.e., the first time a particular scientific approach is deployed in the context of drug development.

This result is robust to a series of checks and holds independently in leading subsets of the data. The finding is also consistent with results from the medical literature. For example, examining the targets employed by drugs approved by the FDA between 2000 and 2005, Zheng et al. (2006) observe that "it appears that the majority of the successful targets have been continuously explored for deriving new therapeutic agents." In the same vein, Zimmerman et al. (2007) note that "unfortunately, few new drugs act at novel molecular targets." and Shih et al. (2018) that "industry output in terms of successful projects in this period [1996-2016] has come primarily from a limited set of well-validated therapeutic mechanisms."[37]

There are several economic theories that could rationalize our main empirical results. Focusing on the demand side, at any point in time, firms have a set of potential projects with a distribution of discounted excepted net present values (NPVs). The elasticity of drug candidate supply with respect to a demand shock depends on the shape of the associated density function of NPVs of candidate drugs. It is easy to imagine situations where the number of available infra-marginal projects is smaller for more innovative drugs than for less novel projects.

There may also be explanations linked to the supply side. For example, the development of potential breakthrough drugs may simply depend on basic science, which is relatively insensitive to demand (Ward and Dranove; 1997). The large risk associated with the development of potential breakthroughs (Hara, 2003) may also contribute to our findings. Sams-Dodd (2005) and Torbert (2003) flesh out some of the sources of this increased risk, indicating that the usual process of novel

[37] The time period shown in brackets has been added to the quote the purposes of clarity.

target validation is challenging and may take several years. Even when usual validation standards are met, questions may remain regarding the target's interaction with the broader organism. These risks may increase the spread of the potential returns of breakthrough drugs, implying that a smaller percentage of these types of drugs are marginally profitable. Put differently, a disproportionate share of breakthrough drugs would have been extremely profitable both before and after the passage of Medicare Part D.

Additional reasons may stem from the organization and interactions between the pharmaceutical and academic sectors. For example, Kneller (2010) finds that academic institutions and small biotech firms have a disproportionally large role in the generation of scientifically novel FDA-approved drugs. Relative to well-funded large pharmaceutical firms, these organizations may be more mission-oriented and resourced constrained, and hence less reactive to changes in market conditions.[38] A more nuanced aspect may be rooted on the potentially uneven applicability of novel scientific approaches across markets. Specifically, Lowe (2010) suggests that novel scientific approaches (i.e., targets) may be intrinsically more adept for the targeting of small-market (orphan) diseases, which might involve early stage research that is less likely to meet expected profitability standards required by large pharmaceutical firms.

Finally, it may be that given the density of NPVs of potential products in the current U.S. system of pharmaceutical pricing, the demand shock from Part D was not sufficiently large to meaningfully change the incentives for developing highly innovative products. As we note above, such were likely quite valuable for firms both before and after the passage of Part D. This is just another way of saying that there may be a long tail of highly positive NPV innovative products, while the density of zero NPV innovative products may be quite small. Under such a distribution of potential

[38] Of course, understanding the effect of demand shocks on these firms would require also carefully considering the incentives of the venture capital firms (and the investors in those firms) that provide the funding for much of this early stage research. Future work should focus on the investment responses for these early stage investors.

NPVs, a larger change in potential market size could generate a meaningful change in supply of innovative products than a smaller change. This is particularly true if these more novel investments are motivated by firms attempting to capitalize on the potential option value of early stage investments — which would be more valuable in a setting with very large potential returns for positive outcomes.

For this reason, we suggest great caution in extrapolating the estimates of our results to far larger changes in the size of the market. Our study, and all other studies that rely on small changes in revenue to identify the magnitude and nature of the innovative response, effectively provide local average treatment effects. In particular, these estimates are local to the density of potential innovative products at any given point in time. There is little reason to suggest we can extrapolate from this local estimate to far larger changes to the returns to innovation, as this will depend on the shape of the density of investment returns. It also may vary based on whether there is a positive or a negative shock. For example, if the overall distribution of net returns to innovative projects is unimodal, and firms currently pursue projects in the right tail, then the density is downward sloping at the zero NPV project. This implies that the density to the left of NPV equal zero project is larger than the density to the right, so that the loss of investment following a negative shock would be larger than our estimated gain in investment following the positive Medicaid Part D shock.

Overall, our results demonstrate that the clinical trial activity caused by the passage of Part D occurred for less scientifically novel approaches. Furthermore, this measure of novelty is correlated with other indicators of novelty that are believed to affect welfare. That said, far more work is needed to understand the welfare implications of our novelty results and to determine whether similar results would be seen for larger financial shocks.

References

Acemolgu, D. and J. Linn, (2004), "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry," *Quarterly Journal of Economics*, 119(3): 1049-1090.

Angell, M., (2010), "Bad Pharma, Bad Medicine," *Boston Review*, accessed at: http://bostonreview.net/angell-big-pharma-bad-medicine

Arcidiacono, P., Ellickson, P., Landry, P., and Ridley, D. B. (2013), "Pharmaceutical Followers," *International Journal of Industrial Organization*, *31*(5), 538-553.

Bast, J. (2004), "The Pros and Cons of Importing Drugs from Canada" Press Release by the Heartland Institute, 4/19/2014.

Bokhari, F. and G. Fournier (2013), "Entry in the ADHD Drugs Market: Welfare Impact of Generics and Me-Too's," *Journal of Industrial Economics*, *61*(2), 339-392.

Boyd, D. B. (1999). Is rational design good for anything? ACS Publications

Branstetter, L., Chatterjee, C., and M. Higgins (2016) "Regulation and Welfare: Evidence from paragraph IV Generic Entry in the Pharmaceutical Industry," *The RAND Journal of Economics*, *47*(4), 857-890.

Blume-Kohout, M. and N. Sood (2013) "Market Size and Innovation: Effects of Medicare Part D on Pharmaceutical Research and Development," *Journal of Public Economics,* 97: 327–336.

Cerda, R. (2007), "Endogenous Innovations in the Pharmaceutical Industry," *Journal of Evolutionary Economics,* 17 (4): 473–515.

Chandra, A., Garthwaite, G., Stern, A, "Characterizing the Drug Development Pipeline for Precision Medicine," in *Economic Dimensions of Personalized and Precision Medicine*, 2019, National Bureau of Economic Research, Inc.

Chaudhuri, S., Goldberg, P., and P. Jia (2006) "Estimating the effects of global patent protection in pharmaceuticals: a case study of quinolones in India," *American Economic Review*, *96*(5), 1477-1514.

Danzon, P. (2000), "Making Sense of Drug Prices," *Regulation*, 23(1): 56-63.

Doshi, P., Dickersin, K., Healy, D., Vedula, S. S., & Jefferson, T. (2013). Restoring invisible and abandoned trials: a call for people to publish the findings. *BMJ*, *346*, f2865.

Dranove, David, Craig Garthwaite, and Manuel Hermosilla (2014). *Pharmaceutical profits and the social value of innovation*. National Bureau of Economic Research Working Paper No. 20212.

Drews, J. (2003). Strategic trends in the drug industry. *Drug discovery today*, *8*(9), 411-420.

Dubois, P., de Mouzon, O., Scott-Morton, F., and P. Seabright. 2015. "Market Size and Pharmaceutical Innovation" *The RAND Journal of Economics*, *46*(4), 844-871.

Duggan, M. and F. Morton (2010). "The Effect of Medicare Part D on Pharmaceutical Prices and Utilization," *American Economic Review,* 100 (1): 590–607.

Dutta, A. (2011), "From Free Entry to Patent Protection: Welfare Implications for the Indian Pharmaceutical Industry," *The Review of Economics and Statistics*, *93*(1), 160-178.

Ernst & Young (2012), "Beyond Borders: Global Biotechnology Report"

Finkelstein, A. (2004), "Static and Dynamic Effects of Health Policy: Evidence from the Vaccine Industry" *Quarterly Journal of Economics*, 119(2): 527-564

Fishman, M. C., & Porter, J. A. (2005). A new grammar for drug discovery. *Nature*, *437*(7058), 491.

Gaessler, F., & Wagner, S. (2019). Patents, data exclusivity, and the development of new drugs. *Available at SSRN 3401226*.

Granlund, D. (2010), "Price and Welfare Effects of a Pharmaceutical Substitution Reform," *Journal of Health Economics*, *29*(6), 856-865.

Engelhardt, G. V., & Gruber, J. (2011). Medicare Part D and the financial protection of the elderly. *American Economic Journal: Economic Policy*, *3*(4), 77-102.

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. *RAND Journal of Economics*, 16-38.

Hara, T. (2003). *Innovation in the pharmaceutical industry: the process of drug discovery and development*. Edward Elgar Publishing.

Hermosilla, M. (2019). Rushed Innovation: Evidence from drug Licensing. *Forthcoming in Management Science.*

Johnson, M. A. and G. M. Maggiora (1990). Concepts and applications of molecular similarity. Wiley

Kesselheim, A. S., Wang, B., & Avorn, J. (2013). Defining "innovativeness" in drug development: a systematic review. *Clinical Pharmacology & Therapeutics*, *94*(3), 336-348.

Ketcham, J. and K. Simon (2008), "Medicare Part D's effects on Elderly drug costs and utilization" *American Journal of Managed Care*, 14(11): 14-22.

Kneller, R. (2010). The importance of new companies for drug discovery: origins of a decade of new drugs. *Nature Reviews Drug Discovery*, *9*(11), 867.

Krieger, J. L., Li, D., & Papanikolaou, D. (2018). *Developing novel drugs* (No. w24595). National Bureau of Economic Research.

Krieger, J.L. (2019). *Trials and Terminations: Learning from Competitors' R&D Failure*. Working Paper Harvard Business School.

