Patent Policy and Costly Imitation: Evidence from Paragraph IV Patent Challenges in the Pharmaceutical Industry*

Olena Izhak
Stora Enso

Tanja Saxell
VATT and Helsinki GSE

Tuomas Takalo
Bank of Finland and Helsinki GSE

October 11, 2023

Abstract

We estimate the effects of effective patent length and scope on patent challenges by generic firms using two quasi-experimental approaches: one based on changes in patent laws and another on the allocation of patent applications to examiners. We find that one year increase in effective patent length increases the probability of a successful patent challenge by five percentage points, whereas broader protection reduces successful patent challenging. We also match our patent length and scope estimates with a standard theory of patent policy with costly imitation. Overall, our results suggest that making pharmaceutical patents shorter but broader would increase welfare.

Keywords: Patent policy, pharmaceuticals, generic entry, innovation, imitation
JEL: I18, K20, L13, O34, O31

*We thank Vanessa Behrens, Vincenzo Denicolò, Alberto Galasso, Ashvin Gandhi, Matthew Grennan, Dietmar Harhoff, Jussi Heikkilä, Ari Hyytinen, Xavier Jaravel, William Kerr, Igor Letina, Matthew Mitchell, Petra Moser, Mattia Nardotto, Benjamin Roin, Mark Schankerman, Carlos Serrano, Morten Sæthre, Ashley Swanson, Emanuele Tarantino, Otto Toivanen, Robert Town, Janne Tukiainen, Hannes Ullrich, Rosemarie Ziedonis, and audiences of numerous seminars and conferences for useful comments and discussions. We gratefully appreciate the hospitality of the Economics Departments at Boston and Stanford Universities, and Health Care Management Department of Wharton School where parts of this research has been conducted. We also gratefully acknowledge funding from the Yrjö Jahnsson Foundation. Contacts: olena.izhak@gmail.com, tanja.saxell@vatt.fi, tuomas.takalo@bof.fi.

1 Introduction

Patent policy aims at stimulating innovation by granting temporary exclusive rights to innovators at the cost of reduced competition. Balancing this classic patent policy tradeoff between competition and innovation incentives is at the core of the Drug Price Competition and Patent Term Restoration Act of 1984 (aka the Hatch-Waxman Act) which, one the one hand, lengthened effective terms of new drug patents but, on the other hand, introduced a mechanism for generic firms to challenge those patents before their expiration. In this paper we estimate the effects of effective patent length and scope on such patent challenging using two quasi-experimental approaches. We find that one year increase in effective patent length increases the probability of a successful patent challenge by roughly five percentage points. We also find a negative effect of broader patent scope on successful patent challenging. We then match our estimates of these effects with a model of patent policy with costly imitation. Collectively, our results suggest that a policy with shorter and broader patents for pharmaceuticals would increase welfare.

The U.S. pharmaceutical industry provides a well-defined setting to assess the effects of patent policy on the imitation of patented innovations: The Hatch-Waxman Act introduced generic drug applications with Paragraph IV (PIV) certifications. In such an application a generic firm certifies noninfringement or invalidity of a new drug patent, allowing the U.S. Food and Drug Administration (FDA) to authorize generic entry before the patent expires. We use this setting to measure the probability of a successful *PIV patent challenge* by a generic entrant. To estimate the impact of effective patent length on this PIV patent challenge probability, we exploit two patent law reforms inducing quasi-experimental variation across patents in their effective terms depending on the prosecution time at the U.S. Patent and Trademark Office (USPTO): The Agreement on Trade-Related Aspects of Intellectual Property (TRIPS) of 1994 changed the statutory patent term from 17 years from the grant date to 20 years from the first filing date. Subsequently the American Inventors Protection Act (AIPA) of 1999 significantly expanded the patent term adjustments (PTAs), which compensate for delays in patent prosecution. We show how the two reforms affected the effective terms of patents with grant lags exceeding three years, whereas the effective terms of patents with shorter grant lags were hardly changed.

Using difference-in-differences (DiD) regressions, we find that TRIPS decreased the effective

terms of patents prosecuted at least three years by 17 percent, compared to other patents. This shorter effective patent length reduced the probability of PIV patent challenge around 8–10 percentage points. In contrast, AIPA increased the effective terms of patents with the long prosecution lags by 10 percent, which in turn increased the probability of PIV patent challenge by seven percentage points. These results, together with supporting evidence from ordinary least squares (OLS) regressions, suggest a positive effect of longer patent duration on PIV patent challenges.

Using various measures of patent scope and OLS regressions, we also find evidence of broader patent scope reducing the probability of PIV patent challenge. To address the endogeneity of patent scope, we develop instrumental variables (IV) for multiple claim-based scope measures by exploiting differences in the propensity of patent examiners to grant broader or more claims. The evidence from our IV regressions – though inconclusive given the lack of fully randomized examiner assignment (see Righi and Simcoe, 2019) – implies that broader patent scope wards off patent challenges: For example, a 10 percent increase in the number of related chemical compounds or process steps (Markush groups) in the first independent claim decreases the probability of PIV patent challenge by approximately two percentage points.

Our empirical results are in line with the theory of costly imitation pioneered by Gallini (1992) suggesting that shorter but broader patents would increase welfare. In particular, our results corroborate the prediction of the theory according to which longer patent term is an inefficient way to promote new drug development since it increases successful challenges to new drug patents, thus diluting its positive effect on the originators' innovation incentives and causing unnecessary imitation costs. To assess Gallini's (1992) policy prescription in more detail, we use our data moments in a variant of the costly imitation model. In our model the prospect of obtaining a patent incentives new drug development. Patent expiration enables generic entry resulting in savings for consumers and a loss of profit for the originator firm. There is a possibility of early generic entry via a PIV patent challenge, whose costs are affected by patent scope. Optimal patent policy must balance incentives to develop new drugs, and costs and consumer savings from generic entry.

Based on the DiD and IV estimates, we extrapolate the elasticity of PIV patent challenge with respect to effective patent length and scope to be around 3 and -1, respectively. Using *either* of these elasticity estimates in our model, we can calibrate the effects of changes in patent length and scope on innovation and welfare. Irrespective of the elasticity estimate we use, the result confirms

Gallini's (1992) prediction that shorter patent length, which is compensated for originators by broader scope, would increase welfare. This conclusion is at odds with the way Hatch-Waxman Act balances competition and innovation incentives in the pharmaceutical industry. Naturally our straightforward approach, while allowing an analysis of pharmaceutical patent policy without a need to collect or estimate difficult-to-obtain information on, e.g., R&D costs, and private and social value of new drugs, may render the precise policy assessment sensitive to model details (as discussed in Sections 7 and 8).

Literature. Our work adds to the empirical literature on the effects of patent policy–for surveys, see, e.g., Moser (2013) and Sampat (2018). Our empirical approach is inspired by Abrams (2009) and Kyle and McGahan (2012) who study the effects of TRIPS on innovation in the pharmaceutical industry, and by Sakakibara and Branstetter (2001) who study the effects of patent claim scope. Quian (2007), Lerner (2009) and Gilchrist (2016), too, find evidence that supports Gallini (1992)'s theory. We differ from these studies, e.g., in that we consider imitation as an outcome variable, the effects of AIPA, and patent policy structure. We also draw on Kuhn and Thompson (2019), Sampat and Williams (2019), Farre-Mensa et al. (2020) and Feng and Jaravel (2020) who develop similar examiner-leniency IVs as we do. Unlike these papers, we use examiner variation to validate widely used claim-scope measures in their effectiveness to prevent patent challenging by generic entrants.

Methodologically, our policy evaluation is similar in spirit to the sufficient statistics approach (Chetty, 2009; Kleven, 2021). We estimate the effects of effective patent length and scope on patent challenging and use these estimates to recover patent challenging elasticities for our patent policy analysis. Here, related papers are Denicolò (2007) and Budish et al. (2016) who use estimates from earlier literature to construct innovation elasticities for their analyses of patent policy.

Our work also relates to the mature theoretical literature on the optimal design of patent length and scope, dating back to the seminal works of Nordhaus (1969, 1972). This literature – see Denicolò (1996) for a synthesis – characterizes the conditions in which making patent protection "shorter-but-broader" or "longer-but-narrower" would increase welfare. In particular, Wright (1999) isolates imitation elasticities as key determinants of the optimal design of patent length and scope. We use quasi-experimental estimates to measure similar imitation elasticities to make policy recommendations.

2 Data

The FDA is our source of the data concerning patents protecting new drugs, successful patent challenges by generic entrants, and characteristics of new drugs. We obtain information on further patent characteristics from the USPTO and the European Patent Office (EPO). We next explain how we construct the variables used in our main empirical analyses. Online Appendix A.1 contains further details of our data construction.

2.1 Identifying New Drug Patents

We construct our sample of new drug patents from 2001 – 2013 annual editions of the Orange Book, which lists patents protecting FDA approved new drugs and their expiration dates. Our sample thus excludes drug patents that expired before 2001 and patents of drugs whose marketing authorization expired before 2001. Sample truncation arising from the exclusion of these old patents is hardly biasing our estimates since PIV patent challenges only began to grow in the late 1990s (Figure 1).

A more serious truncation bias is likely to stem from the patents with long grant lags filed at the end of our observation period, since such patents were still pending in 2013. To mitigate this truncation bias, we only use patents filed before 2009. Our results are similar without this sample restriction. Long grant lags could also cause a reverse distortion at the beginning of our observation period if a patent was filed before 1980, but our data contains only 14 such patents. We also drop all patents with grant lags exceeding five years in robustness analyses. Our final sample consists of 3517 new drug patents granted between 1980 – 2013 and listed in the Orange Book for 1484 FDA approved new applications.

2.2 Identifying Successful Challenges of New Drug Patents

Under the Hatch-Waxman Amendments, a generic firm can seek FDA approval to enter the market before the expiration of an originator firm's patent by filing a drug application with a PIV certification that the patent is invalid or noninfringed by the generic product. The generic firm must notify the originator firm of its PIV application. If the originator does not sue the generic for patent infringement within 45 days from notification, or sues but the court rules in favor of the generic or issues no ruling within 30 months, the FDA can authorize generic entry before the patent expires.

In this case, the PIV patent challenge is successful. In contrast, if the court verdict favors the originator firm, the PIV patent challenge fails, and the FDA is unable to authorize generic entry until the patent expires. See, e.g., Branstetter et al. (2016), for more details of the PIV patent challenge process.

Our outcome of interest is an indicator for whether a new drug patent listed in the Orange Book has successfully been challenged by a generic entrant at least once. We thus focus on successful PIV challenges of new drug patents, rather than challenge attempts. This focus follows the outcome of costly imitation in Gallini's (1992) theory and the empirical literature on entry (e.g., Ciliberto and Elie, 2021). To identify the successfully challenged new drug patents, we first obtain a list of approved generic drugs with a PIV certification from the FDA. To match these PIV approved generic drugs with associated challenged patents, we read their FDA approval letters. Some of the approval letters are available from the Drugs@FDA database. We obtain more approval letters by submitting the Freedom of Information Act requests to the FDA. Even though we fail to identify challenged patents for all PIV approved generic drugs, we show in Online Appendix A.1 that there is little measurement error in our outcome variable. For example, the probability of a successful PIV patent challenge at the active ingredient level is almost the same regardless of whether it is calculated based on PIV approved generic drugs or our granular patent challenge information.

Figure 1 depicts the number of new drug patents challenged by generic entrants in our sample by the year of generic drug approval. If we observe multiple successful PIV challenges to the same new drug patent, we use the one with the earliest generic drug approval date. While using a different measure, Figure 1 confirms the finding documented by, e.g., Branstetter et al. (2016) that PIV patent challenges became de facto possible only after a series of well-known legal and policy changes at the turn of the millennium.

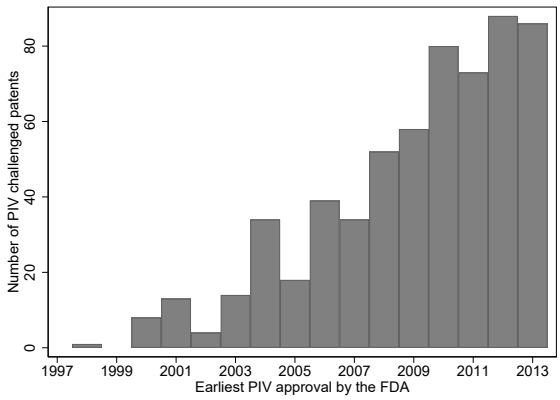


Figure 1: Successfully PIV Challenged New Drug Patents.

Notes: The figure shows the number of new drug patents that have successfully been challenged via PIV certifications by the year of generic drug approval by the FDA. We use the challenge with the earliest generic drug approval date in the case of several successful challenges to the same patent.

2.3 Measuring Effective Patent Length

We measure the effective length of a new drug patent – the period when a patent is in force and protects an approved new drug – by subtracting from the expiration date of the patent either its grant date or the approval date of a drug protected by the patent, whichever date is later. If the patent protects several new drugs, we use the earliest approval date. We use the USPTO Patent Examination Research Dataset (PatEx) to determine patent grant dates and the Orange Book to identify the dates of patent expirations and new drug approvals.

