The Effects of Tax Credits for Rare Disease R&D: Evidence from the Tax Cuts and Jobs Act

Ben Berger*

February 7, 2025

Abstract

Incentivizing private research and development for drugs that treat rare diseases presents significant economic and policy and challenges. Countries have enacted policies to stimulate the development of medications for these conditions; however, the effectiveness of such policies remains largely unexplored. In this study, I explore the impacts of a reduction of the Orphan Drug Credit, a US federal tax incentive for R&D spending in the clinical stage of rare disease drug development. I show that this reduction disproportionately increased the after-tax cost of R&D for firms with taxable income in the US and that this resulted in a modest decline in the utilization of US sites in rare disease clinical trials. Nonetheless, there is no evidence to suggest that the reduction in the credit adversely affected the overall number of clinical trials for rare diseases. These findings highlight practical weaknesses to the use of R&D credits to stimulate innovation in the biopharmaceutical industry, including regulation of market entry and the global scope of clinical research activity. However, I conclude that tax credits for rare disease drug development may nonetheless improve domestic access to experimental treatments by onshoring clinical research.

* Contact: b.a.berger.94@gmail.com.

I am grateful to my advisors Amitabh Chandra, David Cutler, and Mark Shepard for their wisdom and direction. I would also like to thank Ariel Dora Stern for her invaluable advice. I am also thankful to participants at the Harvard Kennedy School Economics and Social Policy Seminar, the 2022 Global Conference on Regulatory Science organized by the Harvard-MIT Center for Regulatory Science, and the 2022 and 2023 NBER Summer Institute Innovation Boot Camps.

Rare diseases adversely impact the health of affected patients, reducing life expectancy, causing severe pain, impairing cognitive function and motor skills, and worsening mental health. Evidence is also emerging that these diseases have large aggregate economic impact. A recent study found that 379 rare diseases, affecting a total of 15.5 million Americans, cost almost $1 trillion in 2019 (Yang et al. 2022), more than a quarter of the cost of all chronic diseases.[1] Thus, while rare diseases may individually make up only a tiny fraction of the cost of disease, together they account for over one quarter of the cost.

Patients have looked towards development of novel therapies to more effectively treat rare diseases. However, the rare nature of these diseases presents a fundamental economic challenge for incentivizing private research and development. Like treatments for more common diseases, regulators require evidence of the safety and efficacy of rare disease therapies generated in clinical trials with human subjects. These trials have high rates of failure, and firms considering starting trials can expect to spend over $100m (DiMasi et al. 2016).[2] Incurring these costs can be profitable for firms because positive trial findings provide scientific rationale for regulators to approve drugs for sale. Moreover, patents and other forms of legal exclusivity grant monopoly rights to firms developing branded drugs, meaning that firms can expect to earn super-normal profits upon market entry. However, small populations of rare disease patients have historically reduced the profits that firms expect to reap from developing therapies for these diseases. Lackluster profits weaken incentives to develop rare disease therapies and contribute to the lack of options for effective medical care for rare diseases.[3]

Governments have implemented various policies to encourage development of rare disease therapies, commonly referred to as "orphan drugs." The first such law was the Orphan Drug Act (ODA), passed by the US Congress in 1983. Under ODA, drugs desig-

[1]Chronic diseases are estimated to cost Americans $3.7 trillion per year (Waters & Graf 2018).
[2]Firms and other research-funding institutions including universities and government incur additional costs from basic science, preclinical animal studies, and post-market monitoring.
[3]Studies have consistently found that market expansions increase development of biopharmaceuticals and conversely that contractions depress development (Acemoglu & Linn 2004, **?**, Dubois et al. 2015).

nated orphan by the US Food and Drug Administration (FDA) are eligible for tax credits for clinical R&D, seven years of market exclusivity, clinical trial grants, and user fee waivers. These policies are intended to incentivize private R&D both by "pushing" drug development through reducing the cost of R&D and "pulling" drugs to market by increasing the benefits of market entry (Gamba et al. 2021). Passage of ODA in 1983 was followed by adoption of similar orphan drug laws in Japan in 1993 and the European Union in 2000 (Bagley et al. 2019).

The provisions of ODA cost the US government billions of dollars annually,[4] and yet the evidence that these policies have coaxed firms into developing rare disease therapies is limited.[5] There is even less evidence on which (if any) of these policies have worked.[6] Thus, in this paper, I focus on evaluating the impact of the Orphan Drug Credit (ODC), one of the key provisions of ODA. ODC allows firms conducting clinical research on rare disease drugs in the United States to reduce their US federal tax bill, effectively reducing the after-tax cost of US R&D. I first develop a simple model of international biopharmaceutical development in which the costs of R&D spending are impacted by the local tax regime, including the corporate tax rate and available R&D credits. I show that ODC substantially reduces the cost of US R&D for established biopharmaceutical firms, which have approved products and flows of US taxable income. In contrast, tax credits have little inherent value to startups, which face significant risk of never having a tax bill to offset.

[4]The US Treasury estimates that orphan drug tax credits reduced government revenue by $2.7 billion in 2023 (U.S. Department of the Treasury 2023). Orphan drug exclusivity reduces generic competition for rare disease therapies, increasing the costs faced by public payers including Medicare and Medicaid. Waivers for user fees, which are a critical component of FDA funding, cost the agency on the order of $3 million per New Drug Application (Center for Drug Evaluation and Research 2023).

[5]Yin (2008) finds that after passage of the Orphan Drug Act, the growth rate of clinical trials for rare diseases was higher than the growth rate of non-rare diseases. Yin's difference-in-differences model implies that ODA increased orphan drug clinical trial starts by 69 percent. However, this finding rests upon a parallel trends assumption that may be untenable: that rare disease trials would have grown at the same rate as non-rare disease trials in the absence of ODA. If other factors such as the expansion of health plans' pharmacy benefits or changes in basic science would have led to higher growth of rare disease trials in the absence of ODA, then this methodology would tend to overestimate the impact of ODA.

[6]Because Yin (2008) uses the passage of ODA for identification, his difference-in-differences design cannot separately estimate the effects of its various provisions (i.e. tax credits, exclusivity, research grants, and PDUFA waivers).

To test the incidence and effects of ODC, I combine publicly-accessible data sources to assemble a dataset of clinical trials for rare and non-rare diseases and determine whether their lead sponsors are established US drugmakers in the sense of having one or more approved drugs at the time of trial submission to a major US-based clinical trial registry. I show that 45 percent of rare disease trials are sponsored by established firms (compared to 57 percent of non-rare disease trials), indicating a high level of heterogeneity in firms' exposure to ODC. I then leverage a reduction of the ODC rate resulting from the Tax Cuts and Jobs Act of 2017 (TCJA) to estimate a triple differences model, which acts as an empirical analogue of the theoretical model. I show that the ODC reduction led established firms to modestly reduce their use of US sites in rare disease clinical trials. Nevertheless, I find no evidence that the ODC reduction decreased the overall number of rare disease trials in the six years following passage of TCJA. These results suggest that while ODC incentivizes firms to increase demand for US R&D inputs, the credit primarily leads to cross-country substitution of R&D inputs for use in inframarginal projects.

These results cast doubt on the effectiveness of ODC for bringing rare disease drugs to market. As the cost of ODC to taxpayers is expected to grow at a rate of 20 percent per annum, considering alternative approaches to incentivizing rare disease drug development would benefit both rare disease patients and taxpayers. The lack of a response to the ODC reduction also begs the question: what has driven the general increase in rare disease drug development since passage of ODA in 1983? Other provisions of ODA, particularly orphan drug exclusivity, have been cited by advocates as key incentives driving continued development of new and innovative therapies (National Organization for Rare Disorders 2022). This additional period of market exclusivity has been shown to on average add 0.8 years of additional exclusivity to drugs' effective market life (Seoane-Vazquez et al. 2008). This is a non-trivial incentive, but is unlikely to explain nearly 600 percent growth in rare disease trials in the decade following passage of ODA alone. Early explosive growth and more recent waves of rare disease drug development must be viewed in

the context of evolving knowledge of the origins of disease, advances in clinical trial design and diagnostic technology, and the formation of patient advocacy groups and rare disease consortia.

These findings also highlight broader limitations to the use of R&D tax credits, which are considered by leading scholars to be among the most effective tools for incentivizing R&D spending and innovation (Bloom et al. 2019). When market entry is costly and/or highly regulated, non-refundable credits disproportionately benefit incumbents and may only weakly incentivize R&D spending by prospective entrants. Moreover, in industries with global markets for R&D inputs, the ability of a single country's research-subsidizing policies to lead to development and commercialization of new innovative products is limited by the substitutability of foreign inputs. In an age of renewed interest in industrial policy, it is critical for policy makers to evaluate the conditions under which R&D credits are most likely to truly generate product market innovation.