Kyle, M. and A. McGahan (2012), "Investments in Pharmaceuticals Before and After TRIPS" *Review of Economics and Statistics,* 94(4): 1157-1172.

Lanthier, M., Miller, K. L., Nardinelli, C., & Woodcock, J. (2013). An improved approach to measuring drug innovation finds steady rates of first-in-class pharmaceuticals, 1987–2011. *Health Affairs*, *32*(8), 1433-1439.

Liu, Angus, "Gilead's stealing hep C share with its cut-rate Eplusa and Harvoni generics," *FiercePharma,* Sep 26, 2019. Accessed at: https://www.fiercepharma.com/pharma/gilead-s-authorized-hcv-generics-now-own-over-20-market on 3/21/2020.

Lowe, Derek (November 4, 2010), "Where Drugs Come From: The Numbers" Blog post for "In the Pipeline" https://blogs.sciencemag.org/pipeline/archives/2010/11/04/where_drugs_come_from_the_numbers

Roth, B. L., Sheffler, D. J., & Kroeze, W. K. (2004). Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. *Nature reviews Drug discovery*, *3*(4), 353.

Sams-Dodd, F. (2005). Target-based drug discovery: is something wrong?. *Drug discovery today*, *10*(2), 139-147.

Swinney, David C., and Jason Anthony. "How were new medicines discovered?" *Nature reviews Drug discovery* 10.7 (2011): 507.

Swinney, D. C. (2004). Biochemical mechanisms of drug action: what does it take for success? *Nature reviews Drug discovery*, *3*(9), 801.

Tobert, J. A. (2003). Lovastatin and beyond: the history of the HMG-CoA reductase inhibitors. *Nature reviews Drug discovery*, *2*(7), 517.

Trajtenberg, M. (1990). A penny for your quotes: patent citations and the value of innovations. *RAND Journal of Economics*, 172-187.

Shih, H. P., Zhang, X., & Aronov, A. M. (2018). Drug discovery effectiveness from the standpoint of therapeutic mechanisms and indications. *Nature Reviews Drug Discovery*, *17*(1), 19.

Ward, M. and D. Dranove (1997), "The Vertical Chain of Research and Development in the Pharmaceutical Industry" *Economic Inquiry,* 33(1): 70-87.

Yin, W., Basu, A., Zhang, J., Rabbani, A., Meltzer, D., and C. Alexander (2008), "The Effect of the Medicare Part D Prescription Benefit on Drug Utilization and Expenditures" *Annals of Internal Medicine*, 148(3): 169-177.

Zheng, C. J., Han, L. Y., Yap, C. W., Xie, B., & Chen, Y. Z. (2006). Progress and problems in the exploration of therapeutic targets (vol 11, pg 412, 2006). *Drug Discovery Today*, *11*(15-16), 717-717.

Zimmermann, G. R., Lehar, J., & Keith, C. T. (2007). Multi-target therapeutics: when the whole is greater than the sum of the parts. *Drug discovery today*, *12*(1-2), 34-42.

Table 1: Data overview.

	Trials initiated before 1997			Trials initiated in 1997-2018			Notes
	Small molecules	Biologicals	Total	Small molecules	Biologicals	Total	
A. Full Cortellis dataset							
No. molecules	3,599	1,135	4,734	40,513	23,699	64,212	
No. trials	5,790	1,876	7,666	82,248	48,888	131,136	
B. Subset with TBA information available							
No. molecules	2,424	584	3,008	25,880	13,431	39,311	Novelty metrics constructed
No. trials	3,975	1,023	4,998	58,527	32,282	90,809	from this entire subset (including trials initiated before 1997)
C. Subset with TBA information available and linked demand shock proxy							
No. molecules				24,119	11,883	36,002	Dependent variable based
No. trials				**50,346**	**25,815**	**76,161**	on bolded figures

Table 2: Part D impacts across translational novelty categories.

	Translational novelty category			
	T1	T2	T3	T4
MMS	-1.332***	-0.318***	0.201**	0.269***
	(0.119)	(0.090)	(0.088)	(0.096)
MMS*1[Year=2004-5]	0.342*	0.141	0.063	-0.039
	(0.179)	(0.159)	(0.145)	(0.167)
MMS*1[Year=2006-8]	0.578***	0.171	0.034	0.035
	(0.155)	(0.132)	(0.126)	(0.135)
MMS*1[Year=2009-11]	1.018***	0.482***	0.109	0.195
	(0.154)	(0.134)	(0.129)	(0.145)
MMS*1[Year=2012-14]	1.716***	0.596***	0.205	0.098
	(0.147)	(0.127)	(0.126)	(0.139)
MMS*1[Year=2015-18]	2.047***	0.925***	0.557***	0.468***
	(0.134)	(0.117)	(0.116)	(0.125)

Zero-inflated negative binomial estimates. The dependent variable is the number of initiated trials, aggregated at the level of targeted condition (3-digit ICD-9 code), stage, year of trial initiation and novelty category (N=103,840). Displayed coefficients correspond to interactions of the variables displayed in rows with indicators for those displayed in columns. The estimated model includes year, therapeutic area, development stage, and novelty category fixed effects. Robust standard errors are presented in parentheses. *Legend:* $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

Table 3: Translational novelty effects by molecule size.

	Translational novelty category			
	T1	T2	T3	T4
	A. Trials for small molecules			
MMS	-1.230***	-0.400***	-0.041	-0.322**
	(0.128)	(0.099)	(0.101)	(0.126)
MMS*1[Year=2004-5]	0.276	0.016	-0.008	-0.022
	(0.183)	(0.171)	(0.167)	(0.188)
MMS*1[Year=2006-8]	0.562***	0.149	-0.003	-0.033
	(0.156)	(0.142)	(0.146)	(0.162)
MMS*1[Year=2009-11]	0.835***	0.332**	-0.204	0.193
	(0.165)	(0.157)	(0.159)	(0.198)
MMS*1[Year=2012-14]	1.622***	0.510***	0.134	0.250
	(0.153)	(0.146)	(0.150)	(0.171)
MMS*1[Year=2015-18]	1.804***	0.823***	0.356**	0.599***
	(0.140)	(0.134)	(0.139)	(0.156)
	B. Trials for biologicals (large mols.)			
MMS	-3.372***	-1.251***	-0.055	0.430***
	(0.244)	(0.141)	(0.126)	(0.128)
MMS*1[Year=2004-5]	1.142***	0.537**	0.222	-0.197
	(0.326)	(0.235)	(0.205)	(0.234)
MMS*1[Year=2006-8]	1.333***	0.464**	0.148	-0.055
	(0.298)	(0.199)	(0.171)	(0.184)
MMS*1[Year=2009-11]	2.151***	1.114***	0.483***	0.012
	(0.278)	(0.185)	(0.164)	(0.179)
MMS*1[Year=2012-14]	2.923***	1.356***	0.353**	-0.264
	(0.270)	(0.187)	(0.172)	(0.192)
MMS*1[Year=2015-18]	3.634***	1.860***	0.919***	0.172
	(0.253)	(0.168)	(0.154)	(0.174)

Zero-inflated negative binomial estimates. The dependent variable is the number of initiated trials, aggregated at the level of targeted condition (3-digit ICD-9 code), stage, year of trial initiation and novelty category (N=95,832 in Panel A and N=82,544 in Panel B). Displayed coefficients correspond to interactions of the variables displayed in rows with indicators for those displayed in columns. The estimated model includes year, therapeutic area, development stage, and novelty category fixed effects. Robust standard errors are presented in parentheses. *Legend:* $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

Table 4: Translational novelty effects by development stage.

	Translational novelty category			
	T1	T2	T3	T4
	A. Early stage development (preclinical)			
MMS	-0.500***	0.620***	0.464***	0.248*
	(0.157)	(0.140)	(0.141)	(0.146)
MMS*1[Year=2004-5]	0.325	-0.015	0.248	0.178
	(0.273)	(0.265)	(0.259)	(0.278)
MMS*1[Year=2006-8]	0.350	-0.173	0.052	0.263
	(0.229)	(0.225)	(0.219)	(0.225)
MMS*1[Year=2009-11]	0.714***	0.116	0.305	0.941***
	(0.236)	(0.227)	(0.224)	(0.275)
MMS*1[Year=2012-14]	1.353***	0.347*	0.521**	0.577**
	(0.222)	(0.210)	(0.218)	(0.228)
MMS*1[Year=2015-18]	1.929***	0.690***	0.725***	0.555***
	(0.206)	(0.197)	(0.198)	(0.208)
	B. Late stage development (Phases 1-3)			
MMS	-3.168***	-0.734***	0.068	0.268**
	(0.202)	(0.116)	(0.103)	(0.120)
MMS*1[Year=2004-5]	1.367***	0.289	0.042	-0.060
	(0.274)	(0.183)	(0.166)	(0.192)
MMS*1[Year=2006-8]	1.916***	0.443***	0.086	0.065
	(0.239)	(0.153)	(0.146)	(0.160)
MMS*1[Year=2009-11]	2.348***	0.749***	0.146	-0.024
	(0.237)	(0.155)	(0.148)	(0.162)
MMS*1[Year=2012-14]	3.166***	0.828***	0.235	0.015
	(0.227)	(0.153)	(0.151)	(0.168)
MMS*1[Year=2015-18]	3.332***	1.073***	0.532***	0.487***
	(0.212)	(0.142)	(0.139)	(0.154)

Zero-inflated negative binomial estimates. The dependent variable is the number of initiated trials, aggregated at the level of targeted condition (3-digit ICD-9 code), stage, year of trial initiation and novelty category (N=23,760 in Panel A and N=68,574 in Panel B). Displayed coefficients correspond to interactions of the variables displayed in rows with indicators for those displayed in columns. The estimated model includes year, therapeutic area, development stage, and novelty category fixed effects. Robust standard errors are presented in parentheses. *Legend:* $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

Table 5: Part D impacts across recombinant NOVELTY novelty categories (multi-TBA trials only).