This effective patent length varies for three reasons. First, its starting date depends on whether the FDA drug approval or the USPTO patent grant comes later and is hence affected by approval and grant lags. The USPTO grant lags may also affect effective patent length even when it runs from the FDA approval date. For example, originators may wait for core patent grants before entering clinical trials (see, e.g., Budish et al., 2015). Second, legal reforms changed patent terms in the United States during our observation period. Third, various patent-specific term extensions and adjustments separately compensate for the FDA and USPTO lags. We explain these patent term reforms and USPTO term adjustments in Section 4.1.[1]

[1]The Hatch-Waxman term extensions compensating for the FDA delays, while counted in effective patent length, are less pertinent to our work as we are unaware of exogenous sources of variation that could be used in identification in our data (see, e.g., Grabowski and Vernon, 2000, for the details of these term extensions).

2.4 Measuring Patent Scope

We identify several measures of patent claim scope from the USPTO Patent Claims Research Dataset. We use the number of the phrases "selected from" in the first independent claim as a proxy for the number of "Markush groups". Such Markush groups, ubiquitous in drug patents, are lists of functionally equivalent alternatives and thus capture alternative variants of a drug protected by a patent, for example, substitutable chemical compounds or process steps. A common and acceptable form of a Markush group is *"...selected from the group consisting of A, B, and C."*.[2] We also calculate the number of "or" in the first independent claim, since the conjunction "or" might also be used to introduce variants of a drug protected by a patent.

Similarly, we calculate the number of words in the first independent claim. Some studies (e.g., Kuhn and Thompson, 2019; Marco et al., 2019; Akcigit and Ates, 2023; Lerner et al., 2021) use claim length as an inverse proxy for patent scope. However, Kuhn and Thompson (2019) argue that the use of Markush groups reverses this inverse relationship between claim length and scope in the case of pharmaceutical patents. Following, e.g., Lanjouw and Schankerman (2001, 2004) and Marco et al. (2019), we also use the number of independent claims as a proxy for patent scope.

In the light of the costly imitation theory we view patent scope as a determinant of *patent strength* (Shapiro, 2003) in the sense that broader scope should make PIV patent challenges more difficult – see Section 6 for an explicit definition. Whether or not patent scope or its proxies work in this fashion is ultimately an empirical question. We focus on patent claim scope since its legal foundation is clear: the purpose of patent claims is to mark the boundaries of patent rights (e.g., Merges and Nelson, 1990; Freilich, 2015). Patent claim scope is also at the center of PIV patent challenges. For example, some new drug patents can relatively easily be designed around by replacing one compound or process step with a slightly different one that serves the same function but is uncovered by patent claims. However, if claims are broader or interpreted more broadly by the courts, designing around becomes more difficult (Tang, 2013; Voet, 2016). Moreover, claim scope measures allow us to formulate patent-examiner specific instruments in an attempt to validate these scope measures as affecting successful PIV patent challenges (see Section 5).

[2]Using the phrase "consisting of" as a proxy for a Markush group gives similar results. For more details of Markush groups and their construction, see, e.g., Kuhn and Thompson (2019), Wagner et al. (2022), and the USPTO Manual of Patent Examining Procedure (9^{th} edition) §803.02 `https://www.uspto.gov/web/offices/pac/mpep/s803.html#d0e98237` and §2117 `https://www.uspto.gov/web/ohffices/pac/mpep/s2117.html` (accessed June 22, 2023).

Besides claim scope measures, we create indicators measuring whether a patent in our data covers an active ingredient, a method of use, or some other pharmaceutical invention such as drug formulation. We infer these patent types from the abstracts and claims of patents by using text pattern recognition algorithm and manual verification (see Online Appendix A.1). Active ingredient patents are more likely to provide stronger protection against PIV patent challenges than other new drug patents (Hemphill and Sampat, 2012).

We also use the Orange Book to create indicators measuring whether or not a patent in our sample protects a drug with new chemical entity (NCE), orphan drug, or pediatric exclusivity. Such exclusivity is granted by the FDA upon approval of a new drug. While not strictly measuring patent scope, exclusivity works in a similar fashion as patents by making generic entry substantially more difficult as it prevents generics from using originators' clinical trials data. This data exclusivity period varies across drugs and hence, it is also related to the length of protection (see, e.g., Gaessler and Wagner, 2022). New clinical investigation, NCE and orphan drug exclusivity last, respectively, three, five, and seven years. Pediatric drugs may get six months of exclusivity on the top of other exclusivity periods and patent term.

2.5 Other Patent and Drug Characteristics

We measure several other patent characteristics. We collect backward and forward patent citations from the USPTO Patent Full-Text and Image Database (PatFT), and patent family sizes from the Open Patent Services of the EPO. Citations and family size are common indicators of patent value (Lanjouw and Schankerman, 2004; Gambardella et al., 2008). Backward citations could also serve as proxy for patent scope: for example, careful documentation of prior art can make a patent difficult to invalidate on the grounds of failure to disclose prior art (Lanjouw and Schankerman, 2001; Harhoff and Reitzig, 2004). However, a high number of backward citations may also indicate a patent protecting an incremental invention, which could make the patent less lucrative for PIV patent challenges. From PatEx, we identify patents filed as continuation, continuation-in-part or divisional applications. Such "continuing patents" may differ from other patents across various dimensions (Lemley and Moore, 2004) that may affect propensity to encounter PIV patent challenges. We also retrieve the main three-digit U.S. Patent Classification (USPC) number and the filing year of each patent in our sample from PatEx.

We also measure several other characteristics of the drugs protected by each patent in our data. We determine from the Orange Book whether a patent protects a drug that is available as a capsule, an injection, a tablet, or as some less common form. Some dosage forms may be easier, e.g., to manufacture, distribute, or use by consumers (Hemphill and Sampat, 2011), which could affect generics' incentives to launch PIV patent challenges. From the Drugs@FDA database, we identify whether a patent protects a drug priority reviewed by the FDA. Such priority reviews may reflect drug value. Finally, we identify from the Orange Book the first FDA-approved active ingredient and the latest exclusivity expiration year of drugs related to each patent.

2.6 Summary Statistics

Table 1 reports the summary statistics for our sample of 3517 new drug patents. Over 17 percent of these patents have been successfully challenged via PIV certifications. Effective patent length is on average 12.6 years, varying from a month to 20 years. Using related measures, Grabowski and Vernon (2000), Hemphill and Sampat (2012), and Gilchrist (2016) report average new drug patent lengths of 11.7, 15.9 and 11.8 years, respectively.

A new drug patent has three independent claims on average. An average first independent claim contains 1 Markush group, 3 conjuctions "or", and 117 words. The measures of claim scope, citations and family size have highly skewed distributions. We therefore log transform all these count variables in our regressions. In the case of zero-valued observations, we simply add one before logging the variable, since using an inverse hyberbolic sine transformation yields similar results.

3 Evidence from Ordinary Least Squares Regressions

We set out to test the key predictions of Gallini's (1992) costly imitation theory: an increase in effective patent length increases the probability of a successful patent challenge, whereas broadening patent scope decreases it. We first estimate the following OLS regression using our patent-level data:

$$\mathbb{1}[\text{PIV challenge}_{it}] = \alpha^r + \beta \log(\text{Effective length}_i) + \gamma^{r\prime}\mathbf{X_i} + \delta_t^r + \varepsilon_{it}^r, \tag{1}$$

in which $\mathbb{1}[\text{PIV challenge}_{it}]$ is an indicator variable equaling one if new drug patent i filed in year t is successfully challenged via a PIV certification, $\log(\text{Effective length}_i)$ is the natural logarithm of the

Table 1: Summary Statistics for New Drug Patents.

	Mean	Std. Dev.	Min	Max	N
PIV challenge	0.171	0.377	0	1	3517
Effective length	12.586	3.931	0.096	20	3517
Markush groups	0.704	4.151	0	112	3485
Conjuctions "or"	3.256	9.572	0	184	3485
Words	116.890	153.337	1	2197	3485
Independent claims	3.187	3.840	1	92	3488
NCE exclusivity	0.374	0.484	0	1	3517
Orphan drug exclusivity	0.123	0.328	0	1	3517
Pediatric exclusivity	0.177	0.382	0	1	3517
Method patent	0.311	0.463	0	1	3517
Active ingredient patent	0.226	0.419	0	1	3517
Forward citations	35.705	57.889	0	1297	3517
Backward citations	34.065	63.209	0	1005	3517
Patent family size	13.418	12.031	1	51	3511
Continuing patent	0.589	0.492	0	1	3517
Priority review	0.078	0.269	0	1	3517
Tablet	0.384	0.486	0	1	3517
Capsule	0.160	0.366	0	1	3517
Injectable	0.195	0.397	0	1	3517

Notes: This table reports summary statistics for our sample of 3517 new drug patents. PIV challenge equals 0 if a patent has never been successfully challenged via a PIV certification, and 1 otherwise. Effective length is measured in years and defined as "Expiration date - max{Grant date, Drug approval date}". The third, fourth and fifth row depict the number of Markush groups (i.e., the number of the expressions "selected from"), conjunctions "or", and words, respectively, in the first independent claim. Each exclusivity indicator equals 1 if a patent covers a drug that has been awarded the corresponding exclusivity. The indicators Method patent and Active ingredient patent equal 1 if a patent protects a method of use and an active ingredient, respectively. The indicators Tablet, Capsule, and Injectable equal 1 if the drug protected by a patent has the corresponding dosage form. The Priority review indicator equals 1 if the drug protected by a patent has been priority reviewed by the FDA. The Continuing patent indicator equals 1 if a patent is filed as a continuation, a continuation-in-part or a divisional application. Independent claims, Backward citations, Forward citations, and Patent family size give the number of independent claims included in a patent in our sample, of earlier patents cited by a patent in our sample, of later patents citing a patent in our sample, and of countries where the same patent in our sample has been filed, respectively.

effective length of patent i, the vector $\mathbf{X_i}$ includes measures of patent scope and other characteristics except for patent filing year fixed effects, which control for time trends in patent challenging and are captured by δ_t^r. We cluster standard errors at the level of patents protecting the same active ingredient, since challenge decisions can be correlated if multiple patents cover the same new drug.

Table 2 presents estimates from various specifications of the model (1). In column (1), we presents the estimates for effective patent length from a specification without $\mathbf{X_i}$. The specification in column (2) adds into $\mathbf{X_i}$ characteristics of the drugs protected by each patent: dosage form, the type of FDA exclusivity, and fixed effects for exclusivity expiration year to account for a possible negative bias in our length estimates due to different exclusivity periods.

Finally, column (3) adds patent characteristics: the log number of Markush groups in the first independent claim, indicators for method, active ingredient, and continuing patents; log forward and backward citations; log patent family size, and fixed effects for the main U.S. patent class to control for technology specific idiosyncrasies in patent challenging.

We find a statistically and economically significant relationship between the probability of PIV patent challenge and effective patent length. The estimated coefficient in column (1) implies that doubling the effective term of a patent is associated with a 13 percentage point increase in the probability of PIV patent challenge. The relationship becomes economically weaker but remains statistically significant after adding various controls – see columns (2) and (3) of Table 2.

Table 2 shows a negative, but statistically insignificant, relationship between the probability of PIV patent challenge and the number of Markush groups. Using other measures of claim scope as an alternative for the Markus groups yields similar results, except that the number of conjunctions "or" in the first independent claim gets a statistically significant negative coefficient (see Table A3 in Online Appendix A.3. Estimates for the other measures of the difficulty of patent challenging suggest stronger associations: The probability of PIV patent challenge is over six percentage points lower for patents covering active ingredients, and three percentage points lower for patents covering new methods of use compared to patents protecting, say, new drug formulations. Furthermore, the probability of PIV patent challenge is around eight to nine percentage points lower for patents protecting drugs with NCE or orphan drug exclusivity than for patents protecting drugs with no exclusivity or with clinical investigation exclusivity of shorter duration. In contrast, pediatric exclusivity appears to be associated with an increase in the probability of PIV patent challenge.

Pediatric exclusivity works differently from other exclusivities as it is added to other exclusivity periods and patent term.

The negative estimate of the effect of backward citations might support a negative association of the probability of PIV patent challenge with patent scope, but it could also suggest a positive association with patent value. This positive association between the probability of PIV patent challenge and patent value is also supported by the positive and statistically significant coefficients of patent family size, priority review indicator, and indicators for common orally administered dosage forms (tablet and capsule).

Overall the findings from the OLS regressions suggest that the probability of PIV patent challenge positively correlates with effective patent length and negatively with patent claim scope and other measures of difficulty of patent challenging. However, while we attempt to control for, e.g., drug and patent value, these results may still be driven by unobserved heterogeneity.

Table 2: Probability of Successful PIV Patent Challenge and Patent Characteristics: OLS Estimates.