In Section 1, I provide background on the biopharmaceutical research and development pipeline as well as the particular challenges faced in development of rare disease therapies and the ways in which the Orphan Drug Act has attempted to help firms overcome them. In Section 2, I introduce the model of drug development and highlight how TCJA differentially impacted incentives for established firms and startups. In Section 3, I discuss the data used to assemble my sample of clinical trials, and then in Section 4 I introduce my methodology for using that data to test the model's predictions and present the results. I conclude in Section 5 with a discussion of the limitations of this study and directions for future research.

1 Background

Biopharmaceutical research and development is broadly split into a preclinical stage, in which firms test the viability of new medicines in animal models, and a clinical stage, in

which firms conduct human clinical trials to evaluate the safety and efficacy of compounds that show promise in preclinical studies. Because decisions regarding preclinical R&D depend on the cost of downstream clinical R&D (i.e. firms do not do preclinical R&D if they expect a therapy will not be profitable conditional on reaching the clinical stage) and because rare disease policy has historically emphasized incentivizing clinical R&D, I focus on the latter stage of development.

In the clinical stage of biopharmaceutical R&D, firms test the safety and efficacy of candidate drugs that have made it past pre-clinical studies. The clinical stage begins with one or more studies establishing drugs' safety profiles. Safety studies, also known as Phase I trials, typically do not involve control groups and enroll relatively few patients. If Phase I trials evidence a positive safety profile, firms may proceed to Phase II trials, which are the first trials intended to establish efficacy (whether drugs actually improve health). These trials may or may not randomly assign some patients to a control group that receives placebo or the standard of care. These trials usually provide the first signals of efficacy in humans before transitioning to large, placebo-controlled, and very costly Phase III trials, which confirm clinical value.

Evidence of safety and efficacy generated in human clinical trials is a requirement for marketing authorization in many countries, including the United States, but the R&D expenditure required to complete clinical trials combined with their low success rate make clinical studies significant obstacles for biopharmaceutical commercialization. DiMasi et al. (2016) surveyed firms and found that expected clinical costs capitalized to the time of approval were $172.7m ($261.2m in 2023 dollars), while the rate of successful commercialization was only 12%. Firms considering initiating clinical trials must expect to at least recoup these costs in expectation, implying that the average drug must generate over $2 billion in lifetime profits upon successful approval.

Meeting a threshold of profitability is comparatively challenging for drugs intended to treat rare diseases because market size is a critical component of lifetime profits. Suppose

lifetime after-tax profits discounted to the time of marketing approval (T_m) are given by $nm \sum_{t=T_m}^{T_g-1} \exp((g-r)(t-T_m))$ where nm is market size, which is the product of n, the global number of patients with the targeted disease in the period of approval, and m, the average after-tax income per patient. g is a fixed population growth rate, r is the firm's cost of capital, and T_g is the year of first generic entry. Small market sizes, owing either to few patients (small n) or little ability or willingness for patients to pay (small m) shrink lifetime profits. If market size is small enough, expected lifetime profits fall below the expected cost of clinical studies and firms forgo development. All else equal, the rarity of a disease like Duchenne Muscular Dystrophy (DMD) with a global prevalence of 2.8 cases per 100,000 persons ($n \approx 224{,}000$), generates weaker incentive to develop therapies than diabetes, with 8,400 cases per 100,000 persons ($n \approx 451{,}000{,}000$).[7,8]

Recognizing that small n hinders drug development, a key project of rare disease policy has been augmenting profits and reducing development costs to incentivize more development of rare disease therapies. The Orphan Drug Act's provisions do this primarily by affecting the timing of generic entry, as in the case of Orphan Drug Exclusivity,[9] and subsidizing development costs, as in the case of ODC. In order to benefit from any of the perks of ODA, firms must first seek an orphan drug designation from FDA, which they may do at any time, even prior to obtaining permission to conduct clinical studies. Doing so requires submitting an application to FDA including information on the disease targeted for treatment, diagnosis, or prevention, scientific rationale that the drug demon-

[7]Estimating the global population of DMD patients assumes a global population of 8 billion. Estimates of DMD and diabetes prevalence from Crisafulli et al. (2020) and Cho et al. (2018).

[8]On the other hand, per patient profits are often higher for rare disease drugs. Because rare diseases tend to have few treatments, an innovator firm bringing a new rare disease drug to market can expect to face few competitors and inelastic consumer demand and thus can profitably charge high prices per patient. For example, in 2023, FDA approved Elevidys, the first gene therapy for DMD; its manufacturer Sarepta Therapeutics launched the drug at a price of $3.2 million per patient. Unlike traditional pharmaceuticals, gene therapies are characterized by high production costs; however, even if the per patient manufacturing cost were generously estimated at $1 million, and Sarepta allowed for discounts and paid rebates of 20% of the purchase price, this still leaves Sarepta with over $1.5 million in gross profit per patient (Fidler 2023, Harris 2019).

[9]Orphan Drug Exclusivity entitles manufacturers of designated orphan drugs to 7 years without generic entry; however, because exclusivity is use-specific and overlaps with patent protection, the impact of orphan exclusivity on generic entry is ultimately an empirical question.

strates promise, and evidence that the US population of the disease is sufficiently small for the disease to be considered rare. ODA's definition of a rare disease includes diseases that afflict 200,000 or fewer persons in the US or diseases with "no reasonable expectation of recovering the cost of [development] (Internal Revenue Service 2018)."

From passage of ODA in 1983 through 2017, ODC allowed firms to claim a credit for 50 percent of qualified clinical testing expenses (QCTEs) for designated orphan drugs. QCTEs include 100 percent of both in-house and contract research expenses for clinical testing of rare diseases occurring between the dates of orphan designation and FDA approval. The credit also only applies to expenses for testing in the United States unless the US population is insufficiently large to produce scientific data (Internal Revenue Service 2018). The ODC rate must first be reduced by the top corporate tax rate before claiming it to account for deductions, and while clinical expenses for any drug may be used to claim the Credit for Increasing Research Activities, commonly known as the R&D Tax Credit, firms cannot claim both credits using the same expenses. Lastly, because firms may not have sufficient taxable income to use the entirety of their credits in the year they claim them, they may roll those credits back 1 year or forward up to 20 years.

ODC persisted with the same set of provisions until the Tax Cuts and Jobs Act was signed into law on December 22, 2017, launched a sweeping overhaul of the US tax system. Among its many provisions, TCJA instituted a single corporate tax rate of 21 percent and reduced the ODC rate from 50 percent of QCTEs to 25 percent. Figure 1 shows that this had a noticeable impact on tax expenditures, the loss of tax revenue attributable to ODC. After TCJA, the US Treasury revised its estimates of tax expenditures from the Orphan Drug Credit downwards an average of 53.8 percent. These calculations, however, do not elucidate the incidence of ODC nor do they factor in firms' behavioral responses. In the next section, I consider both by developing a model that structures how TCJA heterogeneously impacted the cost of biopharmaceutical R&D in the US and outlines the factors that contribute to firms' responses to the new tax regime.

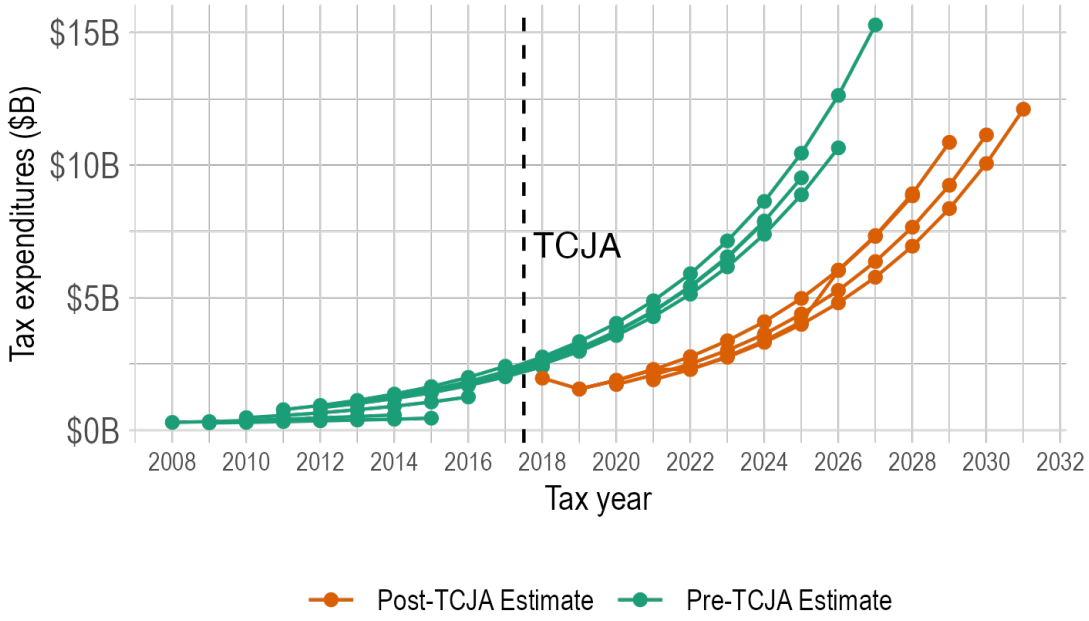


Figure 1: Annual Orphan Drug Credit Tax Expenditure Estimates (U.S. Department of the Treasury 2023)

2 Model of Biopharmaceutical R&D

Consider a firm deciding whether and how to invest in developing a therapy. The firm owns a candidate drug and may test the drug in clinical studies by employing clinical R&D inputs in the United States (z_U with price p_U) or internationally (z_I with price normalized to $p_I = 1$) to generate research output R. The net present value of testing the drug depends on the present value of expected after-tax profits from developing the drug in the United States (Π_U) and internationally (Π_I), the use of R&D inputs, and the treatment of R&D spending by the US and international tax systems $T_U(.)$ and $T_I(.)$, respectively. If $T_j(C) < C$, then R&D spending in geography j is treated preferentially so that the effective cost of R&D is less than its nominal cost. This could reflect tax deductibility of business expenses or credits for R&D spending. The firm tests the drug if the net present value of doing so

is positive: NPV $= \Pi_U + \Pi_I - T_U(p_U z_U) - T_I(z_I) \geq 0$. In other words, the after-tax value of expected lifetime profits must meet or exceed the after-tax cost of R&D.