	Recombinant NOVELTY category		
	R1	R2	R3
MMS	-1.357***	-0.726***	-0.177
	(0.144)	(0.127)	(0.122)
MMS*1[Year=2004-5]	-0.412*	0.110	0.195
	(0.241)	(0.205)	(0.201)
MMS*1[Year=2006-8]	-0.195	0.090	0.164
	(0.196)	(0.171)	(0.165)
MMS*1[Year=2009-11]	0.306	0.496***	0.212
	(0.193)	(0.173)	(0.172)
MMS*1[Year=2012-14]	1.365***	0.904***	0.272
	(0.182)	(0.169)	(0.171)
MMS*1[Year=2015-18]	1.724***	1.150***	0.328**
	(0.167)	(0.158)	(0.167)

Zero-inflated negative binomial estimates. The dependent variable is the number of initiated trials, aggregated at the level of targeted condition (3-digit ICD-9 code), stage, and year of initiation (N=65,384). Displayed coefficients correspond to interactions of variables displayed in rows with indicators for those displayed in columns. All models include year, therapeutic area, development stage, and novelty category fixed effects. Robust standard errors are presented in parentheses. *Legend:* $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

Table 6: Recombinant novelty effects by molecule size.

	Recombinant Novelty category		
	R1	R2	R3
	A. Trials for small molecules		
MMS	-1.523***	-1.033***	-0.649***
	(0.16)	(0.14)	(0.14)
MMS*1[Year=2004-5]	-0.303	0.191	0.349
	(0.26)	(0.23)	(0.22)
MMS*1[Year=2006-8]	-0.0275	0.203	0.231
	(0.21)	(0.19)	(0.19)
MMS*1[Year=2009-11]	0.334	0.432**	0.345*
	(0.21)	(0.20)	(0.20)
MMS*1[Year=2012-14]	1.395***	1.035***	0.456**
	(0.20)	(0.19)	(0.20)
MMS*1[Year=2015-18]	1.587***	1.162***	0.487**
	(0.18)	(0.18)	(0.20)
	B. Trials for biologicals		
MMS	-0.863**	-1.097***	0.0287
	(0.42)	(0.22)	(0.19)
MMS*1[Year=2004-5]	-0.633	0.363	-0.0426
	(0.42)	(0.33)	(0.31)
MMS*1[Year=2006-8]	-0.566	0.434	0.213
	(0.36)	(0.29)	(0.25)
MMS*1[Year=2009-11]	-0.201	1.179***	0.0105
	(0.34)	(0.26)	(0.24)
MMS*1[Year=2012-14]	0.795**	1.302***	0.0966
	(0.33)	(0.26)	(0.24)
MMS*1[Year=2015-18]	1.135***	1.823***	0.144
	(0.33)	(0.23)	(0.23)

Zero-inflated negative binomial estimates. The dependent variable is the number of initiated trials, aggregated at the level of targeted condition (3-digit ICD-9 code), stage, and year of initiation (N=58,168 in Panel A and N=47,168 in Panel B). Displayed coefficients correspond to interactions of variables displayed in rows with indicators for those displayed in columns. The estimated model includes year, therapeutic area, development stage, and novelty category fixed effects. Robust standard errors are presented in parentheses. *Legend:* $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

Table 7: Recombinant novelty effects by development stage.

	Recombinant novelty category		
	R1	R2	R3
	A. Early stage development (preclinical)		
MMS	-0.976***	-0.361*	-0.0424
	(0.28)	(0.20)	(0.22)
MMS*1[Year=2004-5]	-0.521	-0.0231	0.0194
	(0.35)	(0.30)	(0.31)
MMS*1[Year=2006-8]	-0.525*	-0.427*	-0.140
	(0.30)	(0.26)	(0.26)
MMS*1[Year=2009-11]	-0.0318	0.251	0.245
	(0.31)	(0.26)	(0.28)
MMS*1[Year=2012-14]	1.065***	0.789***	0.242
	(0.29)	(0.26)	(0.26)
MMS*1[Year=2015-18]	1.573***	0.798***	-0.101
	(0.28)	(0.23)	(0.26)
	B. Late stage development (Phases 1-3)		
MMS	-1.772***	-0.966***	-0.0786
	(0.20)	(0.19)	(0.16)
MMS*1[Year=2004-5]	-0.383	0.261	0.187
	(0.32)	(0.27)	(0.25)
MMS*1[Year=2006-8]	0.0770	0.503**	0.244
	(0.25)	(0.23)	(0.21)
MMS*1[Year=2009-11]	0.467*	0.681***	0.0845
	(0.25)	(0.23)	(0.22)
MMS*1[Year=2012-14]	1.552***	1.055***	0.240
	(0.25)	(0.24)	(0.23)
MMS*1[Year=2015-18]	1.665***	1.309***	0.292
	(0.22)	(0.22)	(0.22)

Zero-inflated negative binomial estimates. The dependent variable is the number of initiated trials, aggregated at the level of targeted condition (3-digit ICD-9 code), stage, and year of initiation (N=14,740 in Panel A and N=41,118 in Panel B). Displayed coefficients correspond to interactions of variables displayed in rows with indicators for those displayed in columns. The estimated model includes year, therapeutic area, development stage, and novelty category fixed effects. Robust standard errors are presented in parentheses. *Legend:* $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

Table 8: Scientific novelty and *Top-Seller* status.

Translational novelty F.E.	0.736**		
	(0.331)		
Recombinant novelty F.E.		0.334	
		(0.298)	
1[Multi-TBA product]			0.384***
			(0.116)

Probit estimates. The dependent variable is an indicator for whether each drug product appears in the 2017 *MedAd News* top-sellers ranking (N=1,025 in the first and third columns, and N=387 in the second column). In addition to the displayed variables, all models include fixed effects for the year in which each product was launched to the market, and a dummy variable for whether the product is a biologic. The (product-level) novelty fixed effects entering the specifications are obtained from a regression of (trial-level) novelty scores on product, month, area, and stage fixed effects. Robust standard errors are presented in parentheses. *Legend:* $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

Table 9: Scientific novelty and regulatory designation awards.

	Designation type					
	Novelty	Rare dis.	Novelty	Rare dis.	Novelty	Rare dis.
Translational novelty F.E.	2.287***	3.358***				
	(0.185)	(0.239)				
Recombinant novelty F.E.			2.091***	3.423***		
			(0.260)	(0.389)		
1[Multi-TBA product]					1.141***	1.104***
					(0.041)	(0.034)

Probit estimates. The dependent variable is an indicator for whether each drug product has obtained at least one designation of each kind (N=12,075 for the first two and last two columns, N=3,509 for the third and fourth columns). In addition to the displayed variables, all models include fixed effects for the year in which each product initiated its first clinical trial, and a dummy variable for whether the product is a biologic. The (product-level) novelty fixed effects entering the specifications are obtained from a regression of (trial-level) novelty scores on product, month, area, and stage fixed effects. Robust standard errors are presented in parentheses. *Legend:* $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.1$.

Table 10: Product novelty and patent forward citations.

	Forward citations horizon			
	3-year	5-year	3-year	5-year
Translational novelty F.E.	0.266***	0.202**		
	(0.090)	(0.101)		
Recombinant novelty F.E.			0.122	0.054
			(0.140)	0.158)

OLS estimates. The dependent variable is the logger number of forward citations for a product key patent. Regressions for 3-year citations use data of products whose patents were granted no later than 2014; those for 5-year citations, those granted no later than 2012 (N=5,521, 4,709, 1,645, 1,401, 5,521, and 4,709 respectively for columns one through six). The (product-level) novelty fixed effects entering the specifications are obtained from a regression of (trial-level) novelty scores on product, month, area, and stage fixed effects. Robust standard errors are presented in parentheses. *Legend:* $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

Figure 1: Number of trials by therapeutic area.

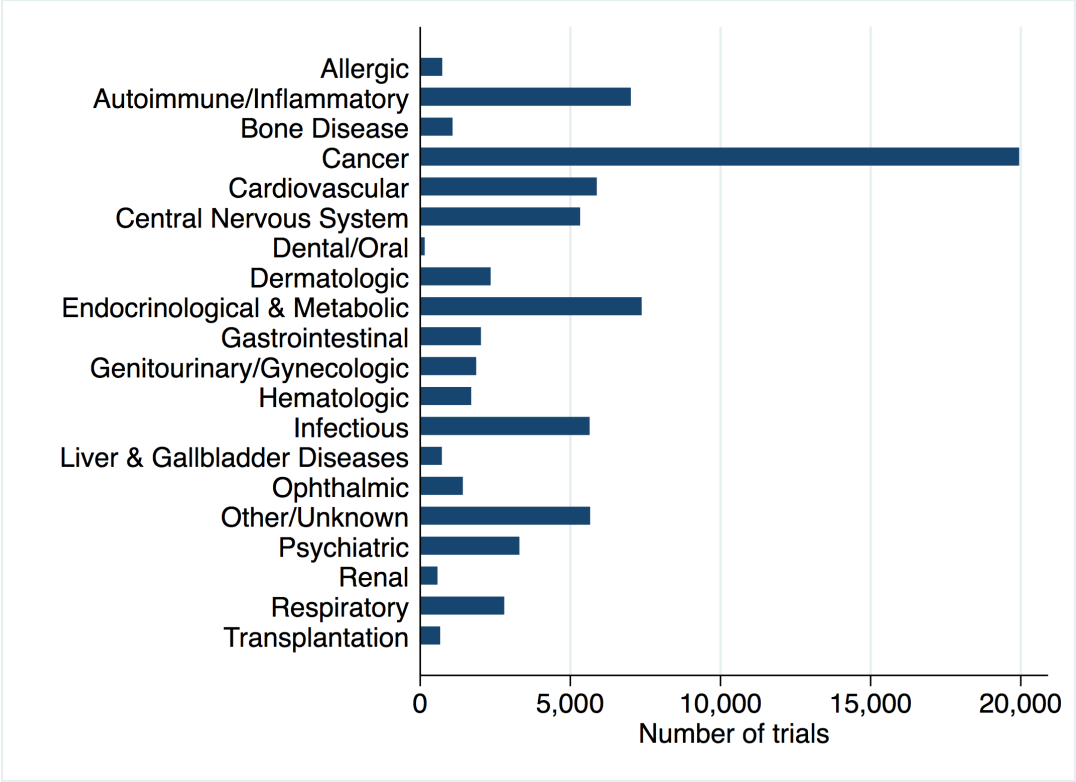


Figure 2: Number of trials by development stage.