	(1)	(2)	(3)
log(Effective length)	0.129	0.069	0.065
	(0.015)	(0.013)	(0.013)
New chemical exclusivity		-0.091	-0.075
		(0.025)	(0.026)
Orphan drug		-0.093	-0.092
		(0.024)	(0.025)
Pediatric exclusivity		0.087	0.086
		(0.035)	(0.035)
Priority review		0.179	0.178
		(0.046)	(0.047)
Tablet		0.181	0.183
		(0.027)	(0.028)
Capsule		0.106	0.112
		(0.034)	(0.035)
Injectable		-0.041	-0.033
		(0.024)	(0.024)
log(Markush groups+1)			-0.012
			(0.012)
Method patent			-0.029
			(0.017)
Active ingredient patent			-0.063
			(0.021)
log(Forward citations+1)			0.003
			(0.007)
log(Backward citations + 1)			-0.021
			(0.008)
log(Patent family size)			0.010
			(0.006)
Continuing patent			0.024
			(0.015)
Filing year FE	×	×	×
Exclusivity expiry year FE		×	×
USPC FE			×
Mean dep. variable	0.171	0.171	0.173
Observations	3517.000	3517.000	3483.000
R-squared	0.065	0.224	0.237

Notes: This table reports coefficients from OLS regressions of the probability of successful PIV patent challenge on effective patent length, measures of patent scope, and controls. FE stands for fixed effects. Standard errors, in parentheses, are clustered at the level of patents protecting the same drug.

4 Effect of Patent Length

In Section 4.1, we describe how two patent policy reforms, TRIPS of 1994 and AIPA of 1999, affected the effective lengths of patents depending on their grant lags and filing date. In Section 4.2, we estimate the effects of the reforms and the resulting changes in effective patent length on the probability of PIV patent challenge using difference-in-differences regressions.

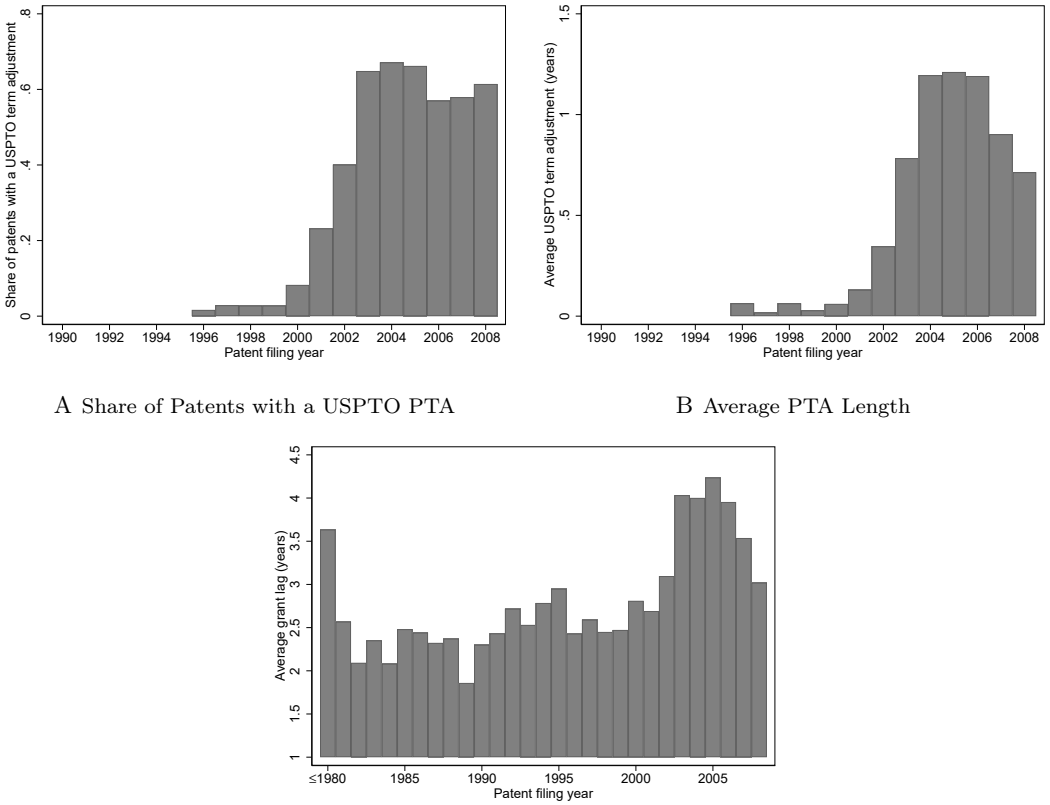
4.1 Patent Term Reforms in the United States and Pharmaceutical Patents

Prior to TRIPS, the United States had a 17-year standard patent term counting from the grant date. TRIPS introduced a 20-year standard patent term measured from the filing date pertaining to patents filed on and after June 8, 1995. For patents filed prior to June 8, 1995, the standard patent term was changed to *either* the new 20-year term from filing *or* the old 17-year term from the grant, whichever expires later. Patents that were issued prior to June 8, 1978, were kept in the old 17-year term regime, but our sample includes no such old patents.

This change in the standard patent term due to TRIPS treats patents differently depending on whether or not they are granted within three years from filing: Patents granted within three years from filing receive the same 20-year standard term from filing regardless of whether or not they are filed before or after TRIPS. In contrast, patents with grant lags exceeding three years filed before TRIPS received the 17-year standard term from the grant date, which fully compensates for grant lags. But similar patents filed after TRIPS receive the 20-year standard term from filing, thus losing some of effective protection time.

To compensate patentees for this loss in effective patent life because of delays in the USPTO approval process, TRIPS also introduced PTAs (which were initially called patent term extensions). These PTAs only apply to patents filed after TRIPS, and can add a maximum of five years to the patent term. The USPTO calculates the length of a PTA automatically, taking into account only certain delays caused by the USPTO itself. Initially eligible delays were limited but, subsequently, AIPA expanded the list of reasons which may give rise to PTAs for patents filed on and after May 29, 2000. In particular, AIPA introduced compensation for grant lags exceeding three years, thus mitigating the adverse impact of TRIPS on the effective length of patents with grant lags exceeding threeg years.

We determine PTAs and grant lags for our sample of drug patents from PatEx. Panels A and B of Figure 2 document how PTAs were rare and short before AIPA, which increased their provision and length substantially. The share of patents with a PTA rises from 2 percent in 1996 to 66 percent in 2005 (panel A). The average PTA length increases from less than a month in 1996 to around 15 months in 2005 (panel B). Even after AIPA, the increase in the length of PTAs was gradual, reflecting increasingly slow patent prosecution at the USPTO in the early years of the millenium (panel C). (The long grant lags observed in the early years in panel C are a consequence of the truncation of our sample discussed in Section 2.1.)



A Share of Patents with a USPTO PTA

B Average PTA Length

C Average Grant Lag

Figure 2: USPTO Grant Lags and PTAs of Pharmaceutical Patents.

Notes: Panel A of this figure shows the share of new drug patents in our sample with a USPTO PTA by patent filing year. Panel B shows the average length of a PTA in our sample by patent filing year, including patents without PTAs. Panel C figure shows the average grant lag of new drug patents in our sample by patent filing year.

Comparing panels A and B with panel C indicates that PTAs fail to fully compensate for the adverse impact of TRIPS on the effective length of new drug patents with long grant lags, especially before AIPA. To confirm this suggestion, we regress effective patent length on grant lag separately

for the periods of pre-TRIPS, between TRIPS and AIPA, and post-AIPA. Figure 3 shows the results: Before TRIPS, effective patent length is relatively invariant to grant lag. TRIPS disproportionately shortens the effective length of patents prosecuted over three years, especially before AIPA, which partially restores the effective length of such patents.

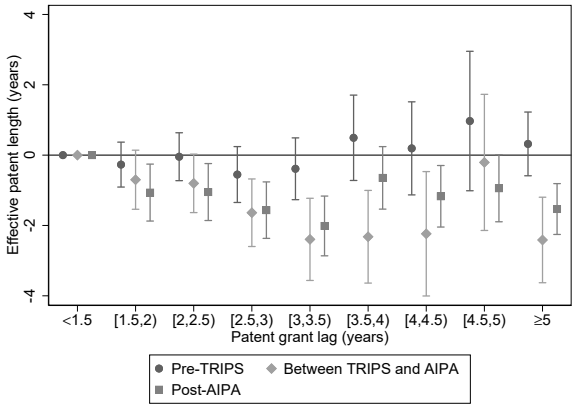


Figure 3: Effective Patent Length Before and After TRIPS and AIPA.

Notes: This figure shows OLS estimates and 95 percent confidence intervals of the raw data relationships between effective patent length (in the y-axis) and patent grant lag (in the x-axes). We estimate the relationships separately for three time periods: The pre-TRIPS period includes patents filed before June 8, 1995, the period between TRIPS and AIPA includes patents filed between June 8, 1995 and May 29, 2000, and the post-AIPA period includes patents filed on or after May 29, 2000. Patents with grant lags exceeding five years are binned together (denoted by "5-" on the x-axis). In each regression, the comparison group consists of patents granted less than 1.5 years from the filing date. Each dot shows the averages of the x- and y-axes variables within each equal-sized bin.

According to the costly imitation theory, these changes in the effective length of patents with long grant lags should be matched by a decrease in the probability of PIV patent challenge especially between TRIPS and AIPA. Consistent with this prediction, Figure 4 indicates that between TRIPS and AIPA, the probability of PIV patent challenge is lower for patents prosecuted over three years compared to patents with shorter grant lags. There is no similar difference between the two patent groups in other periods. These patterns in PIV patent challenges and grant lags motivate our research design.

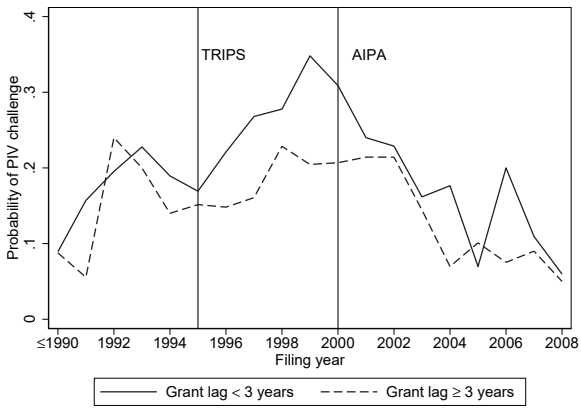


Figure 4: Probability of Successful PIV Patent Challenge by Patent Filing Year and Grant Lag.

Notes: This figure shows the probability of successful PIV patent challenge by patent filing year. The dashed and solid lines depict the groups of patents with a grant lag more and less than three years, respectively. Patents filed before the end of year 1990 are binned together (denoted by "-1990" on the x-axis). Those oldest patents in our sample encounter only few successful PIV patent challenges with no systematic difference depending on the grant lag.

4.2 Difference-in-Differences Estimations and Results

We estimate the following DiD model using our patent-level data:

$$\begin{aligned} \mathbb{1}[\text{PIV challenge}_{it}] \;=\; & \alpha^s + \beta_1 \mathbb{1}[\text{Grant lag}_i \geq 3\text{years}] \\ & + \beta_2 \mathbb{1}[\text{Grant lag}_i \geq 3\text{years}] \times \mathbb{1}[\text{Post-TRIPS}_i] \\ & + \beta_3 \mathbb{1}[\text{Grant lag}_i \geq 3\text{years}] \times \mathbb{1}[\text{Post-AIPA}_i] + \gamma^{s\prime}\mathbf{X_i} + \delta_t^s + \varepsilon_{it}^s, \end{aligned} \tag{2}$$

in which $\mathbb{1}[\text{Grant lag}_i \geq 3\text{years}]$ is an indicator variable equaling one if patent i has at least a three-year grant lag, $\mathbb{1}[\text{Post-TRIPS}_i]$ is an indicator variable equaling one if the filing date of patent i is on or after June 8, 1995, and $\mathbb{1}[\text{Post-AIPA}_i]$ is an indicator variable equaling one if the filing date of patent i is on or after May 29, 2000. The coefficients of interest, β_2 and β_3, measure changes in the probability of PIV patent challenge after TRIPS and AIPA, respectively, for patents prosecuted at least three years (treatment group), compared to other patents (control group).

Table 3: Probability of Successful PIV Patent Challenge, Effective Patent Length, and Patent Law Changes by Patent Grant Lag.