Suppose that to enter the US market, the firm must produce a threshold level of research $R \geq \underline{R}$. If the firm does so, the drug is approved in the United States with probability q. Further suppose that R&D inputs produce research output according to a constant elasticity of substitution production function with elasticity σ: $R = A(\alpha z_U^{\frac{\sigma-1}{\sigma}} + (1-\alpha) z_I^{\frac{\sigma-1}{\sigma}})^{\frac{\sigma}{\sigma-1}}$. Because research is costly, the firm produces exactly the amount of research needed for approval, and so the problem of the firm is to choose z_U and z_I to minimize after-tax development costs subject to $R = \underline{R}$. Then the firm chooses whether to develop given this optimal choice of R&D inputs.

$$\log\left(\frac{z_U}{z_I}\right) = \sigma \log\left(\frac{\alpha}{1-\alpha}\frac{1}{p_U}\right) + \sigma \log(T_I'(z_I)) - \sigma \log(T_U'(p_U z_U)) \tag{1}$$

Equation 1 shows that relative demand for US R&D depends on the relative after-tax marginal cost of R&D. Policies that reduce the marginal cost of R&D spending in the US will tend to encourage more R&D stateside. On the other hand, preferential policies abroad reduce the relative demand for US-based R&D. Moreover, σ represents the elasticity of US R&D intensity with respect to the marginal cost of R&D, such that an X percent increase in the marginal cost of US R&D causes a σX percent decrease in the intensity of US R&D spending.

I model the cost of R&D spending in the US as depending on the tax system in two primary ways. First, R&D spending is tax-deductible and thus reduces the portion of operating income that is taxable. Second, R&D may be used to claim tax credits, which directly offset tax liability. I now show that the impact of both provisions on the effective price of US R&D depends crucially on whether firms are established US drugmakers in the sense of having current taxable income.

$$T'_U(C) = \begin{cases} (1-\tau)(1-s(C)) & \text{if established firm} \\ 1 - q\exp(-rT_m)(\tau + (1-\tau)s) & \text{otherwise.} \end{cases} \tag{2}$$

Equation 2 expresses the marginal cost of a dollar of US R&D spending for established firms and startups where τ is the firm's marginal tax rate and $s = s(C)$ is the firm's marginal credit rate. First consider the case of an established firm. If this type of firm spends an additional \$1 on US R&D, it can reduce its US taxable income by the same amount, decreasing its tax bill by $\$\tau$. The firm can also claim a credit on the dollar of R&D and use it to immediately offset tax liability, but it must reduce the credit by the tax rate so that the value of the credit is $(1-\tau)s$. Thus, the marginal cost of US R&D spending for an established firm is $T'_U(C) = 1 - \tau - (1-\tau)s = (1-\tau)(1-s)$.

On the other hand, startups with little or no taxable income in the US face a drastically different outlook. Because these firms do not have current taxable income, they cannot use deductions or credits to reduce their current tax bill, and while they can roll deductions and credits forward to years in which they do have taxable income, because market entry is dependent on FDA approval, the probability of ever earning taxable income in the US is limited by q. Given a fixed cost of capital r, for any $q < 1$ and $T_m > 0$ years until marketing approval, an established firm's marginal cost of US R&D is less than a startup's. From Equation 1, this implies that, all else equal, established firms will tend to spend a greater fraction on US R&D inputs than startups.[10]

TCJA had two key provisions that impacted τ and s. First, it instituted a flat corporate tax rate of 21 percent. Prior to this, the top corporate rate was 35 percent and applied to all income above \$10 million. Because R&D expenses are tax-deductible, this increased both after-tax expected income and the after-tax cost of US R&D. Second, the reduction of

[10]In practice, I show that firms with FDA-approved drugs sponsor trials with *fewer* US sites than other firms. This may reflect greater substitutability of foreign R&D inputs or lower relative price of using those inputs among established firms. The empirical design employed in Section 4 addresses this concern in part by controlling for time-invariant productivity differences across firms with firm fixed effects.

Table 1: Marginal Cost of US Biopharmaceutical R&D at Start of Phase I

Established firm	Rare disease drug	Post-TCJA	τ	s	T'_U	$\Delta \log(T'_U)$
No	No	No	0.35	0.10	0.971	
No	No	Yes	0.21	0.10	0.980	0.00896
No	Yes	No	0.35	0.50	0.953	
No	Yes	Yes	0.21	0.25	0.972	0.0193
Yes	No	No	0.35	0.10	0.585	
Yes	No	Yes	0.21	0.10	0.711	0.195
Yes	Yes	No	0.35	0.50	0.325	
Yes	Yes	Yes	0.21	0.25	0.592	0.601

All estimates assume an 80 month lag between clinical study initiation and approval, an 8% cost of capital, and an approval success rate of 11.83%. Prior to TCJA, which set a single corporate rate of 21%, I assume firms faced the top corporate rate of 35%. Moreover, I assume that firms developing non-rare disease drugs face a marginal credit rate of 10% from the R&D tax credit.

ODC from 50 percent to 25 percent of QCTEs increased the after-tax cost of R&D for rare disease drugs. In Table 1, I show how these provisions led to heterogeneity in the after-tax marginal cost of US R&D across firms, drugs, and policy environments at the beginning of the clinical stage. Established firms consistently have the lowest US R&D costs because they are able to immediately and with certainty use R&D spending to reduce their US tax bill. Prior to being reduced by TCJA, established firms saved \$0.26 per dollar due to ODC, a 44 percent reduction in marginal cost. In contrast, ODC decreased startups' marginal cost by less than \$0.02 per dollar of US rare disease R&D (2 percent).

Now consider investment over two time periods: $t = 0$ prior to TCJA and $t = 1$ after TCJA. Define $\lambda_t = \sigma \log\left(\frac{\alpha_t}{1-\alpha_t}\frac{1}{p_{Ut}}\right) + \sigma \log(T'_{It}(z_{It}))$. Then $\log\left(\frac{z_{Ut}}{z_{It}}\right) = \lambda_t - \sigma \log(T'_{Ut}(p_{Ut} z_{Ut}))$. Taking the first difference of both sides, I obtain $\Delta \log\left(\frac{z_{Ut}}{z_{It}}\right) = \Delta\lambda - \sigma \Delta \log(T'_U)$.[11] Therefore, the percent effect of TCJA on US research intensity is proportional to $\Delta \log(T'_U)$. In the last column of Table 1, I show that this term varied significantly across types of firms and drugs. TCJA increased $\log(T'_U)$ for firms beginning clinical development of non-rare

[11] $\Delta \log(T'_U) = \log(T'_{U1}(p_{U1} z_{U1})) - \log(T'_{U0}(p_{U0} z_{U0}))$.

and rare disease drugs by 0.00896 and 0.0193 log points, respectively; accordingly these increases correspond to 0.9 and 1.9 percent increases in the after-tax marginal cost. In contrast, TCJA increased $\log(T'_U)$ by 0.195 log points for non-rare disease drugs and 0.601 for rare disease drugs, corresponding to 22 and 82 percent increases in marginal cost, respectively. Thus, this model predicts that TCJA decreased US research intensity among all firms but also that the magnitude of the decline was substantially greater among established firms and particularly among those firms' rare disease R&D activities. Note that if ODC remained a 50 percent credit, then TCJA would have increased the marginal cost of US rare disease R&D spending by established firms by 22 percent, the same as the cost increase for non-rare disease drugs. This counterfactual calculation motivates the empirical specification outlined in Section 4 which relies on an assumption of parallel trends in a proxy for R&D intensity.

3 Data

I combined several public resources to construct a dataset of clinical trials for rare and non-rare disease drugs. I started by identifying a group of candidate trials using the public trial registry ClinicalTrials.gov. I then classified trials as rare or non-rare by determining whether trial descriptors listed in the registry corresponded to diseases listed on Orphanet, a European database which "aims to provide high-quality information on rare diseases."