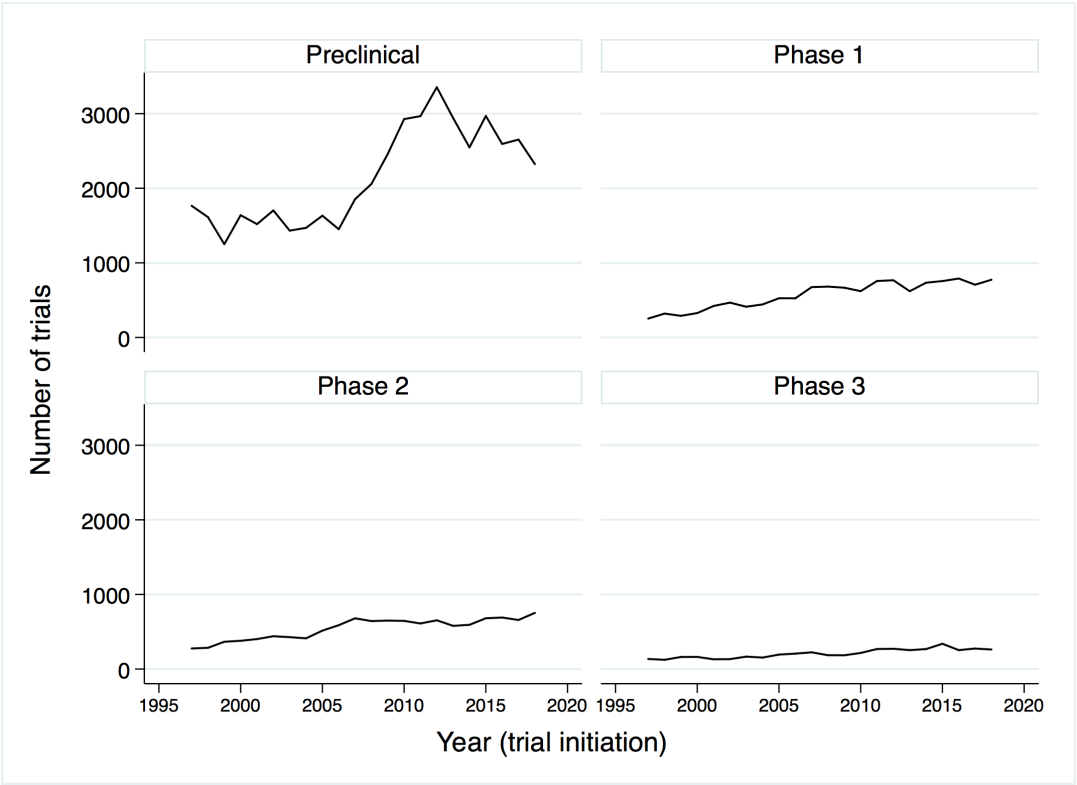


Figure 3: Medicare Market Share (MMS) variation across targeted conditions.

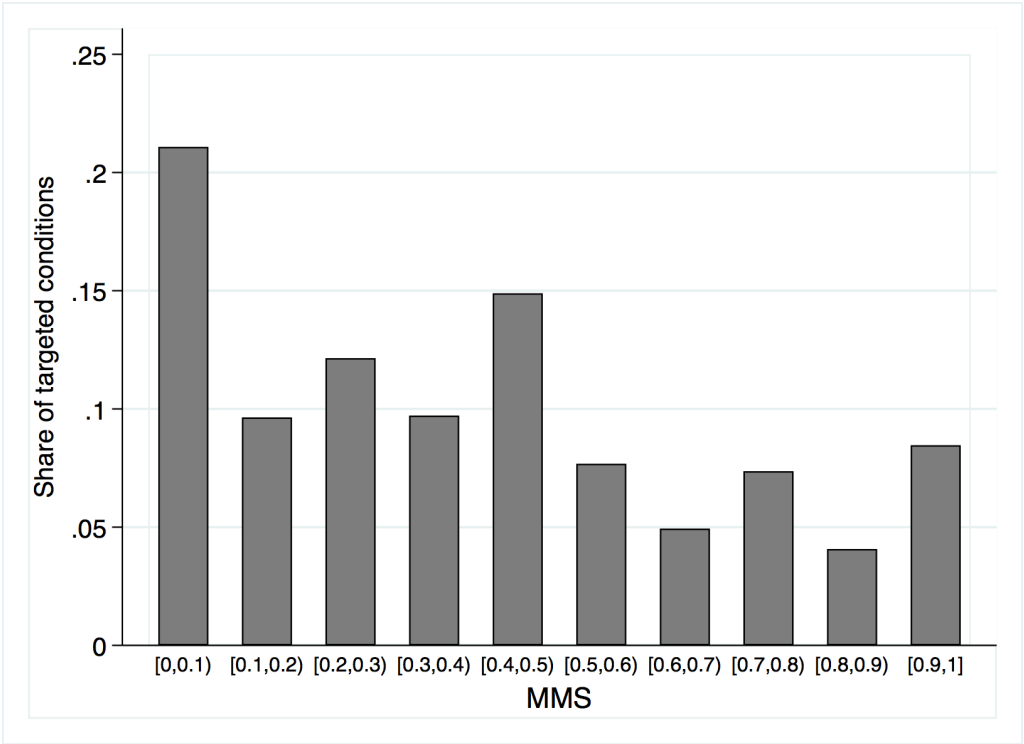


Figure 4: Distributions of novelty scores.

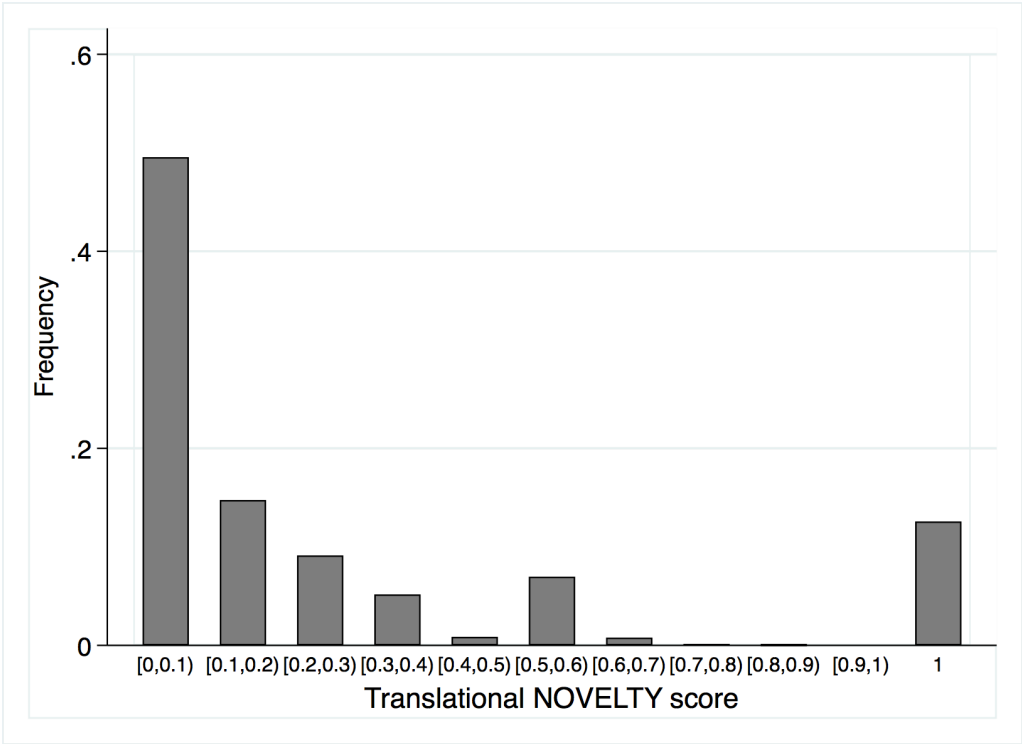


Figure 5: Number of initiated trials, by Medicare orientation and translational NOVELTY category.