Outcome	PIV challenge (1)	PIV challenge (2)	PIV challenge (3)	PIV challenge (4)	PIV challenge (5)	log(Effective length) (6)
Grant lag ≥3 years	-0.021	0.016	0.026	0.021	0.022	0.044
	(0.027)	(0.025)	(0.025)	(0.026)	(0.028)	(0.028)
Grant lag ≥3 years, Post-TRIPS	-0.080	-0.114	-0.106	-0.099	-0.100	-0.171
	(0.042)	(0.039)	(0.039)	(0.044)	(0.041)	(0.043)
Grant lag ≥3 years, Post-AIPA	0.072	0.084	0.072	0.066	0.072	0.100
	(0.041)	(0.037)	(0.037)	(0.042)	(0.039)	(0.045)
Mean dependent variable	0.171	0.171	0.173	0.167	0.177	2.457
Observations	3517	3517	3483	2998	3066	3517
R-squared	0.045	0.221	0.234	0.234	0.240	0.020
Filing year FE	×	×	×	×	×	×
Drug controls		×	×	×	×	
Exclusivity expiration year FE		×	×	×	×	
Patent controls			×	×	×	
USPC FE			×	×	×	
Excluded filing years	None	None	None	1995, 2000	None	None
Grant lag	Any	Any	Any	Any	≤ 5 yrs	Any

Notes: This table reports estimates of the effects of TRIPS and AIPA on the probability of successful PIV patent challenge and effective patent length. Columns (1)–(5) show coefficients from an OLS regression of the PIV challenge indicator and column (6) of log(Effective length) on three indicators for patents with a grant lag exceeding three years and controls. FE stands for fixed effects. Drug controls include the indicators NCE exclusivity, Orphan drug exclusivity, Pediatric exclusivity, Priority review, Capsule, Injectable, and Tablet. Patent controls include log(Markush groups+1), log(Backward citations+1), log(Forward citations+1), log(Patent family size), and the indicators Active ingredient patent, Method patent, and Continuing patent. The samples in columns (4) and (5) exclude patents filed in years 1995 and 2000 and patents with a grant lag exceeding five years, respectively. Standard errors, in parentheses, are clustered at the level of patents protecting the same active ingredient.

Column (1) of Table 3 reports coefficients from a regression that uses the full sample of new drug patents but only controls for filing year fixed effects. Patents prosecuted at least three years experience an eight percentage point reduction in the probability of PIV patent challenge after TRIPS initially, compared to other patents. AIPA almost neutralizes this effect: the corresponding increase in the probability of PIV patent challenge is seven percentage points after AIPA.

Column (2) of Table 3 reports results after controlling for characteristics of the drugs protected by the patents in our sample, including the exclusivity expiration year fixed effects. Column (3) shows results from a specification that further controls for patent characteristics, including the main U.S. patent class fixed effects. These two specifications attempt to account for potential differences between the treatment and control groups stemming from observable characteristics, unobserved

heterogeneity, and possible compositional changes over time across these groups. Adding these controls makes the effects of TRIPS and AIPA stronger and more precisely estimated compared to column (1).

We also estimate the DiD model of equation (2) using a sample that excludes patents filed in 1995 and 2000. This sample restriction attempts to address the potential anticipation effects of the TRIPS and AIPA reforms. For example, applicants expecting a decrease in patent terms due to TRIPS could have advanced patent filing and, analogously, applicants expecting an increase in patent terms due to AIPA could have postponed filing. The coefficient estimates from this regression, reported in column (4) of Table 3, remain similar to the baseline estimates, suggesting no clear anticipation effects. The coefficient β_3 is, however, less precisely estimated, perhaps due to a smaller sample.

Next, we estimate the DiD model using a sample of patents prosecuted within five years. This sample restriction further mitigates concerns arising from possible compositional changes in the treatment group. For example, long grant lags observed in the first and the last years of our data increase the number of patents falling into the treatment group (see panel C of Figure 2). This restriction also excludes the patents for which the five-year maximum length of PTAs is binding. The results from this estimation reported in column (5) are similar to the baseline results.

Finally, column (6) confirms the raw data relationships described by Figure 3: Using DiD regressions (when only controlling for patent filing year fixed effects), we find that TRIPS decreases the effective length of patents prosecuted at least three years by 17 percent (2 years compared to the mean of 12.6 years), whereas AIPA increases the effective patent length by 10 percent. As shown by columns (1)–(5), these changes in effective patent length following TRIPS and AIPA coincide with, respectively, a 8–11 percentage point decrease and a 7–8 point increase in the probability of PIV patent challenge. Overall, these results are consistent with the costly imitation theory suggesting that the probability of PIV patent challenge is increasing with effective patent length.

Robustness. While we use a rich set of controls, and while the raw data trends presented in Figure 4 suggest otherwise, the DiD estimates of the effects of TRIPS and AIPA might nonetheless reflect differential pre-trends in the probability of PIV patent challenge between treated and untreated patents. To mitigate this concern, we estimate an event study specification detailed in Online Appendix A.2. The results show no pre-trend in the outcome. The event study estimates

are also consistent with the DID estimates, suggesting a decrease in the probability of PIV patent challenge between TRIPS and AIPA.

We make many other robustness checks in Online Appendix A.2. To address the concern that longer patent duration would increase PIV patent challenges even if these challenges were random, we control for the length of exposure of a patent to PIV patent challenges. We also show that the estimated effects on the probability of PIV patent challenge cannot be explained by differential changes in patent scope or value after TRIPS and AIPA, nor by the introduction of provisional applications in 1995. As mentioned, multiple patents can cover the same new drug and we therefore cluster standard errors at the level of patents covering the same active ingredient to allow for unobservables in their PIV patent challenges to be correlated. We additionally identify the chain of patents covering the same new drug and exclude some or all but one of the patents in the chain. When we estimate equation (2) using different subsamples and specifications, the effects of TRIPS and AIPA only become stronger and more precisely estimated.

We also find additional evidence of longer effective patent length encouraging PIV patent challenges: TRIPS disproportionately shortened effective terms of continuing patents, leading to a more negative estimate of β_2 using the sample of these patents. Finally, we document how AIPA also mandated earlier disclosure of patent applications and discuss how this change may affect the interpretation of effective patent length in the post-AIPA period.

5 Effect of Patent Scope

In identifying the effect of a change in patent scope on PIV patent challenges we cannot resort to an ideal experiment of random assignment of scope over patents. Inspired by Kuhn and Thompson (2019), Sampat and Williams (2019), Farre-Mensa et al. (2020), and Feng and Jaravel (2020) we instead develop IVs for patent scope based on the "leniency" of patent examiners. Our approach exploits the differences across examiners in their propensity to grant broader or more claims as a source of variation in patent scope, together with the assignment of patent applications to examiners at the USPTO. Previous research (e.g., Cockburn et al., 2003; Lemley and Sampat, 2012) indicates differences in the examiners' decision making which translates into different patent outcomes. Since patent prosecution typically consists of several rounds of claim rejections and modifications re-

quired by an examiner (Kuhn and Thompson, 2019; Marco et al., 2019), systematic differences across examiners plausibly generate systematic differences in patent claim scope. Furthermore, prosecution-history estoppel enhances the role of patent examination in affecting new drug patent claim scope at the core of PIV patent challenges (Tang, 2013).

The second stage of our two-stage least squares (2SLS) analysis consists of estimations of equation (1) using instrumented scope measures. We instrument four measures of scope of a new drug patent: the number of independent claims, the numbers of Markush groups, conjunctions "or" and words in the first independent claim (the last measure follows Kuhn and Thompson, 2019). For each scope measure x_{ijt} of new drug patent i reviewed by examiner j and filed in year t, we construct the corresponding instrument z_{ijt} as

$$z_{ijt} = \frac{\sum_{\tau=\underline{\tau}_j}^{t-1} \sum_{k=1}^{n_{j\tau}} x_{kj\tau}}{\sum_{\tau=\underline{\tau}_j}^{t-1} n_{j\tau}}, \tag{3}$$

in which $x_{kj\tau}$ is the scope measure of patent k reviewed by examiner j and filed in year τ, $n_{j\tau}$ is the number of patents reviewed by examiner j in filing year τ, and $\underline{\tau}_j$ is the earliest filing year of any patent reviewed by examiner j. Hence, z_{ijt} gives the "examiner j's historical average" – the cumulative average of a scope measure over all patents assigned to examiner j up to one year preceding filing year t of new drug patent i. Although we include the previous new drug patents granted by examiner j in z_{ijt}, we exclude their parents and continuations, which are usually assigned to the same examiner (Righi and Simcoe, 2019).

Table 4 shows marked variation across the examiners of our sample patents in granting broader or more claims. For example, the average number of words in the first independent claim is 147, but the toughest examiner only allows 11 words on average, whereas the most lenient examiner allows 456 words on average. This variation unlikely arises from a small sample size: as also shown by Table 4, the average number of patents reviewed by an examiner is 626.

Table 4: Heterogeneity Across Examiners in Patent Measures.

	Mean	Std. Dev.	Min	Max
Examiner average of Markush groups	0.786	1.037	0.000	7.000
Examiner average of conjunctions "or"	3.633	4.100	0.000	21.737
Examiner average of words	147.115	59.068	11.000	455.776
Examiner average of independent claims	2.397	0.443	1.000	4.455
Patents reviewed by an examiner	625.515	660.926	1	3653

Notes: This table reports summary statistics for 579 examiners who have reviewed the new drug patents in our sample. Each of the first four rows shows summary statistics for a examiner-specific patent scope measure, averaging over all patents reviewed by an examiner of a new drug patent in our sample. The last row shows summary statistics for the number of patents reviewed by an examiner of a new drug patent in our sample.

In the first-stage of our 2SLS analysis, we regress each claim scope measure on the corresponding instrument and controls. In the main 2SLS specifications, we control for the same patent and drug characteristics as in column (3) of Table (2) except for the instrumented claim scope measure. While some of these controls (e.g., patent filing year and U.S. patent class fixed effects) may also capture examiner specialization, we also add USPTO Technology Center fixed effects. Technology Centers are responsible for examination in broad technological areas. Each Technology Center typically contains a few dozen Art Units, which are groups of examiners specializing in narrow technology areas. Within a Technology Center, a patent application is assigned to an Art Unit and finally to an examiner. Since we only observe a small number of new drug patents per Art Unit, we use Art Unit fixed effects only in robustness analysis.

The exclusion restriction in our setting holds if, conditional on covariates, examiners' propensity to grant broader claims is uncorrelated with such application characteristics, e.g., drug or patent value or quality, that correlate with PIV patent challenges. The validity of this exclusion restriction is supported by a growing literature (e.g., Lemley and Sampat, 2012; Sampat and Williams, 2019; Kuhn and Thompson, 2019; Farre-Mensa et al., 2020; Feng and Jaravel, 2020) although, e.g., Righi and Simcoe (2019) are more critical. These previous studies indicate that examiner assignment is independent of application characteristics at the time of filing. For example, examiner assignment is based on the last digit of the application number in Art Units. Such assignment plausibly implies that examiner characteristics are uncorrelated with the value or quality of applications. While Righi and Simcoe (2019) show that examiners specialize in narrow technology fields, they find no evidence

of more valuable or broader applications being allocated to certain examiners. Moreover, since our new drug patents form a relatively homogeneous technology field, examiners might be less likely to be specialized within this sample. Nevertheless, the validity of this exclusion restriction is debatable and our IV results must be interpreted cautiously.

Table 5 reports the 2SLS regression results. The first stage coefficients of panel B and F-statistics suggest strong instruments. Estimates of the instrumented scope measures of panel A suggest a negative effect of broader patent scope on the probability of PIV patent challenge: A 10 percentage increase in the Markush groups in the first independent claim decreases the probability of PIV patent challenge some two percentage points. A higher number of Markush groups may thus indicate broader scope by reflecting a higher number of variants of the drug protected by the patent. The number of conjuctions "or", which might also reflect different variants of the drug protected by the patent, has a smaller effect. Additional words in the first independent claim perform similar to Markush groups, supporting a positive relationship between claim length and scope in the case of pharmaceutical patents as argued by Kuhn and Thompson (2019). These coefficients of the scope measures are statistically significant but smaller in magnitude compared to the OLS estimates reported in Table A3 in Online Appendix A.3. Such an upward bias in the OLS estimates could arise, e.g., if originators seek broader protection for more valuable drugs which, at the same time, attract more PIV patent challenges.

The question of whether or not the number of independent claims is a useful proxy for patent claim scope has been debated in the literature (see, e.g., Kuhn and Thompson, 2019 and Marco et al., 2019 for different points of view). Our coefficient estimate of the number of independent claims is close to zero in magnitude and statistically insignificant, suggesting that additional independent claims in a pharmaceutical patent fail to protect against PIV patent challenges.

Table 5: Probability of Successful PIV Patent Challenge and Patent Scope: IV Estimates.