Clinical Trials

To identify candidate trials, I first obtained an extract of all 483,239 studies registered on ClinicalTrials.gov as of February 20, 2024, from the Aggregate Analysis of ClinicalTrials.gov (AACT) database. The AACT database compiles information on studies' submission dates, clinical phases, eligibilities, enrollments, site locations, interventions, and

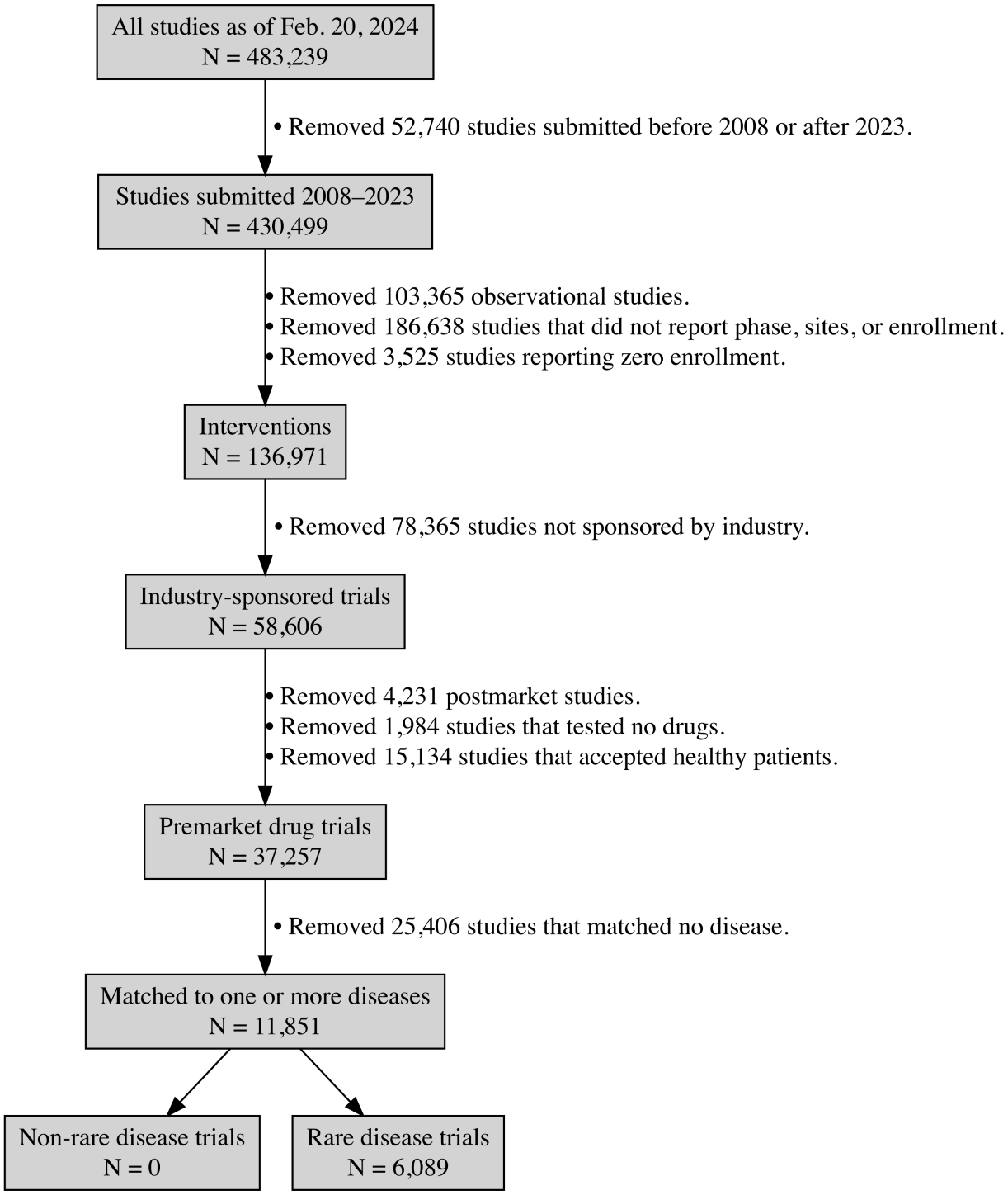


Figure 2: Sample Construction

many more characteristics. In Figure 2, I depict how I used this extract to create my sample of rare and non-rare disease trials. First, because clinical trials were severely underreported before new regulations were passed by Congress in 2007 (Phillips et al. 2017, Swanson et al. 2021), I removed studies submitted prior to 2008 or after 2023. Next, I further eliminated studies that were observational (i.e. studies that did not assign specific interventions to participants), studies that did not report clinical phase, sites, or enrollment, and studies that reported zero enrollment. I then removed studies with primary sponsors outside of industry, and finally removed postmarket studies (Phase IV trials), studies that did not test any drugs, and studies that enrolled healthy patients. This process yielded a sample of 37,256 premarket drug trials. Trial characteristics were primarily sourced from the AACT database. To determine whether trial sponsors had approved drugs in the United States, I joined data on US-approved drugs from Drugs@FDA using natural language processing to attain high merge performance.[12]

Diseases

I then aimed to separately identify trials that tested interventions in patients with rare and non-rare diseases. I did so by using Medical Subject Heading (MeSH) terms, which are elements of a controlled medical vocabulary that are reported on ClinicalTrials.gov for the primary conditions studied in trials. By determining MeSH terms that correspond to diseases reported on Orphanet, I am able to link the two databases and identify rare disease trials.

Because Orphanet determines whether diseases are rare according to the European Union definition, which differs from that of the United States, I first determined whether Orphanet diseases would be classified as rare in the United States. Among 4,184 diseases listed in Orphanet, 4,075 (97%) are considered rare diseases in the European Union. For these diseases, I assume they are also considered rare in the United States unless Orphanet

[12]I used natural language processing (NLP) techniques to perform two complementary tasks: consolidating trial sponsors with similar names and determining which sponsors have FDA-approved drugs. R code to replicate the following process is available on Github.

reports estimated US prevalence greater than 1 in 1,000.[13] This criterion classifies all but one of the European rare diseases as rare in the United States. The remaining 109 diseases on Orphanet are uncommon but not considered non-rare in Europe. These include diseases such as Crohn's disease, Hashimoto's disease, and Tourette's syndrome. I manually verified whether each of these diseases were considered rare in the United States by first surveying the epidemiological literature to determine whether US prevalence is consistent with orphan designation eligibility and then further determining whether any drug had ever received an orphan designation for the disease. I classified 10 diseases with any evidence of orphan designation eligibility as rare in the United States, classified another 88 diseases as non-rare, and failed to classify the remaining 10 diseases.

For each Orphanet disease, I obtained MeSH terms using disease names and identifiers reported by Orphanet.[14] In total, 16,616 terms linked to 2,384 Orphanet diseases (57% of all diseases in the database). I then linked diseases to trials with an exact merge on MeSH terms and classified trials as rare disease trials if they matched any rare disease. Of the 37,256 premarket drug trials, I matched 11,850 to Orphanet diseases, including 6,089 to one or more rare diseases and 5,761 to only non-rare diseases.[15] In the following analyses, I consider two samples of trials: a restricted sample that includes only the 11,850 that matched one or more (rare or non-rare) Orphanet diseases, and an unrestricted sample that includes all 37,256 premarket drug trials, including 25,406 that matched no Orphanet disease.

[13] US prevalence of 1 in 1,000 corresponds to roughly 330,000 Americans as of 2023, significantly above the threshold of 200,000 persons for a disease to be considered rare. However, treating diseases within the prevalence range of 6-9 Americans per 10,000 as non-rare reclassifies only 2 diseases (sarcoidosis and American trypanosomiasis) both of which have prevalence on the cusp of 200,000 and both of which have received orphan drug designations in the past 5 years.

[14] To obtain MeSH descriptors for Orphanet diseases, I queried the UMLS Metathesaurus for UMLS concept codes for each disease. Then using those concept codes and any UMLS concept codes ascribed to diseases by Orphanet, I queried the Metathesaurus for related concepts that are narrower in definition. For example, the concept code C0029434 for Gynandroblastoma is related to the narrower concept code C0346178 for Ovarian gynandroblastoma. Lastly, I queried each concept for MeSH terms. For example, the Orphanet disease oxoglutaric aciduria maps to 5 terms including: "Oxoglutaricaciduria," "Oxoglutaric Aciduria," and "2-Ketoglutarate Dehydrogenase Deficiency."

[15] See Section 5.1 for examples of trials by group.