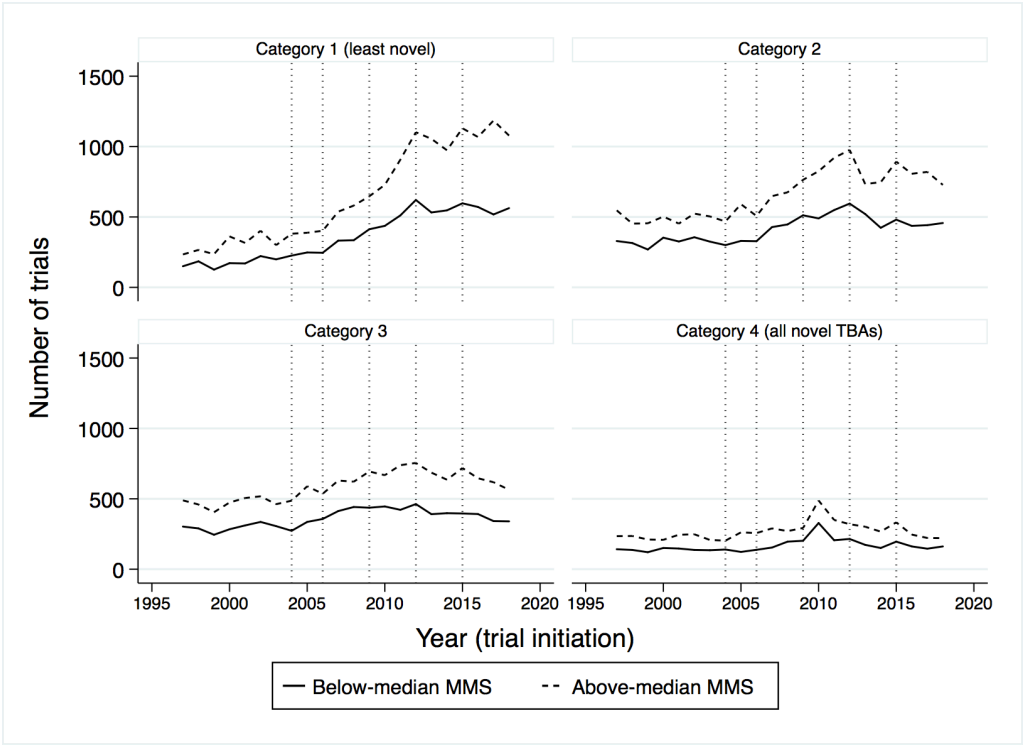


Figure 6: Part D impacts on translational novelty (number of initiated trials in categories 1 and 4).

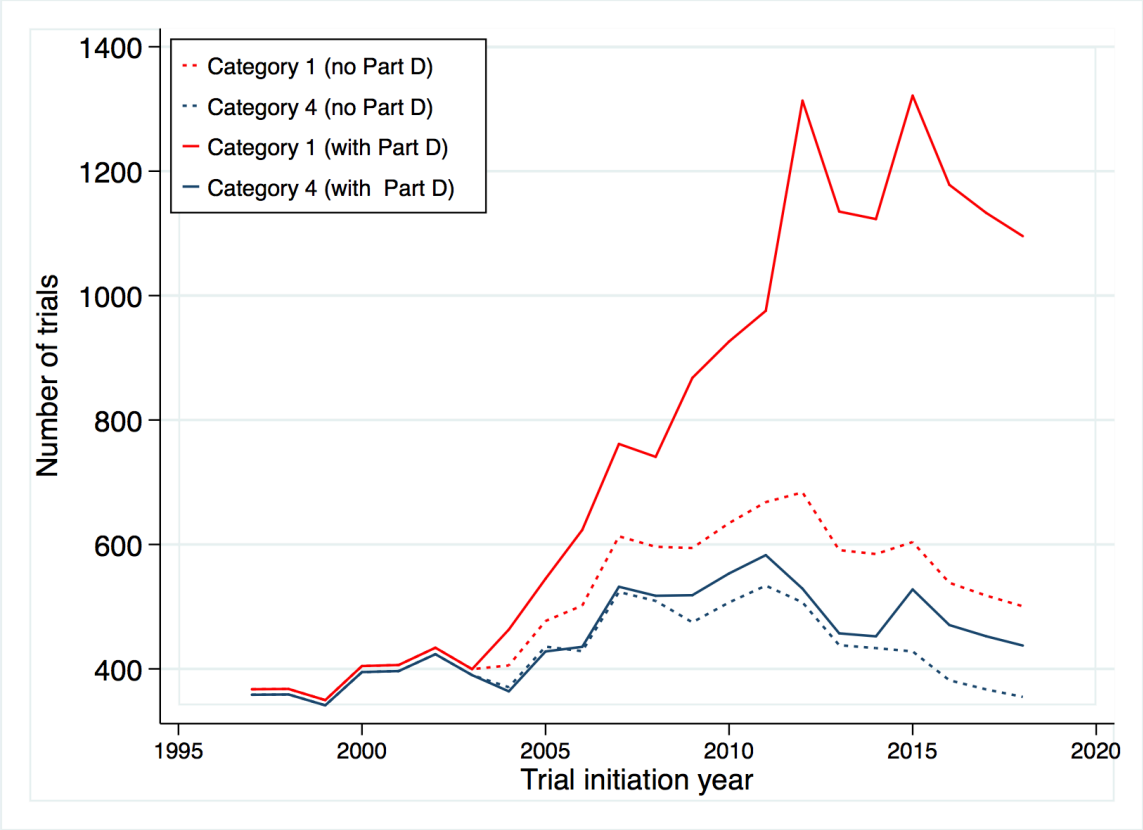


Figure 7: Percentage of trials deploying TBAs with regulatory validation.

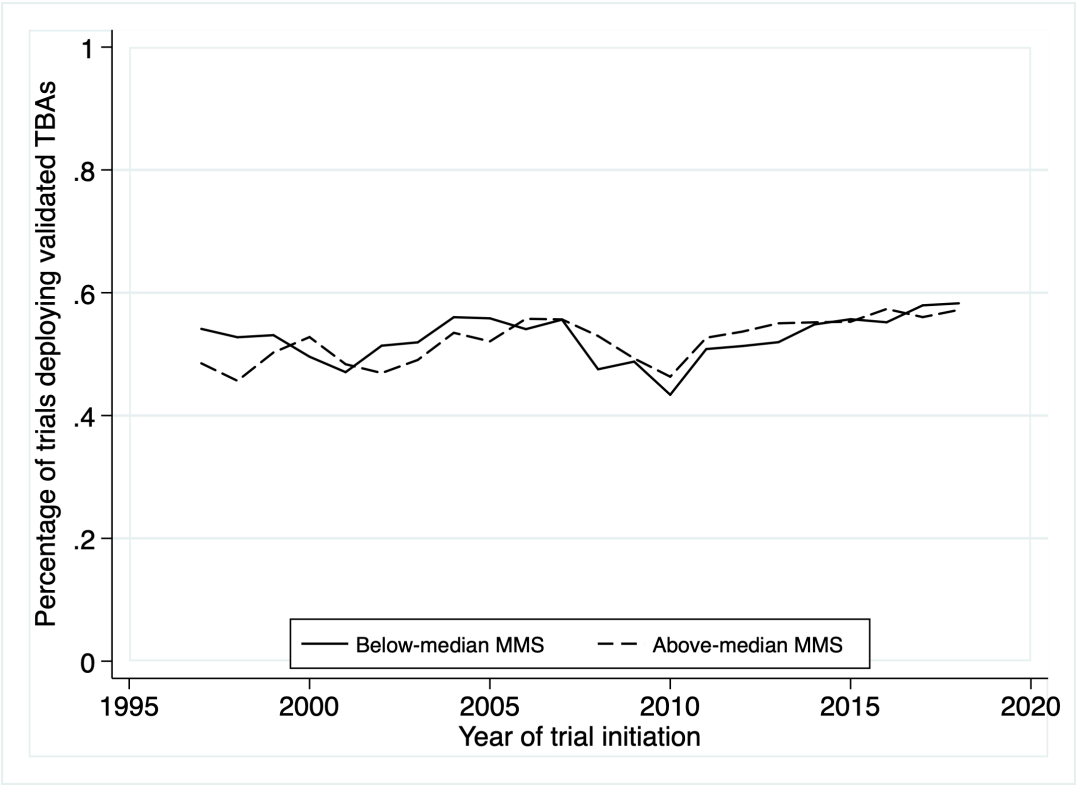


Figure 8: Share of multi-TBA trials, by Medicare orientation.