	(1)	(2)	(3)	(4)
PANEL A: Second stage estimates				
Instrumented variables:				
log(Markush groups+1)	-0.220			
	(0.090)			
log(Conjunctions "or"+1)		-0.105		
		(0.052)		
log(Words)			-0.237	
			(0.095)	
log(Independent claims)				0.069
				(0.132)
PANEL B: First stage estimates				
Instruments:				
log(Examiner historical average of Markush groups+1)	0.239			
	(0.048)			
log(Examiner historical average of conjunctions "or"+1)		0.225		
		(0.034)		
log(Examiner historical average of words)			0.238	
			(0.062)	
log(Examiner historical average of independent claims)				0.242
				(0.073)
Observations	3445	3445	3445	3447
First stage F-statistic	25.286	44.860	14.523	10.934
Technology Center FE	×	×	×	×
Filing year FE	×	×	×	×
Drug controls	×	×	×	×
Exclusivity expiration year FE	×	×	×	×
Patent controls	×	×	×	×
USPC FE	×	×	×	×

Notes: This table reports 2SLS estimates of the effects of patent scope on the probability of successful PIV patent challenge. Panel A shows the main coefficient from the second stage regressions of the PIV challenge indicator on the instrumented scope measures and controls. Panel B shows the main coefficient from the first stage regressions of the scope measures on the corresponding instruments and controls. The first stage F-statistic test is on the excluded instruments. FE stands for fixed effects. Drug controls include the indicators NCE exclusivity, Orphan drug exclusivity, Pediatric exclusivity, Priority review, Capsule, Injectable, and Tablet. Patent controls include log(Effective length), log(Backward citations+1), log(Forward citations+1), log(Patent family size), and the indicators Active ingredient patent, Method patent, and Continuing patent. We use a full sample of new drug patents for the regressions, and construct the instruments using data on all granted patents reviewed by the examiners of these new drug patents. Robust standard errors are reported in parentheses.

Robustness. The IV regression results appear to suggest that broadening patent claim scope hampers PIV patent challenges. We assess the robustness of the results to specification changes in Online Appendix A.3. The results remain unchanged when we exclude control variables (Table A4) or include of an additional control for trends varying with the Technology Center (Table A5).

We also use Art Unit fixed effects instead of Technology Center fixed effects (Table A6). The number of patents per Art Unit is typically small, only 21 on average. Reflecting this challenge, the point estimates from these specifications, while similar in magnitude, are less precise than from the main specifications. Only the coefficient of the number of words remains statistically significant.

We have also experimented with modified instruments. Our results do not change if we exclude all new drug patents reviewed by examiner j from z_{ijt}, or if we allow τ to run from $\underline{\tau}_j$ to $\min\{\bar{\tau}_j, 2009\}$, in which $\bar{\tau}_j$ is the last filing year of any patent reviewed by examiner j, and only impose $k \neq i$ in equation (3). Also, using two alternative measures of scope, the numbers of characters and the phrase "consisting of" in the first independent claim, and corresponding alternative instruments, yields similar IV regression results. Our results are also robust to using a binary variable of whether or not there is at least one Markush group in the first independent claim.

6 A Model

Overall, our empirical results support the claims of Gallini's (1992) costly imitation theory according to which longer patent terms increase the probability of a successful patent challenge and broader scope reduces it. While these results alone are informative for patent policy design, we construct a variant of the costly imitation model with flexible cost functions to explore policy implications for the pharmaceutical industry.

6.1 Assumptions

Consider a pharmaceutical market with an originator drug producer (firm B) and a generic drug producer (firm G). The originator firm can invest in developing a new drug, automatically protected by a patent. The generic firm can invest in challenging the new drug patent. Ex ante licensing is not possible.

Some of these simplifying assumptions reflect industry practices. For example, originators tend to rely on patents (see, e.g., Hall et al., 2014), partly because of the mandatory disclosure stipulated by the Hatch-Waxman Act significantly limits the originators' possibilities to use secrecy (Tang, 2013). There is also little evidence of generic entry via licensing prior to PIV patent challenges. On the contrary, our data, as also, the results e.g., in Higgins and Graham (2009), Hemphill and Sampat

(2011) and Branstetter et al. (2016), show that PIV patent challenges occur frequently. However, although the Hatch-Waxman Act awards the first generic entrant via a PIV patent challenge a 180-day exclusivity period during which no further generic entry is allowed, there may in practice be many generic entrants. Allowing many generic challengers would complicate the model without necessarily changing its empirical implications.[3]

The success $Y_f : \{0,1\} \to \{0,1\}$, $f = B, G$, of the firm f's investment has a Bernoulli distribution with parameter $p_f \in [0,1]$. Time $t \in [0,\infty)$ is continuous but, for brevity, we assume that the firms act sequentially at $t = 0$ (as in Gallini, 1992) by directly choosing the their success probabilities p_f. (Alternatively, we may think that the firms choose an investment from a collection of projects indexed by p_f.) The associated investment cost functions $C_f : [0,1] \to [0,\infty)$ are twice continuously differentiable with the standard properties $\partial C_f/\partial p_f > 0$ and $\partial^2 C_f/\partial p_f^2 > 0$ for $p_f > 0$, and $C_f(0) = \partial C_f(0)/\partial p_f = 0$. The cost functions are sufficiently convex to satisfy second-order conditions. Gallini (1992) only allows for fixed costs.

We consider two patent policy variables, length $T \in [0,\infty)$ and scope $b \in [0,\infty)$. Following a common practice (see Budish et al., 2015, for an exception), we assume that a marketing authorization and a patent are granted to a new drug simultaneously upon the investment success realization $y_B = 1$ at $t = 0$, from which patent length is counted. Hence, T reflects *effective* patent length.

Following Gallini (1992), we assume that if a new drug patent expires at $t = T$, then $\partial C_G(p_G, b)/\partial b > 0$ for $t < T$ and $b = 0$, with $C_G(p_G, 0) = 0$ $\forall p_G$, for $t \geq T$ – see Maurer and Scotchmer (2002) for a rationale for this modeling of patent scope. In words, a broader patent makes imitation via successful patent challenging more difficult whereas the patent expiration makes imitation costless. For example, developing a noninfringing substitute could become more difficult with broader new drug patent claims that cover, e.g., variants of the drug through Markush groups (see Sections 2.4 and 5). Some imitation costs stemming from the FDA approval process and associated litigation may also be related to the scope of new drug patents. In *Allergan, Inc v. Exela Pharmasci, Inc*, for

[3]With one generic challenger, we do not need to differentiate between successful non-infringement and invalidity challenges. A generic firm has a stronger incentive to file an infringement challenge since the challenged patent is still valid and enforceable against other generics. The prior work (see, e.g., Gallini, 1992; Wright, 1999; Maurer and Scotchmer, 2002) suggest that allowing multiple entrants via non-infringement challenges would not qualitatively change the results (unless there is substantial cross-ownership between the originator and generic entrants – see Newham et al., 2018). If generic entry occurs via invalidation rather than non-infringement, however, then a waiting game among generics might arise (Henry and Ponce, 2011). But in our context the 180-day exclusivity period dilutes the generics' incentives to wait.

example, the Federal Circuit found Exela's design-around formulation sufficiently different to avoid infringement of Allergan's patent, but the same difference prompted the FDA to request Exela to perform expensive bioequivalence studies (Voet, 2016). Higgins and Graham (2009) report the average cost of a PIV patent challenge to be $5 million. While the assumption of costless post-patent entry is made for simplicity, the Hatch-Waxman Act greatly reduced the costs and lags of generic entry after patent expiration (see, e.g., Grabowski and Vernon, 2000; Tang, 2013).

After the realizations $y_f \in \{0,1\}$ of Y_f, $f = B, G$, the firms compete in the market. The net cash flow from selling a drug is $\widetilde{\pi}_N \in [0, \infty)$, in which subscript $N \in \{0,1,2\}$ denotes the number of competing drugs in the market. The pharmaceutical market will exist only if $y_B = 1$; otherwise $N = 0$ and $\widetilde{\pi}_0 = 0$. Conditional on $y_B = 1$, our assumptions imply that $N = 1$ only if $t < T$ and $y_G = 0$; otherwise $N = 2$. As usually, $\widetilde{\pi}_1 > 2\widetilde{\pi}_2$. (The assumption of equal net cash flows upon generic entry can be relaxed at the cost of complicating the notation.)

Reminiscent of Wright (1999), the shape of $C_G(p_G, b)$ turns out to be a key determinant of our patent policy evaluation. We introduce the following definitions:

$$\epsilon_p(p_G) := p_G \frac{\partial^2 C_G / \partial p_G^2}{\partial C_G / \partial p_G}, \tag{4}$$

and

$$\epsilon_b(p_G) := p_G \frac{\partial^2 C_G / (\partial p_G \partial b)}{\partial C_G / \partial b}. \tag{5}$$

The elasticity of the marginal cost of patent challenging, $\epsilon_p(p_G)$, provides a measure of the convexity of the generic's cost function. The convexity itself implies that $\epsilon_p(p_G) > 0$ at least for $p_G \in (0,1)$ and, for simplicity, we make $\epsilon_p(p_G) > 0 \; \forall p_G$. In turn, $\epsilon_b(p_G)$ is the elasticity of the impact of patent scope on patent challenging costs. Since $\partial C_G / \partial b > 0$, the sign of $\epsilon_b(p_G)$ is given by the sign of $\partial^2 C_G / (\partial p_G \partial b)$.

Assumption 1. $\frac{\partial^2 C_G}{\partial p_G \partial b} > 0$.

According to Assumption 1, the effect of patent scope on patent challenging costs is the stronger the easier is patent challenging. In Online Appendix A.4 we relax Assumption 1 and find counterintuitive effects. For example, if $\partial^2 C_G / (\partial p_G \partial b) < 0$, the probability of a successful patent challenge *increases* with patent challenging costs and scope, which is not only odd but also against our empir-

ical results. Unsurprisingly, this assumption is often implicit in the costly imitation models. Also, our results concerning patent length are independent of Assumption 1 (see Online Appendix A.4).

We consider a two-stage game in which the originator firm first chooses $p_B(b,T) \in [0,1]$. In the second stage, after observing y_B, the generic firm chooses $p_G(y_B,b,T) \in [0,1]$, which results in the outcome y_G. The firms collect their payoffs depending on the realizations of Y_f, $f = B, G$, and patent length T. Denote the firm f's expected profit by Π_f. A subgame perfect equilibrium of this game is a pair $(p_B^*, p_G^*(y_B(p_B)))$ such that for $y_B(p_B) \in \{0,1\}$, $p_G^*(y_B(p_B)) = \arg\max_{p_G \in [0,1]} \Pi_G(p_G, y_B(p_B))$ and $p_B^* = \arg\max_{p_B \in [0,1]} \Pi_B(p_B, p_G^*(y_B(p_B)))$. In what follows, we present the main arguments leading to our results, relegating to Appendix 1 their technical proofs.

6.2 Equilibrium Analysis

Consider the second stage of the game after the realization of Y_B. Clearly, if $y_B(p_B) = 0$, new drug market fails to arise, and $p_G^* = 0$. We therefore focus on the subgame in which the market exists, $(p_B^*, p_G^*(y_B(p_B) = 1))$, and suppress the argument $y_B(p_B)$ for brevity.

Given $y_B(p_B) = 1$ the generic firm's problem can be expressed as

$$\max_{p_G \in [0,1]} \Pi_G = p_G \int_0^\infty e^{-rt} \widetilde{\pi}_2 dt + (1 - p_G) \int_T^\infty e^{-rt} \widetilde{\pi}_2 dt - C_G(p_G, b), \tag{6}$$

in which $r \in (0, \infty)$ denotes the firms' common discount rate. The first integral on the right-hand side of equation (6) captures the generic's profits if, with probability p_G, it successfully challenges the new drug patent. The second integral captures the profits if, with probability $1 - p_G$, the patent challenge fails and the generic entry is postponed until the patent expiration.[4] The last term captures the costs of patent challenging.

Using $\pi_2 := \widetilde{\pi}_2 / r$ we can write the first-order condition for the problem (6) as

$$\left(1 - e^{-rT}\right) \pi_2 - \frac{\partial C_G(p_G^*, b)}{\partial p_G} = 0. \tag{7}$$

Equation (7) identifies for each patent policy $(b,T) \in [0,\infty)^2$ a unique probability that the new

[4] As Tang (2013, p.1083) writes "...the generic manufacturer can decide to push market entry before the patent expires by filing a PIV challenge...Even if it loses the PIV challenge, it can still market its drug immediately after the patent expires..."

drug patent is successfully challenged. We may hence consider a new drug patent "probabilistic" (Lemley and Shapiro, 2005), with the endogenous *strength* of $1 - p_G^*(b,T)$. As shown by equation (7) the mapping $p_G^*(b,T)$ is qualitatively invariant to many model details. On the other hand, modifications to the model quantitatively affecting $p_G^*(b,T)$ should be reflected to our empirical estimation results.[5] Proposition 1 establishes the main properties of the mapping $p_G^*(b,T)$.

Proposition 1: *Increasing patent length or narrowing patent scope increases the probability of a successful PIV patent challenge.*

Proposition 1 reproduces the standard results arising from the models of patent policy with costly imitation like ours: Longer patent duration makes waiting for patent expiration less attractive and hence encourages patent challenging, whereas broader patent scope discourages patent challenging by increasing its costs.