Table 2: Trial Characteristics

Characteristic	**Rare disease trial** N = 6089	**Non-rare disease trial** N = 5762	**Other trial** N = 25406
Phase			
1	2,253 (37%)	1,707 (30%)	7,964 (31%)
2	2,271 (37%)	2,106 (37%)	9,186 (36%)
3	1,565 (26%)	1,949 (34%)	8,256 (32%)
Controlled	2,217 (36%)	3,178 (55%)	13,557 (53%)
Randomized control	2,167 (36%)	3,130 (54%)	13,344 (53%)
Active control	793 (13%)	1,307 (23%)	6,196 (24%)
Open label	4,113 (68%)	2,691 (47%)	12,671 (50%)
Trial sites	26.2 (40.8)	43.2 (74.2)	29.3 (65.7)
Share of sites in US	0.43 (0.41)	0.38 (0.40)	0.42 (0.44)
Enrollment	119.3 (187.2)	271.3 (740.1)	254.0 (748.6)
Enrolls children	1,522 (25%)	316 (5.5%)	2,941 (12%)
Terminated	902 (15%)	799 (14%)	2,751 (11%)
Established sponsor	2,738 (45%)	3,309 (57%)	12,369 (49%)

[1] n (%); Mean (SD)

Table 2 presents summary characteristics by group for the unrestricted sample. Rare disease trials are notable in several respects: only 36% of rare disease trials are designed to enroll a control group, substantially fewer than the 55% of non-rare disease trials that do so. Similarly, only 13% of rare disease trials test a treatment against an active comparator (a treatment used in clinical practice), compared to 23% of non-rare trials. Rare disease trials are also more likely to be open label (68% vs. 47%),[16] more likely to enroll children, (25% vs. 5.5%), and less likely to be sponsored by a firm with FDA-approved drugs (45% vs. 57%).[17] Rare disease trials also include fewer sites and enroll fewer patients at those

[16]In open label studies, health providers and patients are both aware of treatment assignment. In contrast, "masked" or "blinded" studies withhold information from investigators, participants, and/or other involved parties. Open label studies are often more practical, but may introduce bias via differential assessment of outcomes and differential attrition (Schulz & Grimes 2002).

[17]In Table 6, I show that these differences are not explained by differences across groups in the distribution

sites than trials for non-rare diseases.

4 Empirical Model and Results

The model in Section 2 highlights how TCJA had heterogeneous impacts on the marginal cost of US R&D. The model thus predicts that TCJA reduced demand for US pharmaceutical R&D inputs and that the largest (percent) declines in US R&D should occur among established drugmakers, particularly in their development of rare disease drugs. The model also predicts that TCJA should lead established firms to increase development of rare disease drugs less than they increase development of other drugs. In this section, I outline the empirical counterparts of these predictions and present regression estimates.

4.1 US Research Intensity

Because determining the amount of R&D spending on clinical trials by geography for both publicly- and privately-held corporations requires data that is very challenging to obtain, I proxy US clinical R&D intensity using the share of clinical trial sites located in the United States. The primary downside of this proxy is that it assumes that R&D inputs are used proportionately across sites. For example, it assumes that the same number of patients are recruited across sites and that the same amount of the R&D input per patient is used at each site. The latter assumption is characteristic of a well-conducted clinical trial that creates comparable conditions for patients regardless of site. The prior assumption is less tenable and thus best viewed as an approximation.

To estimate the effect of the reduced ODC, I use a triple difference model that leverages variation in the timing of trials, the diseases they study, and the types of firms sponsoring those trials. This is motivated in large part by the heterogeneous impacts of TCJA on the marginal cost of US R&D along these dimensions, as demonstrated in Table 1. While a

of clinical phase, as the same patterns emerge within phase.

standard difference-in-differences model that uses only variation across trials in timing and diseases of study could be used, incorporating variation across sponsors serves two useful purposes. First, it leverages additional variation in the effect of TCJA to improve the precision of my estimates. Second, it facilitates controlling for heterogeneous time trends across types of trials. This latter point is particularly important because the benefits from recent innovations in trial design and the use of precision medicine biomarkers may disproportionately reduce the costs of rare disease trials and lead to differing trends in drug development for reasons unrelated to the tax system.

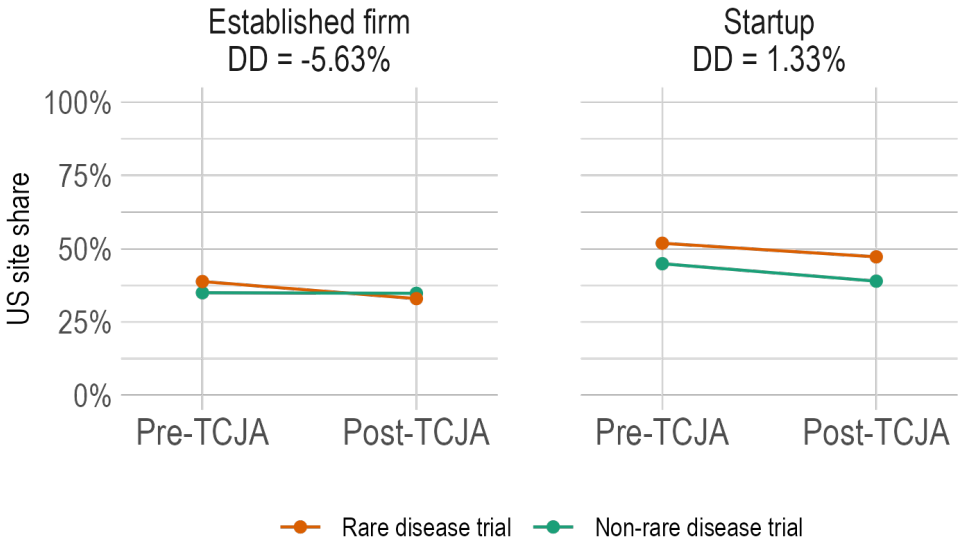
Consider trial i that is sponsored by firm f and is submitted to ClinicalTrials.gov in year t. The trial either studies a rare disease, $r = 1$, or does not, $r = 0$. The trial enrolls patients across one or more sites. Let US site share$_{iftr}$ represent the fraction of sites that are located in the United States.

$$\begin{aligned} \text{US sites}_{iftr} = \; & \beta_1 \text{TCJA}_t + \beta_2 \text{TCJA}_t \times \text{RD trial}_r \\ & + \beta_3 \text{Established}_{ft} \times \text{TCJA}_t + \beta_4 \text{Established}_{ft} \times \text{RD trial}_r \times \text{TCJA}_t \\ & + \gamma_1 \text{Established}_{ft} + \gamma_2 \text{Established}_{ft} \times \text{RD trial}_r \\ & + X_i' \Pi + \alpha_{fr} + u_{iftr}. \end{aligned} \tag{3}$$

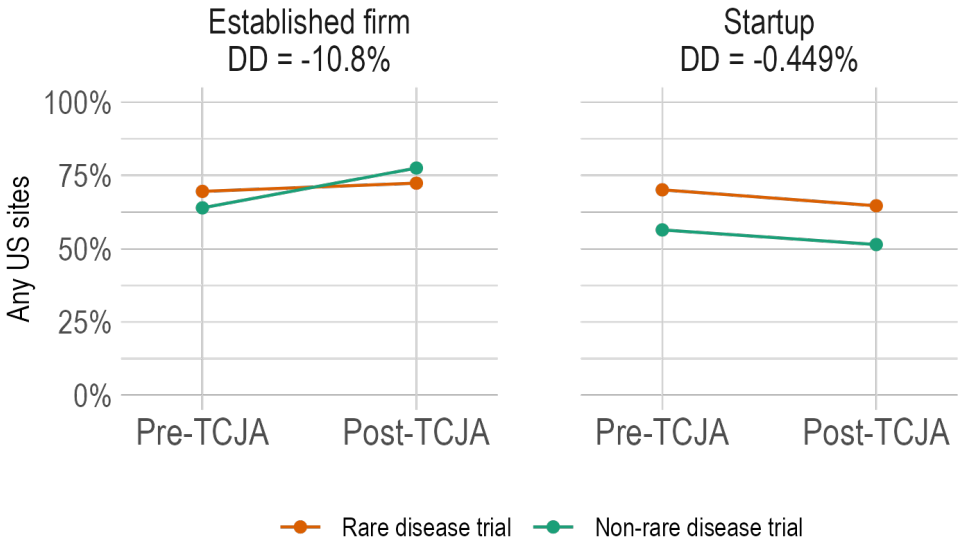
Equation 3 models US site share as a function of Established$_{ft}$, a dummy variable indicating whether sponsor f had an approved drug prior to year t, RD trial$_i$, a dummy variable indicating whether the trial studied any rare diseases, and TCJA$_t$, a dummy variable indicating whether the trial was submitted in 2018 or later. The model also includes trial-level covariates X_i, sponsor-by-rare-disease fixed effects, and year-by-rare-disease fixed effects.[18]

Identification of the causal effect rests on a counterfactual assumption, the additional post-TCJA change in the US site share of rare disease trials relative to non-rare trials among

[18]Covariates include dummies for Phase 2, Phase 3, combination phase trial, Phase 2 and combination phase (i.e. a Phase 2/3 trial), open label design, use of a control/comparator group, and enrollment of children. I also control for the natural logarithm of the number of trial sites and the natural logarithm of enrollment.