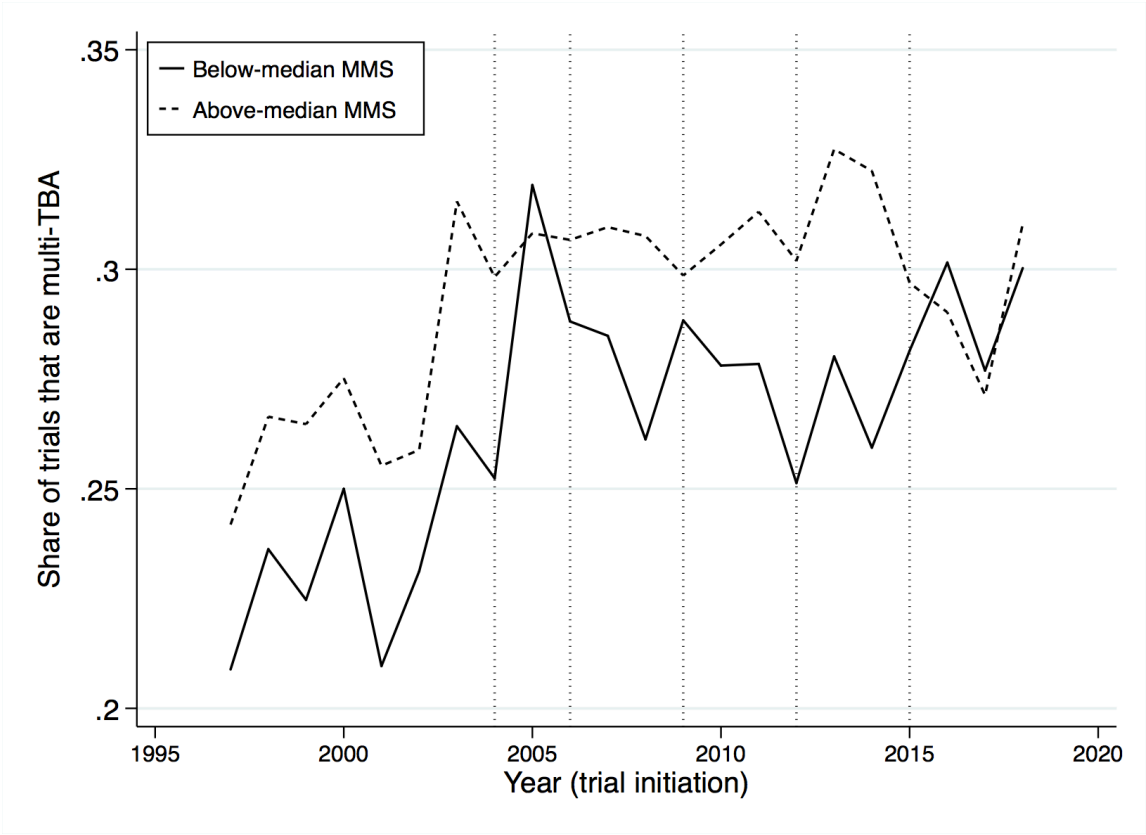
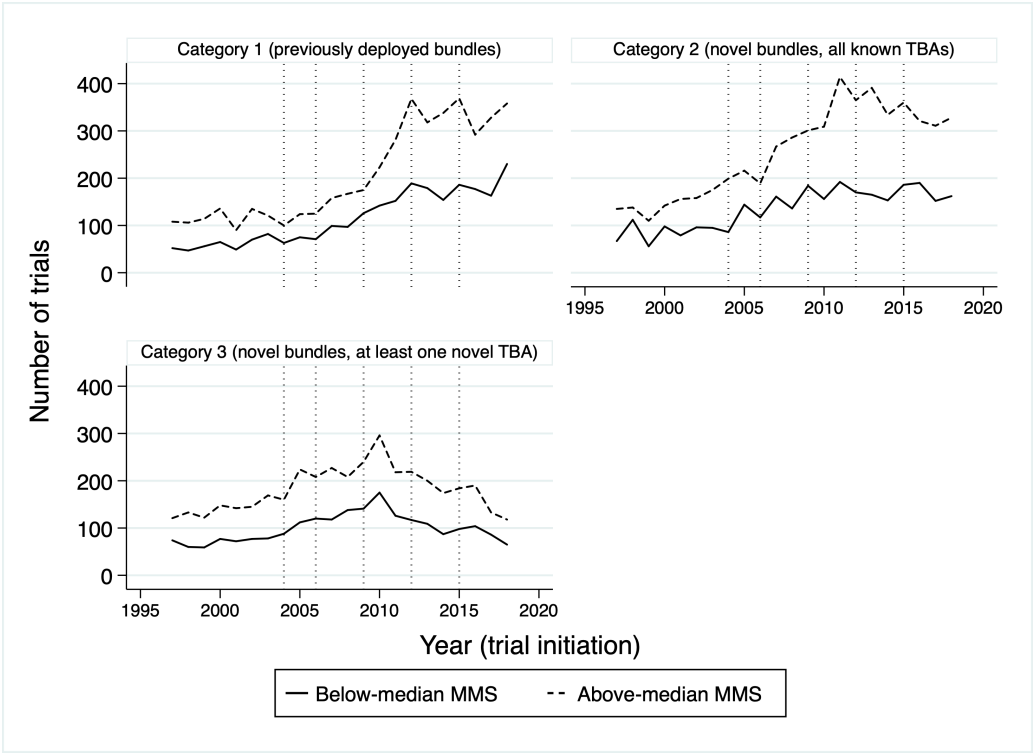


Figure 9: Number of initiated trials, by Medicare orientation and recombinant NOVELTY category (multi-TBA trials only).



Appendix for:
"Expected Profits and the Scientific Novelty of Innovation"

David Dranove, Craig Garthwaite, Manuel Hermosilla

February 2020

A. Is the availability of TBA information correlated with Part D exposure?

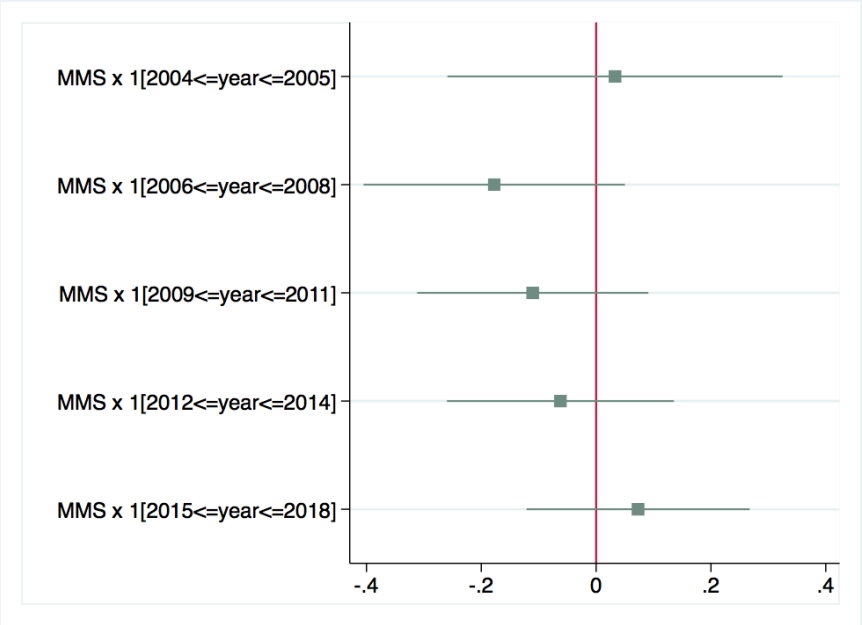
As we note in Section 3.1, TBA information is available for about 60% of the drug products listed in the Cortellis dataset. In this appendix we investigate whether the distribution of missing TBA information may be correlated with a product's exposure to the Part D shock. In particular, we test whether the Part D passage is associated to a higher likelihood that TBA information is available for a product.

To implement our test we consider the 57,902 products tested during the sample period (1997-2018) and which can also be linked to the MMS proxy. As shown by Table 1 of the paper, TBA information is available for about 36,000 of these. Since many of these products have multiple indications and are associated to more than one trial, aggregation is required to format a dataset at the product level. We do so by considering a product's largest MMS score (among all targeted indications) and the earliest date in which development is observed (among all observed trials). With this dataset, we estimate the following specification:

$$y_i = \text{Logit}\Big(\alpha_0 + \alpha_1 \cdot \text{MMS}_i + \sum_p \beta_p \cdot \text{MMS}_i \cdot 1[t(i) \in p] + \eta_{a(i)} + \lambda_{t(i)}\Big),$$

where y_i is an indicator that is activated if TBA information is available for product i. As for specification (2) in the paper, p corresponds to a series of time periods and $t(i)$ corresponds to the earliest year that the earliest trial for product i is observed. Parameters η and λ correspond to area and year fixed effects, respectively. From this specification, we would conclude that exposure to the Part D shock is correlated with the availability of TBA information if estimates for parameters $\{\beta\}_p$ were different from zero. Figure A.1 presents these estimates, along with 95% confidence intervals (from robust standard errors). None of the estimates is statistically significant from zero. These results therefore suggest that a product's exposure to the Part D program is unrelated to whether TBA information is available for it.

Figure A.1: Logit coefficient estimates for a model testing whether Part D exposure influenced the availability of TBA information at the product level.



We plot 95% CI intervals computed from robust standard errors.

B. Replicating main results with a "time-independent" novelty metric

In Section 4.3 we noted that the formulation of our baseline novelty metric could bias our estimates towards finding responses among less novel drugs. This occurs because (i) Part D effects unfold in the latter part of our sample period, and (ii) due to the organic accumulation of TBA deployments, our baseline novelty metric regards trials initiated later in the sample as less novel. Here we introduce an alternative novelty formulation which is "time-independent" in the sense that TBAs' organic accumulation of deployments as time passes does not mechanically determine measured novelty. We label this metric as "percentile novelty" (PNOVELTY).

To formally introduce PNOVELTY, consider again a single-TBA trial i for TBA k with D_k previous deployments. PNOVELTY is defined as:

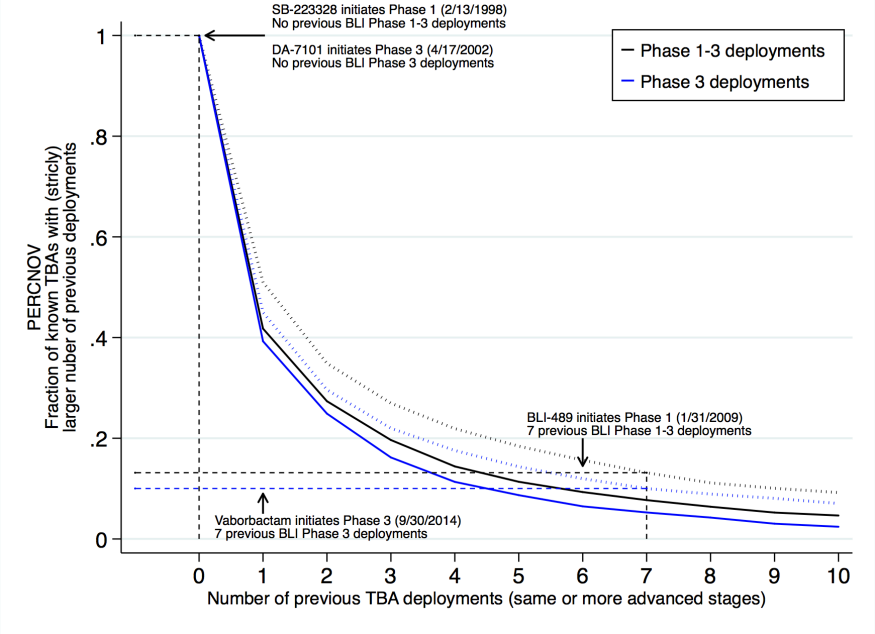
$$\text{PNOVELTY}_i = \frac{\{k' \in \mathcal{K}_i : D_{k'} > D_{k(i)}\}}{|\mathcal{K}_i|} \in [0, 1],$$

where $\mathcal{K}_i$ represents the set of known TBAs for trial i. This set includes all previously deployed TBAs at the same or more advanced stages than that associated to trial i. PERCNOV is thus defined as the fraction of known TBAs with a strictly larger number of previous deployments than k. As with our baseline novelty metrics, PNOVELTY=1 when a trial i deploys a TBA for the first time and decreases with the k's number of previous deployments. In contrast to our baseline metric, PNOVELTY can equal 0, which will occur if k has the largest number of deployments compared to all known TBAs. Given this formulation, near-zero PNOVELTY scores identify the trials that can be deemed as least novel in relation to the current distribution of per-TBA deployments. As with our baseline formulation, for multi-TBA trials PNOVELTY is computed by averaging the right-hand-side of the above expression among all included TBAs.