In the first stage the originator firm chooses p_B. The private value of an approved new drug is

$$V^P(T, p_G^*(b,T)) = \int_0^T e^{-rt} \left[\left(1 - p_G^*\left(b,T\right)\right) \widetilde{\pi}_1 + p_G^*\left(b,T\right) \widetilde{\pi}_2 \right] dt + \int_T^\infty e^{-rt} \widetilde{\pi}_2 dt, \tag{8}$$

in which $p_G^*\left(b,T\right)$ is determined by equation (7). The first integral on the right-hand side of equation (8) depicts the originator's profits when its new drug patent is in force. The originator will retain market exclusivity if the generic's patent challenge fails (the first term in the square brackets) but will encounter competition if the challenge succeeds (the second term in the square-brackets). The second integral expresses the originator's profits after the patent expiration.

The originator firm's problem can be expressed as

$$\max_{p_B \in [0,1]} \Pi_B = p_B V^P(T, p_G^*(b,T)) - C_B(p_B),$$

in which $V^P(T, p_G^*(b,T))$ is given by equation (8) and the last term captures the costs of developing

[5] Even equation (7) in itself is fairly resilient to some changes in model details. For example, assuming that a generic's profit flow is smaller when $t \geq T$ than when $t < T$ makes both a PIV patent challenge and waiting for patent expiration less lucrative, and these effects tend to cancel out each other: Consider, e.g., free entry after patent expiration which drives profits to zero. Then the generic's problem $\max_{p_G \in [0,1]} \Pi_G = p_G \int_0^T e^{-rt} \widetilde{\pi}_2 dt - C_G(p_G, b)$ leads to the same first-order condition (7).

a new drug. The first-order condition for this problem is

$$V^P(T, p_G^*(b,T)) - \frac{\partial C_B(p_B^*)}{\partial p_B} = 0. \tag{9}$$

Equations (7) and (9) determine the unique subgame perfect equilibrium (p_B^*, p_G^*) in which new drug market arises ($y_B(p_B) = 1$).

To facilitate the analysis of patent policy, we define

$$\phi(p_G) := \epsilon_p(p_G) - \frac{p_G}{1 - p_G}, \tag{10}$$

in which $\epsilon_p(p_G) > 0$ is defined by equation (4). Then, we have the following result:

Proposition 2: *Broader patent scope increases incentives to develop new drugs. Increasing (decreasing) patent length increases incentives to develop new drugs if* $\phi(p_G^*) > 0$ ($\phi(p_G^*) < 0$).

Propositions 1 and 2 suggest, as is intuitive, that the sign of $\partial p_B^*/\partial b$ is the reverse of the sign of $\partial p_G^*/\partial b$: broader patent scope makes successful patent challenging more difficult which in turn enhances incentives to develop new drugs.

In contrast, an increase in patent length has both a direct and an indirect effect on incentives to develop new drugs. The direct effect is positive: given the generic's incentive for patent challenging, the originator's market exclusivity lasts longer in expectation. The indirect effect via p_G^* is negative: longer patent duration enhances incentives for patent challenging. Hence, an increase in patent length can have a positive or a negative effect on incentives to develop new drugs depending on whether the direct or indirect effect dominates, which in turn depends on the sign of $\phi(p_G^*)$.[6] However, the case $\phi(p_G^*) < 0$ is not only counterintuitive but also unlikely: If $\phi(p_G^*) < 0$, the originator would have an incentive to shorten the effective length of its patent. Moreover, as shown in Section 7, our empirical results suggest $\phi(p_G^*) > 0$.

[6] With mild additional assumptions ($\partial\phi/\partial p_G < 0$ and $\lim_{T\to\infty} \phi(p_G^*(T)) < 0$), these direct and indirect effects of patent length would create the inverted-U relationship between patent length and innovation incentives (with the peak at some T' solving $\phi(p_G^*(T')) = 0$), which has been discovered in the literature (see, e.g., Horowitz and Lai, 1996; Takalo, 1998; Gallini, 2002; Quian, 2007).

6.3 Welfare Analysis

Denote welfare flow from a new drug by $\widetilde{w}_N \in [0, \infty)$ when $N \in \{0, 1, 2\}$ drugs compete in the market. As usual, $\widetilde{w}_2 > \widetilde{w}_1 > \widetilde{w}_0 = 0$. Analogous to the private value of a new drug given by equation (8), we can write the social value of an existing new drug as

$$\begin{aligned} V^S(b,T) &= \int_0^T e^{-rt} \left[\left(1 - p_G^*(b,T)\right) \widetilde{w}_1 + p_G^*(b,T)\, \widetilde{w}_2 \right] dt \\ &\quad + \int_T^\infty e^{-rt} \widetilde{w}_2 dt - C_G(p_G^*(b,T), b), \end{aligned} \tag{11}$$

in which $p_G^*(b,T)$ is identified by equation (7). The first and second integral on the right-hand side of equation (11) give welfare from a new drug before and after its patent expires, respectively. The last term captures the generic's patent challenging cost.

We formulate the patent policy problem as follows:

$$\max_{b \in [0,\infty),\, T \in [0,\infty)} V^S(b,T) \tag{12}$$

subject to

$$p_B^*(b,T) = \overline{p}_B.$$

This often-used formulation (see, e.g., Gallini, 1992; Dubois et al., 2022) seeks to characterize an optimal patent policy reform that minimizes welfare distortions while keeping incentives to innovate unchanged at a desired level $\overline{p}_B$. For our purpose, the formulation is especially useful as we are aiming at matching its results with the local effects identified in Section 4 and 5, which do not necessarily hold globally.

Recalling that $\epsilon_b(p_G) > 0$, the solution to the problem (12) can be expressed as follows:

Proposition 3: *i) If* $\epsilon_b(p_G^*) > \phi(p_G^*) > 0$, *reducing patent length and increasing scope is efficient; ii) If* $\phi(p_G^*) < 0$, *reducing both patent length and scope is efficient; iii) If* $\phi(p_G^*) > \epsilon_b(p_G^*)$, *increasing patent length and reducing scope is efficient.*

If $\epsilon_b(p_G^*) > \phi(p_G^*)$ (parts i) and ii) of Proposition 3), long-lived patents are inefficient irrespective

of the sign of $\phi(p_G^*)$. This result is familiar from the models of costly imitation. In our context, long patent duration is ineffective in promoting new drug development, since it also increases incentives for patent challenging.

When shorter patent length is desirable, the sign of $\phi(p_G^*)$ suggests optimal changes to patent scope. If $\phi(p_G^*) > 0$ (part i)), shorter patent length has an adverse effect on incentives to develop new drugs, which should be compensated for originators by making new drug patents broader. If $\phi(p_G^*) < 0$ (part ii)), shorter patent length has a *positive* effect on incentives to develop new drugs, and patents can be made narrower without jeopardizing new drug development. Thus the sign of $\phi(p_G^*)$ also determines whether patent length and scope are "substitutable" or "complementary" policy tools with regard to new drug development (see Belleflamme and Peitz, 2015, for the terminology).

Nonetheless, even in the presence of costly imitation, narrower and longer-lived patents could be efficient as suggested by part iii) of Proposition 3. If $\epsilon_b(p_G^*)$ is small, broader patent scope has only a relatively small impact on patent challenging but a relatively large impact on its costs. Thus, distortions caused by broader patents can even be larger than distortions caused by longer patents. Coming up with a sensible cost function generating this outcome is, however, difficult.

Example 1. Assume that the generic firm has a constant elasticity cost function

$$C_G(p_G, b) = \frac{c(b)p_G^{\eta_G}}{\eta_G}, \tag{13}$$

in which $\eta_G > 1$. Assume that the constant $c(b) \geq 0$. scaling the cost function is increasing patent scope, $\partial c/\partial b > 0$. Then, Assumption 1 and, consequently, Propositions 1 and 2 hold.

Equations (4) and (13) imply $\epsilon_p = \eta_G - 1$ and, hence, we can rewrite equation (10) as

$$\phi(p_G) = \eta_G - \frac{1}{1 - p_G}. \tag{14}$$

Applying equation (13) in the definition (5) yields $\epsilon_b = \eta_G$. As a result, $\epsilon_b > \phi(p_G)$. Thus, by Proposition 3, shortening patent length would be desirable irrespective of the sign of $\phi(p_G)$. However, to complete the policy analysis, we should determine whether $\phi(p_G)$ is positive or negative.

7 Implications for Pharmaceutical Patent Policy

Our empirical evidence suggests that longer effective patent length increases the probability of successful PIV patent challenge, as predicted by Proposition 1 of our theoretical model. This effect weakens the efficiency of patent length as a policy tool to promote new drug development. Our evidence also supports the prediction of Proposition 1 about the negative effect of broader patent scope on PIV patent challenges. According to the original Gallini's (1992) model, these results alone would be enough make shorter but broader pharmaceutical patents optimal. Our model using more flexible innovation and imitation cost functions suggests that this conclusion is plausible but not foregone. To push our policy analysis further we resort, in the spirit of the sufficient statistics approach, to the elasticity formulas developed Section 6.

Proposition 3 suggests that to evaluate the welfare effects of pharmaceutical patent policy changes, we only need to compare $\phi(p_G^*) := \epsilon_p(p_G^*) - p_G^*/(1-p_G^*)$ to zero and $\epsilon_b(p_G^*)$. The sign of $\phi(p_G^*)$ tells us whether incentives to develop new drugs are increasing or decreasing in patent length, and also whether patent length and scope are substitutable or complementary policy tools. In turn, the sign of $\phi(p_G^*) - \epsilon_b(p_G^*)$ tells us whether increasing patent length causes larger or smaller distortions than increasing patent scope.

In our data the average probability of PIV patent challenge is 0.17 (see Table 1), which directly provides an estimate of p_G^*. We develop two alternative approaches to recover $\epsilon_p(p_G^*)$, the elasticity of the generic firms' marginal cost function (see equation (4)): one based on the estimated elasticity of PIV patent challenge with respect to effective patent length, $\xi_T(p_G) := (\partial p_G/p_G)/(\partial T/T)$, and another on the corresponding elasticity with respect to patent scope, $\xi_b(p_G) := (\partial p_G/p_G)/(\partial b/b)$.

Recovering $\epsilon_p(p_G^*)$ via Patent Length Estimates. In Appendix 2, we show that $\epsilon_p(p_G)$ can be written as

$$\epsilon_p(p_G^*) = \frac{e^{-rT} rT}{(1-e^{-rT})\xi_T(p_G^*)}. \tag{15}$$

Hence, to calculate $\epsilon_p(p_G^*)$, we need values for r, T and $\xi_T(p_G^*)$. We set $T = 12.59$ corresponding the average effective patent length in our data (see Table 1), and $r = 0.03$ following the value used by Schankerman and Schuett (2022). (We have also experimented with $r = 0.05$, which is used e.g., by Judd et al. (2012) in a similar context.) To obtain a value for $\xi_T(p_G^*)$, we compare estimates of

β_2 of equation (2) in columns (1)–(5) of Table 3 to the average probability of PIV patent challenge, 0.17. This comparison suggests that TRIPS reduced the rate of PIV patent challenge by $47-65$ percent. These figures, and the estimated effect of TRIPS on effective patent length (-17 percent) reported in column (6) of Table 3, suggest that ξ_T could be between 47/17 and 65/17. We use the mean value of this range and set $\xi_T = 3.29$. Inserting $\xi_T = 3.29$, $r = 0.03$, and $T = 12.59$ into equation (15) gives $\epsilon_p \approx 0.25$.

Recovering $\epsilon_p(p_G^*)$ via Patent Scope Estimates. Our level-log regressions imply that an estimate of $\xi_b(p_G^*)$ can directly be obtained by dividing an estimated coefficient of a log scope measure by the average probability of PIV patent challenge, 0.17. The statistically significant coefficient estimates of the IV regressions (columns (1)–(3) of Panel A of Table 5) suggest $\xi_b \in [-0.65, -1, 41]$. Using the average value of this range, we set $\xi_b = -1.03$. To link $\xi_b(p_G^*)$ with $\epsilon_p(p_G^*)$, we assume that the generic's cost function has the constant elasticity form of equation (13), and further stipulate that $c(b) = cb$. Using this functional form, we show in Appendix 2 that $\epsilon_p = -1/\xi_b$. Thus, setting $\xi_b = -1.03$ gives $\epsilon_p \approx 0.97$.

Effect of Patent Length on Incentives to Develop New Drug. We have two alternative ways to calculate $\epsilon_p(p_G^*)$: using $\xi_T(p_G^*)$ yields $\epsilon_p = 0.25$, whereas using $\xi_b(p_G^*)$ yields $\epsilon_p = 0.97$. Assuming that these calculations mean that $\epsilon_p \in [0.25, 0.97]$, and setting $p_G^* = 0.17$ yields $\phi(p_G^*) \in [0.05, 0, 77]$. The marginal cost of PIV patent challenge thus appears to be sufficiently inelastic to keep innovation incentives in the pharmaceutical industry increasing in patent length, which also implies that patent length and scope are substitutes with regard to those innovation incentives.