(a) US Site Share



(b) Any US Sites

Figure 3: US Research Intensity, by Firm Type

Table 3: Trial-level Linear Regressions

Dependent Variables:	Share of sites in US		Any US sites	
Estimate:	DD, Established	DDD	DD, Established	DDD
Model:	(1)	(2)	(3)	(4)
Variables				
TCJA	0.014	-0.008	0.098***	-0.011
	(0.017)	(0.023)	(0.020)	(0.024)
RD trial × TCJA	-0.071***	-0.008	-0.079***	0.033
	(0.019)	(0.030)	(0.022)	(0.029)
Established × TCJA		0.028		0.119***
		(0.029)		(0.032)
Established × RD trial × TCJA		-0.062		-0.115**
		(0.035)		(0.036)
Controls	Yes	Yes	Yes	Yes
Fixed-effects				
Sponsor-RD trial	Yes	Yes	Yes	Yes
Fit statistics				
Observations	6,047	11,851	6,047	11,851
Dependent variable mean	0.355	0.408	0.694	0.656
R^2	0.288	0.566	0.364	0.567

Clustered (Sponsor) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, †: 0.1*

All models include terms for Established$_{ft}$ and Established$_{ft}$ × RD trial$_r$.

established firms would have been the same as the additional change among startups if not for the ODC reduction. Figure 3 depicts this graphically. For both groups of firms, I calculate the additional change in share for rare disease trials. I attribute the difference in the additional change (-5.6% - 1.3% = -6.9%) entirely to the reduced ODC, which substantially increased the cost of US rare disease R&D for established firms but not the cost of rare disease R&D for startups nor non-rare disease R&D for any firm.

Column 1 of Table 3 reports difference-in-differences estimates using only the subset of established firms. Established firms reduced US site share of rare disease trials by 7.1 percent relative to the a slight increase in site share of non-rare disease trials. Under a

typical parallel trends assumption that US site share of rare disease trials sponsored by established firms would have followed the same trend as non-rare disease trials sponsored by those firms if the ODC rate remained at 50 percent, this captures the effect of the ODC reduction on US site share. However, because other time-varying factors may impact the relative productivity of US research inputs for rare disease research, I estimate a triple differences model that accounts for certain sources of this bias.

Column 2 reports estimates of the full model outlined in Equation 3. Startups did not significantly change their use of US trial sites after TCJA. On average, these firms decreased the share of non-rare disease trial sites located in the US by a statistically insignificant 0.8 percentage points (95% CI: [-5.3 p.p., 3.7 p.p]). Similarly, they decreased the share of rare disease trial sites in the US by only 1.7 percentage points (95% CI: [-5.5 p.p., 2.2 p.p]). Established firms also did not significantly change their use of US sites in non-rare trials (+1.9 p.p, 95% CI: [-1.4 p.p, 5.3 p.p]); however, these firms decreased their use of these sites in rare disease trials by a statistically significant 5.1 percentage points ($p < 0.001$, 95% CI: [-8.0 p.p., -2.2 p.p.]). The triple difference coefficient of -0.062 reflects that the additional reduction in US site share of rare disease trials sponsored by established firms relative to the change in site share of non-rare disease trials sponsored by the same firms was 6.2 percentage points greater than the relative change among startups. This coefficient is only marginally significant ($p<0.1$), but its magnitude is only slightly less than the precise 7.1 percentage point reduction implied by the difference-in-differences estimates reported in Column 1, suggesting that there is minimal bias in that initial coefficient. The average rare disease trial sponsored by an established firm post-TCJA has 33 percent of its trials located in the US. Therefore, the triple difference model implies that without the ODC reduction, 39.2 percent of these trials' sites would have been located in the US. The 6.2 point reduction corresponds to an economically significant 16 percent reduction in US sites among these trials.

Figure 4 reports coefficient estimates from a triple differences event study model that

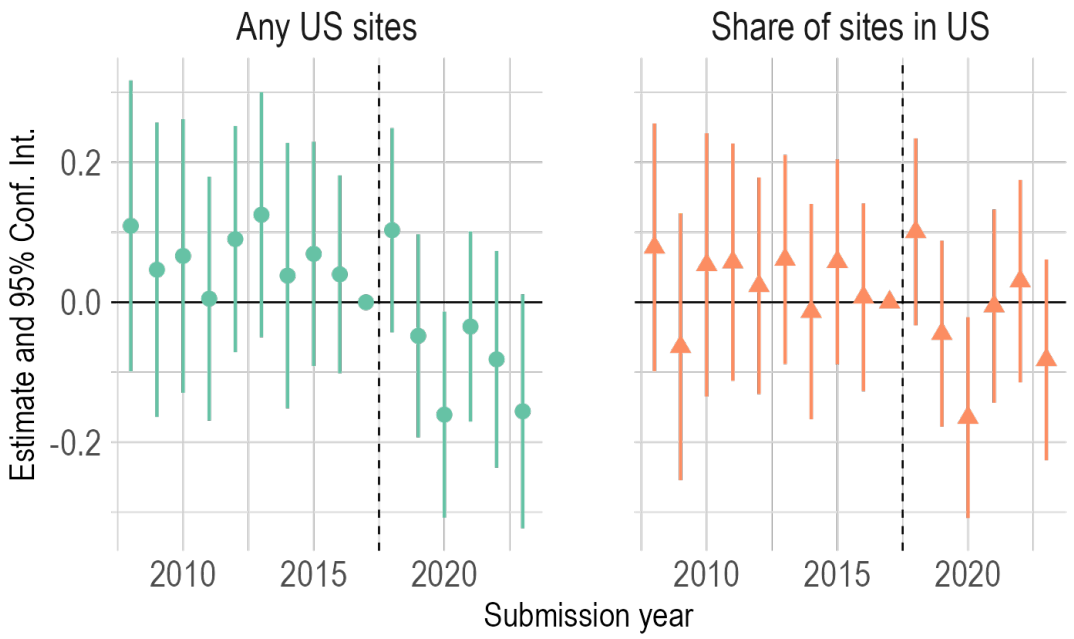


Figure 4: Triple Differences Event Studies, US Research Intensity

allows the effect of TCJA to be fully dynamic.[19] Coefficients that are substantially different from zero prior to enactment of TCJA would indicate a difference in relative (rare vs. non-rare) trends in US site share between established firms and start-ups. If those relative trends differed prior to TCJA, then assuming established firms would have followed the same relative trends as startups after TCJA would be difficult to justify. However, the coefficients are indicative of no such pattern, suggesting that this assumption is reasonable.

Relative declines in rare disease trial US site share among established firms are consistent with a large intensive margin response to TCJA's reduction of the Orphan Drug Credit. Columns 3 and 4 additionally demonstrate an extensive margin response to the

[19]Specifically, the figure reports ordinary least squares estimates of β_s from the following specification:

$$\begin{aligned}\text{US sites}_{iftr} = \; & + \gamma_1 \text{Established}_{ft} + \gamma_2 \text{Established}_{ft} \times \text{RD trial}_r + X_i'\Pi + \alpha_{fr} + \lambda_{tr} \\ & + \sum_{s \neq 2017} \gamma_s \text{Established}_{ft} \times \mathbf{1}\{s = t\} + \beta_s \text{Established}_{ft} \times \text{RD trial}_r \times \mathbf{1}\{s = t\} + u_{iftr}.\end{aligned}$$

reduction. Reporting difference-in-differences estimates, Column 4 indicates that established firms increased the proportion of non-rare disease trials with any US sites by 9.8 percentage points while they increased rare disease trials by 7.9 percentage points less. Furthermore, Column 5 shows that this relative decline was unique to established firms. Startups decreased the proportion of non-rare disease trials with any US sites by 1.1 percentage points and increased the proportion of rare disease trials with any US sites by 2.2 percent. In contrast, established firms increased the proportion of non-rare disease trials with any US sites by 10.7 percentage points and rare disease trials by only 2.5 percentage points. While trials overall tended towards greater inclusion of US sites, the small increase in the proportion of US sites among rare disease trials sponsored by established firms relative to the same increase among non-rare trials by established firms suggests that the reduced ODC substantially decreased rare disease R&D in the US. The triple differences model implies a reduction in the proportion of rare disease trials sponsored by established firms with any US site by 11.5 percent, a 14 percent reduction from the counterfactual proportion implied by the triple differences model.[20]

4.2 Clinical Trials

The previous section demonstrated a significant effect of TCJA's reduction of the ODC rate on the use of US R&D inputs. Yet ultimately the goal of rare disease policy is to encourage firms to test new treatments to bring to market, regardless of where they are tested. Although reducing ODC may cause firms to substitute US R&D inputs for foreign inputs, this need not lead to substantial reductions in the expected profit of initiating clinical studies.