Figure A.2 further illustrates the computation of PNOVELTY by means of an example. We consider Phase 1 and Phase 3 trials for products that deploy a single TBA, beta lactamase inhibitor (BLI). This TBA was first deployed in Phase 1 by a trial for SB-223328 (SmithKline Beecham) and in Phase 3 by a trial for DA-7101 (Dong-A ST Co). Prior to these trials, BLI had not been previously deployed in relevant stages. Hence, deployment distributions (solid curves) are irrelevant and both trials are scored with PNOVELTY=1. Subsequent trials continued to deploy BLI. One of these was the Phase 1 trial for BLI-489 (Wyeth), which began in January 2009. At the time this trial was initiated, BLI had been deployed by 7 other products in Phases 1 through 3. The deployment distribution measured at this time (black dotted curve) shows that 13 percent of then-known TBAs had more than 7 deployments. Thus, this trial is scored with PNOVELTY=0.13. In the same way, when the Phase 3 trial for Vaborbactan (Rempex) started later in September of 2014, BLI had been deployed by 7 products in Phase 3. In this case, about 10 percent of TBAs previously deployed in Phase 3 had accumulated more than 7 deployments (blue dotted curve). Hence, this trial is scored with PNOVELTY=0.1.

Recall that PNOVELTY and our baseline metric coincide when a TBA is being deployed for the first time (both equal 1). This means that for our analysis of the impacts on translational novelty we should expect the classification from the two metrics to differ in categories T1-T3. Table A.1 shows the distribution of trials under both categorizations. However, given the high correlation between the two metrics (0.97), disagreements are infrequent (i.e., less than 2% of trials). In contrast, since for our analysis of recombinant novelty we aggregate non-novel TBA bundles into a single category (R1), the resulting

Figure A.2: Sample computation of translational PNOVELTY—scoring trials that deploy the TBA "beta lactamase inhibitor."



categorizations are identical. Consequently, in what follows we replicate key results for the analysis of translational novelty only.

Table A.1: Distribution of trials across translational novelty categories.

Categorization using	Categorization using PNOVELTY			
the baseline novelty metric	T1	T2	T3	T4
T1	0.250	0.034	0.008	
T2	0.041	0.235	0.033	
T3	0.000	0.023	0.250	
T4				0.1258

Figure A.3: Number of initiated trials, by Medicare orientation and ABSNOV novelty (PNOVELTY) category. (Correlate to Figure 5.)

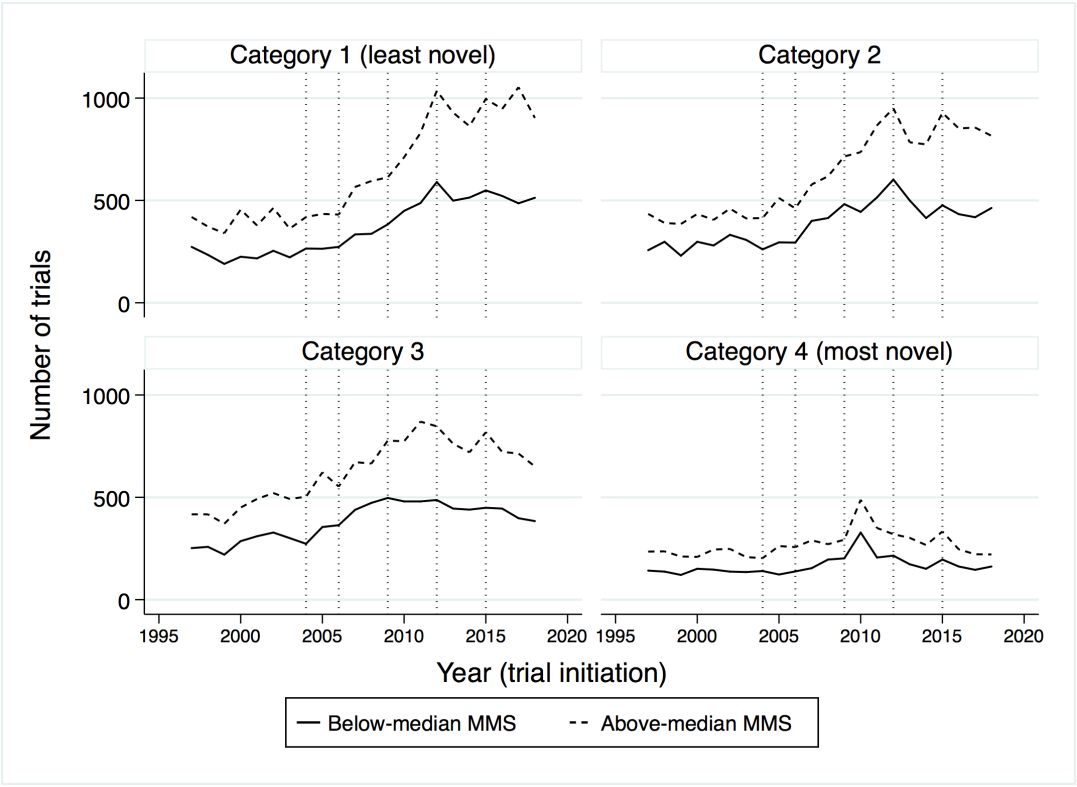
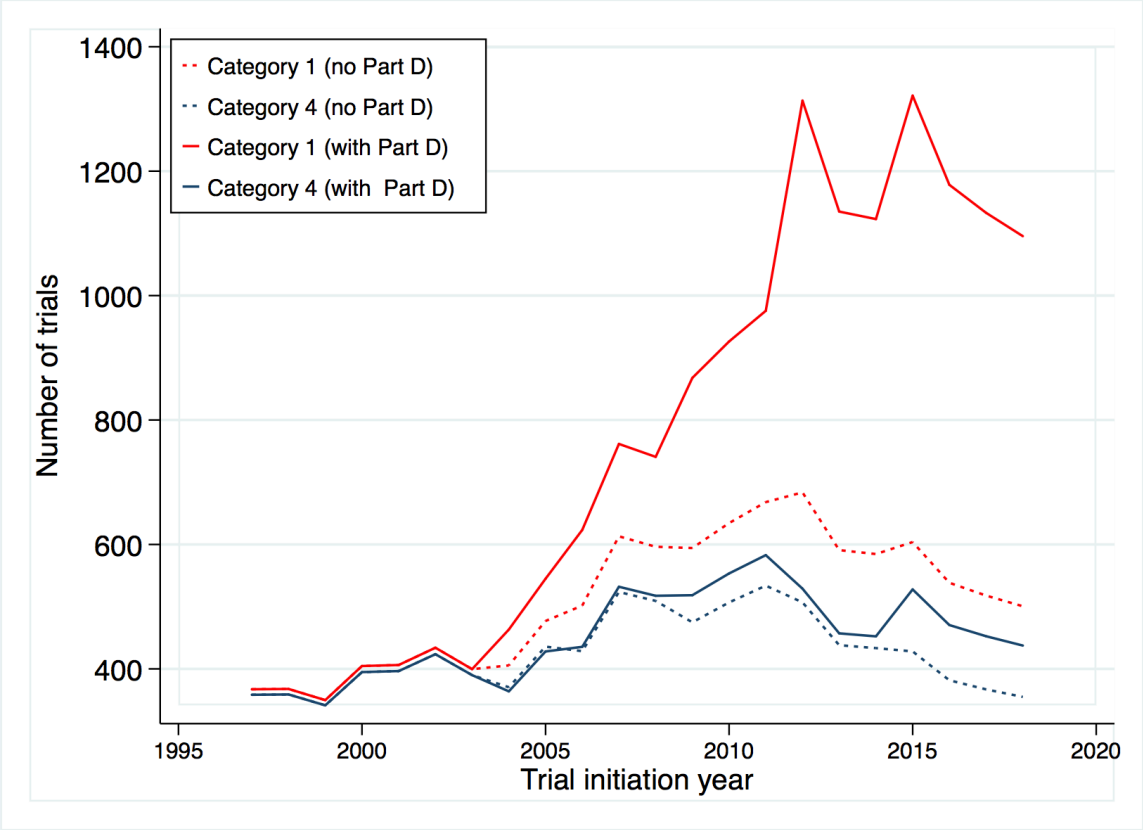


Table A.2: Part D impacts across translational novelty (PNOVELTY) categories. (Correlate to Table 2.)