Evaluating Patent Policy. To complete the evaluation of patent policy, we should assess whether $\phi(p_G^*)$ is larger or smaller than $\epsilon_b(p_G^*)$, the elasticity of the effect of patent scope on patent challenging costs (equation (5)). We estimate $\epsilon_b(p_G^*)$ similarly to $\xi_b(p_G^*)$. In Appendix 2, we show that $\epsilon_b = 1 - 1/\xi_b$. Using $\xi_b = -1.03$ gives $\epsilon_b \approx 1.97$. Thus, $\epsilon_b(p_G^*)$ appears to be clearly larger than our estimates of $\phi(p_G^*)$. In sum, our results suggest $\epsilon_b(p_G^*) > \phi(p_G^*) > 0$, confirming the intuitive argument for the desirability of shorter but broader pharmaceutical patents.

Robustness. Our empirical results provide estimates of the elasticities of PIV patent challenge with respect to effective patent length and scope, $\xi_T(p_G^*)$ and $\xi_b(p_G^*)$. The sufficient statistics approach we use to turn these estimated elasticities into policy prescriptions may be sensitive to some assumptions. While all our estimates of $\xi_b(p_G^*)$ suggest that $\phi(p_G^*)$ is firmly positive but below

$\epsilon_b(p_G^*)$, we use here a functional form assumption which we cannot test.

If we instead use the estimates of $\xi_T(p_G^*)$ in our policy analysis, we need no functional form assumption but $\phi(p_G^*)$ turns out to be close to zero (while positive). As a result, our policy conclusion is sensitive to some estimation results (and to the chosen value for r): If we use the estimated effects of AIPA on effective patent length instead of those of TRIPS, we will obtain somewhat higher values of $\xi_T(p_G^*)$ and hence lower values of $\phi(p_G^*)$. Using our largest estimate of the effect of AIPA from Table 3 makes $\xi_T(p_G^*)$ sufficiently high to render $\phi(p_G^*)$ slightly negative.

Nonetheless, we consider the case $\phi(p_G^*) < 0$ unlikely. For example, if $\phi(p_G^*) < 0$, originators would have an incentive to find means (e.g., licensing or other contractual solutions) to shorten the effective lengths of their patents. In line with this argument, we show in Online Appendix A.3 how originators shy away from continuing patents with long grant lags after TRIPS and argue that such a behavior is consistent with $\phi(p_G^*) > 0$ but not with $\phi(p_G^*) < 0$. Moreover, the estimated effects of AIPA are somewhat ambiguous because of the earlier disclosure of patent applications in the post-AIPA period (see Online Appendix A.2). We conclude that $\phi(p_G^*)$ is likely to be smaller than $\epsilon_b(p_G^*)$, but hardly negative. Finally, even if $\phi(p_G^*)$ were negative, our conclusion of the desirability of shorter pharmaceutical patent lives would remain unchanged.

8 Conclusion

We estimate the effects of patent policy in the U.S. pharmaceutical industry by using data on new drug patents and their challenges by generic entrants. Using DiD regressions exploiting patent term reforms and IV regressions exploiting patent examiner leniency differences, we find that successful PIV patent challenges increase with the effective length of new drug patents and decrease with their scope. According to the costly imitation theory, these results create an argument for shortening patent length, while making their scope broader to restore incentives to develop new drugs. We confirm this policy prescription by matching our data moments with a patent policy model with costly imitation.

Our empirical results should as such be valuable for the design of pharmaceutical patent policy. For example, our DiD regression results provide insights into the debate on pharmaceutical patent length (see, e.g., Budish et al., 2015; Dubois et al., 2022) by pointing out that longer patent terms

can have adverse effects besides reduced consumer surplus. Our IV regression results in turn show that Markush claiming can potentially be used to ward off PIV patent challenges. A future work should build on Wagner et al. (2022) to study the effects of Markush claiming in more detail. Wider applications of the doctrine of equivalents or means-plus-function clauses could constitute another way to broaden the claim scope of drug patents to restrict PIV patent challenges, as proposed by Tang (2013).

When we turn our estimated elasticities into precise policy recommendations, we resort to some strong assumptions. For example, we abstract away from cumulative innovation – see Galasso and Schankerman (2015, 2018) and Sampat and Williams (2019) for the effects of the existence of patents on follow-on innovation. Nonetheless, characterizing the patent policy tradeoff at the heart of the Hatch-Waxman Act in the classic stand-alone innovation framework should constitute the first step towards understanding of the effects drug patent policy also when innovation is cumulative. Moreover, long patent duration tends to create additional harmful effects in the presence of cumulative innovation besides those captured by our work (Boldrin and Levine, 2004; Bessen and Maskin, 2009). Our model also ignores pay-for-delay and other patent litigation settlements (e.g., Hovenkamp and Lemus, 2018; Jacobo-Rubio et al., 2020) However, those settlements in so far they delay generic entry should make successful PIV patent challenges less beneficial for the society and the case for shorter but broader pharmaceutical patents stronger.

References

Abrams, D. S. (2009). Did TRIPS spur innovation? An empirical analysis of patent duration and incentives to innovate. *University of Pennsylvania Law Review*, 157(6):1613–1647.

Akcigit, U. and Ates, S. T. (2023). What happened to U.S. business dynamism? *Journal of Political Economy*, forthcoming.

Belleflamme, P. and Peitz, M. (2015). *Industrial Organization: Markets and Strategies*. Cambridge University Press: Cambridge, UK, 2nd edition.

Bessen, J. and Maskin, E. (2009). Sequential innovation, patents, and imitation. *RAND Journal of Economics*, 40(4):611–635.

Boldrin, M. and Levine, D. K. (2004). The case against intellectual monopoly. *International Economic Review*, 45(2):327–350.

Branstetter, L., Chatterjee, C., and Higgins, M. J. (2016). Regulation and welfare: Evidence from Paragraph IV generic entry in the pharmaceutical industry. *RAND Journal of Economics*, 47(4):857–890.

Budish, E., Roin, B. N., and Williams, H. (2015). Do firms underinvest in long-term research? evidence from cancer clinical trials. *American Economic Review*, 105(7):2044–85.

Budish, E., Roin, B. N., and Williams, H. (2016). Patents and research investments: Assessing the empirical evidence. *American Economic Review: Papers & Proceedings*, 106(5):183–187.

Chetty, R. (2009). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annual Review of Economics*, 1(1):451–488.

Ciliberto, Frederico, M. C. and Elie, T. (2021). Market structure and competition in airline markets. *Journal of Political Economy*, 129(11):2995–3038.

Cockburn, I., Kortum, S., and Stern, S. (2003). Are all patent examiners equal? Examiners, patent characteristics, and litigation outcomes. In Cohen, W. and Merrill, S., editors, *Patents in the Knowledge-Based Economy*, pages 19–53. Washington, DC: The National Academies Press.

Denicolò, V. (1996). Patent races and optimal patent breadth and length. *Journal of Industrial Economics*, 44(3):249–265.

Denicolò, V. (2007). Do patents over-compensate innovators? *Economic Policy*, 22(52):679–729.

Dubois, P., Moisson, P.-H., and Tirole, J. (2022). The economics of transferable patent extensions. TSE Working Paper, n. 22-1377.

Farre-Mensa, J., Hegde, D., and Ljungqvist, A. (2020). What is a patent worth? Evidence from the U.S. patent "lottery". *Journal of Finance*, 75(2):639–682.

Feng, J. and Jaravel, X. (2020). Crafting intellectual property rights: Implications for patent assertion entities, litigation, and innovation. *American Economic Journal: Applied Economics*, 12(1):140–181.

Freilich, J. (2015). The uniformed topology of patent scope. *Stanford Technology Law Review*, 19:150.

Gaessler, F. and Wagner, S. (2022). Patents, data exclusivity, and the development of new drugs. *Review of Economics and Statistics*, 104(3):571–586.

Galasso, A. and Schankerman, M. (2015). Patents and cumulative innovation: Causal evidence from the courts. *Quarterly Journal of Economics*, 130(1):317–369.

Galasso, A. and Schankerman, M. (2018). Patent rights, innovation, and firm exit. *RAND Journal of Economics*, 49(1):64–86.

Gallini, N. T. (1992). Patent policy and costly imitation. *RAND Journal of Economics*, 23(1):52–63.

Gallini, N. T. (2002). The economics of patents: Lessons from recent U.S. patent reform. *Journal of Economic Perspectives*, 16(2):131–154.

Gambardella, A., Harhoff, D., and Verspagen, B. (2008). The value of European patents. *European Management Review*, 5(2):69–84.

Gilchrist, D. S. (2016). Patents as a spur to subsequent innovation: Evidence from pharmaceuticals. *American Economic Journal: Applied Economics*, 8(4):189–221.

Grabowski, H. G. and Vernon, J. (2000). Effective patent life in pharmaceuticals. *International Journal of Technology Management*, 19(1):98–120.

Hall, B., Helmers, C., Rogers, M., Sena, VaniaMarco, A. C., and Miller, R. (2014). The choice

between formal and informal intellectual property: A review. *Journal of Economic Literature*, 52(2):375–423.

Harhoff, D. and Reitzig, M. (2004). Determinants of opposition against EPO patent grants - the case of biotechnology and pharmaceuticals. *International Journal of Industrial Organization*, 22(4):443–480.

Hemphill, C. S. and Sampat, B. N. (2011). When do generics challenge drug patents? *Journal of Empirical Legal Studies*, 8(4):613–649.

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

Henry, E. and Ponce, C. J. (2011). Waiting to imitate: on the dynamic pricing of knowledge. *Journal of Political Economy*, 119(5):959–981.

Higgins, M. J. and Graham, S. J. H. (2009). Balancing innovation and access: Patent challenges tip the scales. *Science*, 326(5951):370–371.

Horowitz, A. W. and Lai, E. (1996). Patent length and the rate of innovation. *International Economic Review*, 37(4):785–801.

Hovenkamp, E. and Lemus, J. (2018). Delayed entry settlements at the patent office. *International Review of Law and Economics*, 54:30–38.

Jacobo-Rubio, R., L., T. J., and Williams, J. W. (2020). The distribution of surplus in the US pharmaceutical industry: Evidence from Paragraph iv patent-litigation decisions. *Journal of Law and Economics*, 63(2):203–238.

Johnson, D. K. N. and Popp, D. (2003). Forced out of the closet: The impact of the American Inventors Protection Act on the timing of patent disclosure. *RAND Journal of Economics*, 34(1):96–112.

Judd, K. L., Schmedders, K., and Yeltekin, S. (2012). Optimal rules for patent races. *International Economic Review*, 53(1):23–52.

Kleven, H. J. (2021). Sufficient statistics revisited. *Annual Review of Economics*, 13(1):515–538.

Kuhn, J. M. and Thompson, N. C. (2019). How to measure and draw causal inferences with patent scope. *International Journal of the Economics of Business*, 26(1):5–38.

Kyle, M. K. and McGahan, A. M. (2012). Investments in pharmaceuticals before and after TRIPS. *Review of Economics and Statistics*, 94(4):1157–1172.

Lanjouw, J. O. and Schankerman, M. (2001). Characteristics of patent litigation: A window on competition. *RAND Journal of Economics*, 32(1):129–151.

Lanjouw, J. O. and Schankerman, M. (2004). Patent quality and research productivity: Measuring innovation with multiple indicators. *Economic Journal*, 114(495):441–465.

Lemley, M. A. and Moore, K. A. (2004). Ending the abuse of patent continuations. *Boston University Law Review*, 84(63):63–123.

Lemley, M. A. and Sampat, B. (2012). Examiner characteristics and patent office outcomes. *Review of Economics and Statistics*, 94(3):817–827.

Lemley, M. A. and Shapiro, C. (2005). Probabilistic patents. *Journal of Economic Perspectives*,

19(2):75–98.

Lerner, J. (2009). The empirical impact of intellectual property rights on innovation: Puzzles and clues. *American Economic Review*, 99(2):343–348.

Lerner, J., Seru, A., Short, N., and Sun, Y. (2021). Financial innovation in the 21st century: Evidence from u.s. patents. Working Paper 28980, National Bureau of Economic Research.

Marco, A. C., Sarnoff, J. D., and deGrazia, C. A. (2019). Patent claims and patent scope. *Research Policy*, 48(9):103790.

Maurer, S. M. and Scotchmer, S. (2002). The independent invention defence in intellectual property. *Economica*, 69(276):535–547.

Merges, R. P. and Nelson, R. R. (1990). On the complex economics of patent scope. *Columbia Law Review*, 90(4):839–916.

Moser, P. (2013). Patents and innovation: Evidence from economic history. *Journal of Economic Perspectives*, 27(1):23–44.

Newham, M., Seldeslachts, J., and Banal Estañol, A. (2018). Common ownership and market entry: Evidence from pharmaceutical industry. DIW Berlin Discussion Paper No. 1738.

Nordhaus, W. D. (1969). *Invention Growth, and Welfare: A Theoretical Treatment of Technological Change*. The MIT Press: Cambridge, MA.

Nordhaus, W. D. (1972). The optimum life of a patent: Reply. *American Economic Review*, 62(3):428–431.