Figure 5 shows that trends in submissions of rare and non-rare disease trials have di-

[20]The 12 percent increase assumes that established firms would have counterfactually increased the share of rare disease trials with any US site by the same amount as they increased the same share for non-rare disease trials. After TCJA, 72.4 percent of established-firm-sponsored rare disease trials had any US site. Therefore, under the aforementioned assumption, 82.3 percent of these trials would have at least one US trial site. A 9.9 p.p. reduction off a base of 82.3 percent is a $100\% \times 9.9/82.3 = 12\%$ reduction.

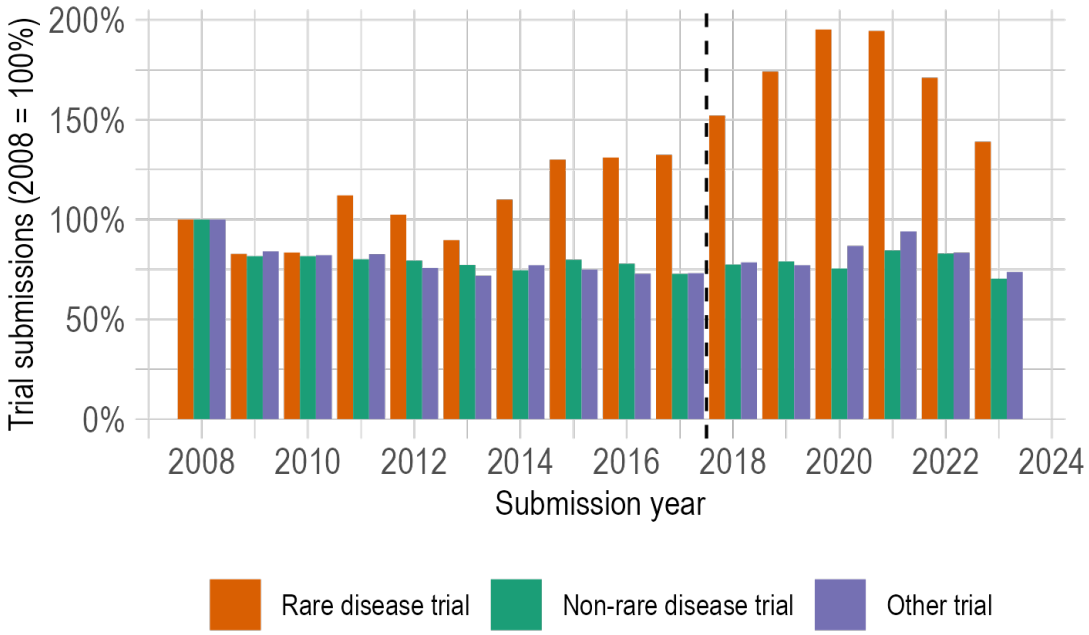


Figure 5: Trial Submission Growth

verged since 2008. By 2017, established firms had slightly increased the number of rare disease trials they sponsored from 155 to 166 (a 0.8 percent annual growth rate), while the number of non-rare disease trials they submitted fell from 307 to 173 (a 6.2 percent annual decline). Startups increased trials for both rare and non-rare diseases, but growth was most dramatic among rare diseases, which increased from 135 to 218 in 2008, (5.5 percent growth) compared to an increase from 145 to only 156 non-rare trials (0.8 percent growth). Among both groups of firms, the growth rate of rare disease trials was significantly higher than non-rare disease trials, suggesting that an identification based on the assumption of

parallel trends between the two types of trials is suspect at best.

$$\begin{aligned}\text{Trial submissions}_{ftr} = \quad & \exp\Big(\beta_1 \text{TCJA}_t + \beta_2 \text{TCJA}_t \times \text{RD trial}_r \\ & + \beta_3 \text{Established}_{ft} \times \text{TCJA}_t + \beta_4 \text{Established}_{ft} \times \text{RD trial}_r \times \text{TCJA}_t \\ & + \gamma_1 \text{Established}_{ft} + \gamma_2 \text{Established}_{ft} \times \text{RD trial}_r \\ & + \alpha_{fr}\Big) + u_{ftr}. \end{aligned} \tag{4}$$

Therefore, to test how firms differentially changed their clinical research output, I estimate the fixed effects Poisson regression in Equation 4. $\text{Trial Submissions}_{ftr}$ is equal to the total number of trials of type r sponsored by f in year t. I estimate this model using an unbalanced panel of firms. Firms are included in the panel in any year in which they are "active," which I define as having started a trial in that year or in any of the previous three years.

Like Equation 3, this is a triple differences model; however, because of the Poisson specification, the interpretations of coefficients reflect percent-scale differences in trial submissions. Like in the other equation, estimating this model identifies a causal effect under an assumption regarding relative changes in the outcome for established firms and startups. Specifically, it assumes that the additional percent change in rare disease trials relative to non-rare disease trials sponsored by established firms would have equaled the additional percent change for rare disease trials sponsored by startups if it were not for the increased ODC. While this assumption is unverifiable, the pre-TCJA growth rate of rare disease trial submissions net of growth in non-rare trials is similar across both established and startup firms. This suggests that similar relative growth trajectories were likely to continue even if TCJA had not been enacted.

Table 4 shows that the reduced ODC rate did little to slow down rare disease trial submissions. The relative increase in rare disease trial submissions after TCJA was 0.217

Table 4: Firm-level Poisson Regressions

Dependent Variables:	All trials	Phase I	Phase II	Phase III
Model:	(1)	(2)	(3)	(4)
Variables				
TCJA	-0.386***	-0.483***	-0.441***	-0.178**
	(0.040)	(0.059)	(0.048)	(0.063)
RD trial × TCJA	0.275***	0.241*	0.320***	0.314*
	(0.063)	(0.095)	(0.087)	(0.140)
Established × TCJA	-0.118	0.081	-0.155	-0.319**
	(0.087)	(0.107)	(0.089)	(0.113)
Established × RD trial × TCJA	0.217*	0.074	0.123	0.398*
	(0.107)	(0.147)	(0.145)	(0.183)
Fixed-effects				
Sponsor-RD trial	Yes	Yes	Yes	Yes
Fit statistics				
Sample size	75,534	75,534	75,534	75,534
Sponsors	4,685	4,685	4,685	4,685
Effective sample size	49,858	29,792	33,395	21,431

Clustered (Sponsor) standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, †: 0.1*

All models include terms for Established$_{ft}$ and Established$_{ft}$ × RD trial$_r$.

log points or 24.2 percent *greater* for established firms than startups (p > 0.1, 95% CI: [-0.05, 0.35]). This is despite the fact that TCJA significantly increased the cost of rare disease R&D in the US for established firms. Columns (2) through (4) show that this relative increase is driven almost entirely by Phase III trials. The relative change in Phase I and II rare disease trials sponsored by established firms was slightly larger in percent terms than and not statistically different from the same relative change in startup-sponsored trials.

5 Discussion

In this paper, I evaluated the incidence and impacts of the Orphan Drug Credit. I show that the credit's benefits accrue primarily to established firms rather than startups because the

latter firms cannot reliably reduce tax liability with the credits they earn from researching rare disease drugs. The 2017 reduction of ODC reduced established firms' use of clinical trial sites for rare diseases in the United States, but had no measurable effect on the number of trials that these firms sponsored.

There are a number of limitations to this study. First, because I do not observe firms' US taxable income, I proxy their exposure to ODC by classifying them as established firms or startups on the basis of having FDA-approved drugs. This measure is imperfect because some firms may not earn substantial taxable income from their approved drugs and others may have taxable income in spite of a lack of previous approvals. Researchers have noted that some large pharmaceutical firms have US taxes far below what would be expected given their US revenues, and have linked this discrepancy to tax avoidance (Setser 2023). If established firms regularly face negative US federal tax liability (for example, because of overseas profit shifting), this would tend to reduce firms' exposure to the ODC reduction and attenuate my estimates. A related concern is that, while the model of pharmaceutical development highlights key TCJA reforms, it simplifies a drastic overhaul of the tax system, which (among other changes) changed the US from a primarily worldwide system of taxation to a primarily territorial system and removed the corporate alternative minimum tax. If these changes changed the relative cost of US rare disease R&D for established firms differently than the relative cost for startups, the interpretation of the estimates in Section 4 cannot be interpreted as the effect of the ODC reduction in isolation. Further analysis with financial or tax data may be able to shed new insights on the incidence and impacts of ODC.

This study also did not consider societal preferences over the location of clinical trial sites. Many rare disease patients likely prefer clinical trials to be located domestically because participating in trials facilitates access to experimental therapies that can be otherwise difficult or impossible to obtain. Traveling abroad may be impractical from a health standpoint for many patients even if travel costs are reimbursed by trial sponsors. To my

knowledge, no study has previously considered this motivation for rare disease tax credits Politicians and health care professionals may also benefit from onshoring high-paying health care jobs. Thus, while R&D credits may have minimal effect on the overall level of drug development, policy makers and their constituents have good reason to fight for domestic pharmaceutical R&D.

References

Acemoglu, D. & Linn, J. (2004), 'Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry', *The Quarterly Journal of Economics* **119**(3), 1049–1090.