	Translational novelty category			
	T1	T2	T3	T4
MMS	-0.663***	-0.260***	-0.035	0.314***
	(0.101)	(0.089)	(0.087)	(0.094)
MMS*1[Year=2004-5]	0.082	-0.007	0.187	-0.095
	(0.167)	(0.153)	(0.142)	(0.162)
MMS*1[Year=2006-8]	0.160	0.065	0.187	-0.039
	(0.143)	(0.128)	(0.123)	(0.132)
MMS*1[Year=2009-11]	0.370***	0.311**	0.411***	0.091
	(0.143)	(0.131)	(0.126)	(0.143)
MMS*1[Year=2012-14]	0.880***	0.582***	0.505***	-0.022
	(0.138)	(0.125)	(0.122)	(0.136)
MMS*1[Year=2015-18]	1.152***	0.949***	0.870***	0.344***
	(0.124)	(0.111)	(0.111)	(0.121)

Zero-inflated negative binomial estimates. The dependent variable is the number of trials initiated each, aggregated at the level of targeted condition (3-digit ICD9 code), stage, year of trial initiation and novelty category (N=103,840). Displayed coefficients correspond to interactions of the variables displayed in rows with indicators for those displayed in columns. The estimated model includes year, therapeutic area, development stage, and novelty category fixed effects. Robust std. errors are presented in parentheses. *Legend:* $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.1$.

Figure A.4: Part D impacts on translational (PNOVELTY) novelty (number of initiated trials in categories 1 and 4). (Correlate to Figure 6.)



C. Falsification and robustness for main results

While the graphical relationships and regression estimates suggest a change in the composition of clinical trial activity following Part D's implementation, there could be a concern that this is a continuation of a secular trend towards drugs targeting conditions of the elderly—perhaps in response to a broader demographic shift caused by the aging baby boomer population. To the extent that such a trend is correlated with MMS, our estimates might be picking up this pre-existing activity. To address this issue, we first revisit Figure 5, which shows the number of clinical trials per year based on whether the indication targets a disease that has an MMS above or below the median level. If secular trends drove the estimates in Table 2, then this should be observable in the data prior to 2003. This not the case. Prior to the passage of Medicare Part D (demarcated by the first dashed vertical line), there was very little difference in level or trend based on MMS. To further address this concern, we estimate a placebo specification of equation (2) using 1997-2003 data, which includes an interaction of MMS with an indicator for the 2000-3 period. Panel A Table A.3 contains the results from this specification. The coefficient on the interaction term measures the change in research activity between 2000 and 2003 compared to earlier time periods. Given that there was little clear evidence that Congress would develop and pass a prescription drug benefit, we do not expect any pre-passage anticipatory behavior. However, if our main estimates are simply the result of a gradual shift in the market, we should find generally similar results from this specification. Across all novelty categories, the estimates on the interaction term are small in magnitude and statistically insignificant. This supports a causal interpretation of our estimates of Table 2 and Figure 5.

A second potential concern is that the prominence of clinical trials for cancer treatments means that our estimates primarily reflect changes in the science of developing oncology products that was coincident with the passage of Part D. A related concern is that a large fraction of cancer drugs are covered under Medicare Part B and therefore the creation of Part D may not represent a substantial profit shock. This concern is mitigated to some degree by the increasing prevalence of oral chemotherapy products that would be affected by Part D. To allay these concerns about the role of oncology products in our results, we re-estimate our main regressions excluding clinical trials for cancer treatments. Estimates are presented in Panel B of Table A.3. These display the same general patterns of our main estimates, except for they are somewhat smaller in magnitude. This smaller magnitude may reflect the fact that oncology products were "protected classes" under Medicare Part D, which increases the expected profitability of these products after the passage of Part D. Overall, these results suggest that the pattern of our main findings with respect to novelty are not driven by cancer-targeted innovation.

Lastly, recall that our approach to measuring the translational novelty of multi-TBA trials is to average the individual novelty scores of each TBAs included in the bundle. This approach could obscure some aspects of novelty, particularly for bundles that combine very novel and very common TBAs into a single product. Such products may entail a high degree of technological uncertainty due to the presence of a novel TBA, but our averaging strategy excludes them from the most-novel category. To investigate the extent to which this drives our results, we re-estimate our models on the full sample of trials, but excluding those with multi-TBAs. Resulting estimates are presented in Panel A of Table A.4. In Panel B, we present the results obtained when we aggregated novelty considering the minimum novelty score across all included TBAs, and in Panel C, when we considered the maximum one. The latter aggregation shifts a number of trials into the most-novel category, which is why we observe that the intensity of Part D effects

is somewhat (although quite mildly) shifted towards more novel categories. Results are otherwise very similar to our main estimates of Table 2, which suggests that aggregation issues have little influence on our conclusions.

Table A.3: Falsification and robustness of main translational novelty effects.

	Translational novelty category			
	T1	T2	T3	T4
	A. Falsification (1997-2003 data)			
MMS	-1.111**	-0.350	0.093	0.388
	(0.497)	(0.442)	(0.418)	(0.362)
MMS*1[Year=2000-3]	0.540	0.039	-0.021	-0.208
	(0.376)	(0.179)	(0.179)	(0.168)
	B. Cancer trials dopped			
MMS	-1.095***	-0.119	0.218**	0.321***
	(0.122)	(0.092)	(0.091)	(0.103)
MMS*1[Year=2004-5]	0.285	0.103	-0.042	-0.043
	(0.181)	(0.162)	(0.149)	(0.175)
MMS*1[Year=2006-8]	0.510***	0.077	-0.072	0.025
	(0.156)	(0.136)	(0.131)	(0.143)
MMS*1[Year=2009-11]	0.872***	0.264*	-0.041	0.113
	(0.158)	(0.140)	(0.135)	(0.152)
MMS*1[Year=2012-14]	1.559***	0.369***	0.016	0.042
	(0.150)	(0.133)	(0.133)	(0.145)
MMS*1[Year=2015-18]	1.723***	0.553***	0.302**	0.365***
	(0.140)	(0.125)	(0.125)	(0.137)

Zero-inflated negative binomial estimates. The dependent variable is the number of trials initiated each, aggregated at the level of targeted condition (3-digit ICD9 code), stage, year of trial initiation and novelty category (N=33,040 in Panel A and N=95,304 in Panel B). Displayed coefficients correspond to interactions of the variables displayed in rows with indicators for those displayed in columns. The estimated model includes year, therapeutic area, development stage, and novelty category fixed effects. Robust std. errors are presented in parentheses. *Legend:* $^{*}p < 0.1$,** $p < 0.05$,*** $p < 0.1$.

Table A.4: Robustness of main translational novelty effects (continued).

	Translational novelty category			
	T1	T2	T3	T4
	A. Multi-TBA trials dropped			
MMS	-1.397***	-0.416***	-0.010	0.060
	(0.127)	(0.094)	(0.094)	(0.102)
MMS*1[Year=2004-5]	0.244	0.061	0.038	0.001
	(0.187)	(0.168)	(0.157)	(0.167)
MMS*1[Year=2006-8]	0.572***	0.123	-0.027	0.096
	(0.162)	(0.140)	(0.138)	(0.142)
MMS*1[Year=2009-11]	0.891***	0.347**	0.102	0.243*
	(0.158)	(0.137)	(0.138)	(0.147)
MMS*1[Year=2012-14]	1.444***	0.474***	0.342**	0.180
	(0.151)	(0.134)	(0.137)	(0.145)
MMS*1[Year=2015-18]	1.828***	0.902***	0.803***	0.611***
	(0.140)	(0.122)	(0.122)	(0.129)
	B. Aggregation via minimum individual novelty			
MMS	-1.413***	-0.255***	0.265***	0.270***
	(0.122)	(0.091)	(0.087)	(0.096)
MMS*1[Year=2004-5]	0.353*	0.266*	-0.035	-0.028
	(0.182)	(0.157)	(0.147)	(0.167)
MMS*1[Year=2006-8]	0.660***	0.190	-0.022	0.047
	(0.153)	(0.132)	(0.128)	(0.136)
MMS*1[Year=2009-11]	1.110***	0.464***	0.073	0.208
	(0.154)	(0.134)	(0.130)	(0.146)
MMS*1[Year=2012-14]	1.804***	0.557***	0.229*	0.124
	(0.148)	(0.129)	(0.127)	(0.139)
MMS*1[Year=2015-18]	2.143***	0.918***	0.533***	0.499***
	(0.136)	(0.117)	(0.116)	(0.126)
	C. Aggregation via maximum individual novelty			
MMS	-1.230***	-0.267***	0.003	0.321***
	(0.117)	(0.089)	(0.088)	(0.088)
MMS*1[Year=2004-5]	0.300*	-0.029	0.029	0.159
	(0.175)	(0.156)	(0.146)	(0.149)
MMS*1[Year=2006-8]	0.452***	0.049	-0.049	0.182
	(0.152)	(0.130)	(0.126)	(0.124)
MMS*1[Year=2009-11]	0.821***	0.286**	0.203	0.267**
	(0.150)	(0.131)	(0.130)	(0.130)
MMS*1[Year=2012-14]	1.499***	0.528***	0.402***	0.206
	(0.145)	(0.126)	(0.125)	(0.127)
MMS*1[Year=2015-18]	1.881***	0.858***	0.799***	0.502***
	(0.132)	(0.113)	(0.114)	(0.118)

Zero-inflated negative binomial estimates. The dependent variable is the number of trials initiated each, aggregated at the level of targeted condition (3-digit ICD9 code), stage, year of trial initiation and novelty category (N=99,792 in Panel A and N=103,840 in Panels B and C). Displayed coefficients correspond to interactions of the variables displayed in rows with indicators for those displayed in columns. The estimated model includes year, therapeutic area, development stage, and novelty category fixed effects. Robust std. errors are presented in parentheses. *Legend:* $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.1$.