Bagley, N., Berger, B., Chandra, A., Garthwaite, C. & Stern, A. D. (2019), 'The Orphan Drug Act at 35: Observations and an outlook for the twenty-first century', *Innovation Policy and the Economy* **19**(1), 97–137.

Bloom, N., Van Reenen, J. & Williams, H. (2019), 'A Toolkit of Policies to Promote Innovation', *Journal of Economic Perspectives* **33**(3), 163–184.

Center for Drug Evaluation and Research (2023), 'Prescription Drug User Fee Amendments', https://www.fda.gov/industry/fda-user-fee-programs/prescription-drug-user-fee-amendments.

Cho, N. H., Shaw, J. E., Karuranga, S., Huang, Y., da Rocha Fernandes, J. D., Ohlrogge, A. W. & Malanda, B. (2018), 'IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045', *Diabetes Research and Clinical Practice* **138**, 271–281.

Clinical Trials Transformation Initiative (n.d.), 'Aggregate Analysis of ClinicalTrials.gov (AACT) Database', https://aact.ctti-clinicaltrials.org/.

Crisafulli, S., Sultana, J., Fontana, A., Salvo, F., Messina, S. & Trifirò, G. (2020), 'Global

epidemiology of Duchenne muscular dystrophy: An updated systematic review and meta-analysis', *Orphanet Journal of Rare Diseases* **15**(1), 141.

DiMasi, J. A., Grabowski, H. G. & Hansen, R. W. (2016), 'Innovation in the pharmaceutical industry: New estimates of R&D costs', *Journal of health economics* **47**, 20–33.

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

Fidler, B. (2023), 'Sarepta prices Duchenne gene therapy at $3.2M', https://www.biopharmadive.com/news/sarepta-duchenne-elevidys-price-million-gene-therapy/653720/.

Gamba, S., Magazzini, L. & Pertile, P. (2021), 'R&D and market size: Who benefits from orphan drug legislation?', *Journal of Health Economics* **80**, 102522.

Harris, E. (2019), 'Breaking Down Pricing Of Cell & Gene Therapies', https://www.cellandgene.com/doc/breaking-down-pricing-of-cell-gene-therapies-0001.

Internal Revenue Service (2018), 'Form 8820'.

Metathesaurus (2021), *in* 'UMLS® Reference Manual [Internet]', National Library of Medicine (US).

National Organization for Rare Disorders (2022), 'NORD Applauds Congressional Efforts to Restore Intent of the Orphan Drug Act', https://rarediseases.org/nord-applauds-congressional-efforts-to-restore-intent-of-the-orphan-drug-act/.

Orphanet (2024), 'Orphadata', https://www.orphadata.com/.

Phillips, A. T., Desai, N. R., Krumholz, H. M., Zou, C. X., Miller, J. E. & Ross, J. S. (2017), 'Association of the FDA Amendment Act with trial registration, publication, and outcome reporting', *Trials* **18**(1), 333.

Schulz, K. F. & Grimes, D. A. (2002), 'Blinding in randomised trials: Hiding who got what', *The Lancet* **359**(9307), 696–700.

Seoane-Vazquez, E., Rodriguez-Monguio, R., Szeinbach, S. L. & Visaria, J. (2008), 'Incentives for orphan drug research and development in the United States', *Orphanet Journal of Rare Diseases* **3**(1), 33.

Setser, B. W. (2023), Cross-border Rx: Pharmaceutical Manufacturers and U.S. International Tax Policy, Technical report, Council on Foreign Relations.

Swanson, M. J., Johnston, J. L. & Ross, J. S. (2021), 'Registration, publication, and outcome reporting among pivotal clinical trials that supported FDA approval of high-risk cardiovascular devices before and after FDAAA', *Trials* **22**(1), 817.

U.S. Department of the Treasury (2023), 'Tax Expenditures', https://home.treasury.gov/policy-issues/tax-policy/tax-expenditures.

Waters, H. & Graf, M. (2018), The costs of chronic disease in the U.S., Technical report, The Milken Institute.

Yang, G., Cintina, I., Pariser, A., Oehrlein, E., Sullivan, J. & Kennedy, A. (2022), 'The national economic burden of rare disease in the United States in 2019', *Orphanet Journal of Rare Diseases* **17**(1), 163.

Yin, W. (2008), 'Market Incentives and Pharmaceutical Innovation', *Journal of Health Economics* **27**(4), 1060–1077.

5.1 Tables

Table 5: Random Sample of Trials

Title	**Trial group**	**Phase**	**Disease(s)**	**Lead sponsor**	**Established sponsor**
BCMA-CD19 cCAR in Multiple Myeloma and Plasmacytoid Lymphoma	Rare	Phase 1	Multiple myeloma	iCell Gene Therapeutics	No
A Study Evaluating Temferon in Patients With Glioblastoma & Unmethylated MGMT	Rare	Phase 1/Phase 2	Glioblastoma	Genenta Science	No

(Continued)

Table 5: Random Sample of Trials *(continued)*

Title	**Trial group**	**Phase**	**Disease(s)**	**Lead sponsor**	**Established sponsor**
Open Label Extension Study Evaluating Safety and Biological Activity of C2L-OCT-01 PR in Acromegalic Patients	Rare	Phase 3	Acromegaly	Ambrilia Biopharma, Inc.	No
A First-in-Human, Phase 1 Study of JAB-3312 in Adult Patients With Advanced Solid Tumors	Rare	Phase 1	Squamous cell carcinoma of the esophagus	Jacobio Pharmaceuticals Co., Ltd.	No

(Continued)

Table 5: Random Sample of Trials *(continued)*

Title	**Trial group**	**Phase**	**Disease(s)**	**Lead sponsor**	**Established sponsor**
Open Label Clinical Trial of Anascorp® in Pediatric Patients With Scorpion Sting Envenomation	Rare	Phase 3	Scorpion envenomation	Instituto Bioclon S.A. de C.V.	No
Suprachoroidal Injection of CLS-TA in Patients With Non-infectious Uveitis	Rare	Phase 3	Intermediate uveitis	Clearside Biomedical, Inc.	No
131I-TLX-101 for Treatment of Newly Diagnosed Glioblastoma (IPAX-2)	Rare	Phase 1	Glioblastoma	Telix International Pty Ltd	Yes

(Continued)

Table 5: Random Sample of Trials *(continued)*

Title	**Trial group**	**Phase**	**Disease(s)**	**Lead sponsor**	**Established sponsor**
A Safety and Efficacy Study of Bevacizumab, Paclitaxel, Carboplatin Compared to Avastin® in Non-Small Cell Lung Cancer	Non-rare	Phase 3	Non-small cell lung cancer	Mabscale, LLC	No
CMP-001 in Combination With Nivolumab in Subjects With Advanced Melanoma	Non-rare	Phase 2	Melanoma	Regeneron Pharmaceuticals	Yes
Long-Term Safety and Efficacy Study of Peginterferon Beta-1a	Non-rare	Phase 3	Multiple sclerosis	Biogen	Yes

Table 6: Trial Characteristics by Phase

Characteristic	**Phase I**, N = 3960		**Phase II**, N = 4377		**Phase III**, N = 3514	
	Rare disease trial N = 2253	**Non-rare disease trial** N = 1707	**Rare disease trial** N = 2271	**Non-rare disease trial** N = 2106	**Rare disease trial** N = 1565	**Non-rare disease trial** N = 1949
Combined with next phase	926 (41%)	485 (28%)	191 (8.4%)	122 (5.8%)	0 (0%)	0 (0%)
Controlled	291 (13%)	518 (30%)	1,002 (44%)	1,266 (60%)	924 (59%)	1,394 (72%)
Randomized control	272 (12%)	499 (29%)	987 (43%)	1,249 (59%)	908 (58%)	1,382 (71%)
Active control	111 (4.9%)	235 (14%)	298 (13%)	380 (18%)	384 (25%)	692 (36%)
Open label	1,977 (88%)	1,175 (69%)	1,312 (58%)	817 (39%)	824 (53%)	699 (36%)
Trial sites	9.9 (12.0)	9.4 (13.8)	22.4 (29.5)	31.7 (42.4)	55.1 (61.4)	85.2 (106.0)
Share of sites in US	0.52 (0.43)	0.47 (0.45)	0.42 (0.41)	0.40 (0.41)	0.32 (0.33)	0.28 (0.33)
Enrollment	67.6 (101.5)	74.0 (102.7)	89.6 (119.7)	161.3 (191.5)	236.8 (285.9)	562.9 (1,199.5)
Enrolls children	357 (16%)	49 (2.9%)	519 (23%)	73 (3.5%)	646 (41%)	194 (10.0%)
Terminated	319 (14%)	214 (13%)	376 (17%)	346 (16%)	207 (13%)	239 (12%)
Established sponsor	832 (37%)	824 (48%)	1,006 (44%)	1,083 (51%)	900 (58%)	1,402 (72%)

[1] n (%); Mean (SD)