Quian, Y. (2007). Do national patent laws stimulate domestic innovation in a global patenting environment? A cross-country analysis of pharmaceutical patent protection, 1978–2002. *Review of Economics and Statistics*, 89(3):436–453.

Righi, C. and Simcoe, T. (2019). Patent examiner specialization. *Research Policy*, 48(1):137–148.

Sakakibara, M. and Branstetter, L. (2001). Do stronger patents induce more innovation? Evidence from the 1988 Japanese patent law reforms. *RAND Journal of Economics*, 32(1):77–100.

Sampat, B. N. (2018). A survey of empirical evidence on patents and innovation. NBER Working Paper No. 25383.

Sampat, B. N. and Williams, H. L. (2019). How do patents affect follow-on innovation? Evidence from the human genome. *American Economic Review*, 109(1):203–236.

Schankerman, M. and Schuett, F. (2022). Patent screening, innovation, and welfare. *Review of Economic Studies*, 89(4):2101–2148.

Shapiro, C. (2003). Antitrust limits to patent settlements. *RAND Journal of Economics*, 34(2):391–411.

Takalo, T. (1998). Innovation and imitation under imperfect patent protection. *Journal of Economics*, 67(3):229–241.

Tang, W. L. (2013). Revitalizing the patent system to incentivize pharmaceutical innovation: The potential claims with means-plus-function clauses. *Duke Law Journal*, 32(1069):1069–1108.

Voet, M. A. (2016). *The Generic Challenge: Understanding Patents, FDA and Pharmaceutical Life-Cycle Management*. BrownWalker Press: Boca Raton, FL, 5th edition.

Wagner, S., Sternitzke, C., and Sascha, W. (2022). Mapping Markush. *Research Policy*, 51:104597.
Wright, D. J. (1999). Optimal patent breadth and length with costly imitation. *International Journal of Industrial Organization*, 17(3):419–436.

Appendix

A.1 Proofs

Proof of Proposition 1: Applying the implicit function theorem to equation (7) yields

$$\frac{\partial p_G^*}{\partial T} = \frac{re^{-rT}\pi_2}{\partial^2 C_G/\partial p_G^2} > 0 \tag{16}$$

and

$$\frac{\partial p_G^*}{\partial b} = -\frac{\partial^2 C_G/\partial p_G \partial b}{\partial^2 C_G/\partial p_G^2} < 0, \tag{17}$$

in which the inequalities follow from $\partial^2 C_G/\partial p_G^2 > 0$ and Assumption 1. □

Proof of Proposition 2: Using the implicit function theorem in equation (9) together with $\partial^2 C_B/\partial p_B^2 > 0$ imply that the signs of $\partial p_B^*/\partial b$ and $\partial p_B^*/\partial T$ are given by the signs of $\partial V^P/\partial b$ and $\partial V^P/\partial T$, respectively. Then, differentiating equation (8) with respect to b and using the definition $\pi_N := \widetilde{\pi}_N/r$ yield

$$\frac{\partial V^P}{\partial b} = -\left(1 - e^{-rT}\right)(\pi_1 - \pi_2)\frac{\partial p_G^*}{\partial b}, \tag{18}$$

in which $\partial p_G^*/\partial b < 0$ by Proposition 1. The claim concerning patent scope follows.

Similarly, differentiating equation (8) with respect to T gives

$$\frac{\partial V^P}{\partial T} = (\pi_1 - \pi_2)\left[re^{-rT}(1 - p_G^*) - \frac{\partial p_G^*}{\partial T}(1 - e^{-rT})\right]. \tag{19}$$

After using equations (4), (7), and (16), we can rewrite this equation as

$$\frac{\partial V^P}{\partial T} = \frac{re^{-rT}}{\epsilon_p}(\pi_1 - \pi_2)\phi(p_G^*)(1 - p_G^*), \tag{20}$$

in which $\phi(p_G)$ is defined by equation (10). Thus the sign of $\partial V^P/\partial T$ is given by the sign of $\phi(p_G^*)$. The claim concerning patent length follows. □

Proof of Proposition 3: Using equation (9), which determines $p_B^*(b, T)$, in applying the implicit function theorem to equation (12) yields

$$\frac{\partial T}{\partial b} = -\frac{\partial V^P/\partial b}{\partial V^P/\partial T}. \tag{21}$$

We may now re-express the planner's problem as $\max_{b\in[0,\infty)} V^S(b, T(b))$. Differentiating $V^S(b, T(b))$

with respect to b gives

$$\frac{dV^S}{db} = \frac{\partial V^S}{\partial b} + \frac{\partial V^S}{\partial T}\frac{\partial T}{\partial b}. \tag{22}$$

After substituting equations (18) and (20) for equation (21), we get

$$\frac{\partial T}{\partial b} = \frac{(1-e^{-rT})\epsilon_p}{re^{-rT}(1-p_G^*)\phi(p_G^*)}\frac{\partial p_G^*}{\partial b}. \tag{23}$$

Let $w_N := \widetilde{w}_N/r$. Then, differentiating equation (11) with respect to b gives

$$\frac{\partial V^S}{\partial b} = \left[\left(1-e^{-rT}\right)(w_2-w_1) - \frac{\partial C_G}{\partial p_G}\right]\frac{\partial p_G^*}{\partial b} - \frac{\partial C_G}{\partial b}. \tag{24}$$

Similarly, for T we get

$$\frac{\partial V^S}{\partial T} = -re^{-rT}(1-p_G^*)(w_2-w_1) + \left[\left(1-e^{-rT}\right)(w_2-w_1) - \frac{\partial C_G}{\partial p_G}\right]\frac{\partial p_G^*}{\partial T},$$

which can be rewritten after some algebra by using equations (4), (7), (10), and (16) as

$$\frac{\partial V^S}{\partial T} = \frac{-re^{-rT}}{\epsilon_p}\left[(w_2-w_1)\,\phi(p_G^*)(1-p_G^*) + p_G^*\pi_2\right]. \tag{25}$$

After using equations (23)–(25), and some algebra, equation (22) can be written as

$$\frac{dV^S}{db} = -\frac{\partial p_G^*}{\partial b}\left[\frac{\partial C_G}{\partial p_G} + \frac{\left(1-e^{-rT}\right)\pi_2 p_G^*}{\phi(p_G^*)(1-p_G^*)}\right] - \frac{\partial C_G}{\partial b}.$$

By using equations (4), (5), (7), (10), and (17), this expression can be further rewritten as

$$\frac{dV^S}{db} = \frac{\partial C_G}{\partial b}\left[\frac{\epsilon_b(p_G^*)}{\phi(p_G^*)} - 1\right]. \tag{26}$$

Since $\partial C_G/\partial b > 0$ by assumption, the sign of dV^S/db is given by the sign of the term in the square brackets of equation (26).

An optimal patent policy reform is characterized by the signs of equations (23) and (26). Equation (26) tells us the optimal direction of patent scope and equation (23) tells us the direction where patent length needs to be adjusted to compensate a change in patent scope. As to the sign of equation (23), Proposition 1 implies that $\partial p_G^*/\partial b < 0$. As a result, the sign of $\partial T/\partial b$ is given by the sign of $-\phi(p_G^*)$. As to the sign of equation (26), recall first the implication of Assumption 1 that $\epsilon_b(p_G^*) > 0$. Thus, if $\phi(p_G^*) < 0$, the term in the square brackets of equation (26) is definitely negative and $dV^S/db < 0$. If $\phi(p_G^*) > 0$, then the first term in the square brackets of equation (26) is positive. Then the term in the square brackets is negative if $\epsilon_b(p_G^*) < \phi(p_G^*)$, and positive if $\epsilon_b(p_G^*) > \phi(p_G^*)$.

To sum up: i) If $\epsilon_b(p_G^*) > \phi(p_G^*) > 0$, $\partial T/\partial b < 0$ and $dV^S/db > 0$. Increasing patent scope and

reducing length is efficient; ii) If $\phi(p_G^*) < 0$, $\partial T/\partial b > 0$ and $dV^S/db < 0$. Reducing both patent scope and length is efficient; iii) If $\phi(p_G^*) > \epsilon_b(p_G^*)$, $\partial T/\partial b < 0$ and $dV^S/db < 0$. Reducing patent scope and increasing length is efficient. □

A.2 Elasticity Calculations for Section 7

Recall from Section 7 the definitions

$$\xi_j(p_G) := \frac{\partial p_G}{\partial j}\frac{j}{p_G},\ j \in \{b, T\}\,. \tag{27}$$

We now prove the claims of Section 7 concerning the relation of $\xi_j(p_G^*)$ to $\epsilon_p(p_G^*)$ of equation (4), and to $\epsilon_b(p_G^*)$ of equation (5).

Let us start with the relationship between $\xi_T(p_G^*)$ and $\epsilon_p(p_G^*)$. Using the first-order condition (7) to eliminate π_2 from equation (16) allows us to rewrite equation (16) as

$$\frac{\partial p_G^*}{\partial T} = \frac{re^{-rT}}{(1-e^{-rT})}\frac{\partial C_G/\partial p_G}{\partial^2 C_G/\partial p_G^2}.$$

Multiplying both sides of this equation by T/p_G^* gives

$$\frac{\partial p_G^*}{\partial T}\frac{T}{p_G^*} = \frac{e^{-rT}rT}{(1-e^{-rT})}\frac{\partial C_G/\partial p_G}{p_G^*\partial^2 C_G/\partial p_G^2}.$$

By using the definitions of the elasticities (4) and (27) this equation can be rewritten as

$$\xi_T(p_G^*) = \frac{e^{-rT}rT}{(1-e^{-rT})\epsilon_p(p_G^*)},$$

i.e.,

$$\epsilon_p(p_G^*) = \frac{e^{-rT}rT}{(1-e^{-rT})\xi^T(p_G^*)}.$$

This equation equals equation (15) of Section 7.

To link $\xi_b(p_G^*)$ with $\epsilon_p(p_G^*)$, we assume that the generic firm's cost function is specified by equation (13), and further stipulate that $c(b) = cb$. Then, equation (13) can be expressed as $C_G(p_G, b) = cbp_G^{\eta_G}/\eta_G$. Using this cost function in equation (7) yields after some algebra

$$p_G^*(b, T) = \left[\frac{\left(1-e^{-rT}\right)\pi_2}{cb}\right]^{\frac{1}{\eta_G - 1}}. \tag{28}$$

Differentiating equation (28) with respect to b gives

$$\frac{\partial p_G^*}{\partial b} = -\frac{1}{b}\left(\frac{1}{\eta_G - 1}\right)\left[\frac{\left(1-e^{-rT}\right)\pi_2}{cb}\right]^{\frac{1}{\eta_G - 1}}.$$

Using equation (28) and multiplying both sides of the equation above by b/p^*_G gives

$$\frac{\partial p^*_G}{db}\frac{b}{p^*_G} = \frac{1}{1-\eta_G},$$

which may be rewritten by using equation (27) as

$$\eta_G = 1 - \frac{1}{\xi_b}. \tag{29}$$

Using equation (13) in equation (4) yields $\epsilon_p = \eta_G - 1$ which, according to equation (29), is equivalent to $\epsilon_p = -1/\xi_b$, as claimed in Section 7. Our claim in Section 7 that $\epsilon_b = 1 - 1/\xi_b$ directly follows from equation (29) and Example 1 of Section 6 in which we show that $\epsilon_b = \eta_G$.

Online Appendix

Online A.1 Data Construction

In this appendix we describe the details of our data sources and variable construction. We also address the issue of missing observations in the outcome variable.

Online A.1.1 Identifying FDA Approved New Drugs and Their Patents: Orange Book and Drugs@FDA

The main data for our empirical analysis comes from the FDA's publication Approved Drug Products with Therapeutic Equivalent Evaluations, commonly known as the Orange Book. We use the annual editions 21–33 (corresponding years 2001–2013) of the Orange Book. The editions 21–32 in electronic format were received from the FDA through a Freedom of Information Act (FOIA) request. The 33^{rd} edition was downloaded from the FDA web-pages. The Orange Book digital publications consist of product, patent, and exclusivity data files. The product file identifies each drug with an FDA application number and, among others, lists information on active ingredient, dosage form, strength, producer, trade name and approval date. The patent file lists the patents included in each drug application. If a drug is covered by the FDA exclusivity the type of exclusivity and its latest expiration date are listed in the exclusivity file. Using the drug application numbers we first combine the Orange Book data files into one, and then link it with the early 2014 release of the Drugs@FDA database, Application and Product Tabs.[7]

The unit of observation in our analysis is a new drug patent. By definition, only new medicines, New Drug Applications (NDAs), may be listed in the Orange Book with corresponding patent information. Each patent thus protects an NDA. To account for the characteristics of NDAs, we aggregate drug application level data into the patent level observations. For example, if a patent protects a drug in a tablet form, we assign value one to the patent level indicator variable Tablet. The indicator variables for two other common drug forms, Capsule and Injectable, are calculated similarly. Some patents protect several active ingredients listed in the Orange Book. So we identify the first FDA-approved active ingredient protected by a given patent. We also identify the first FDA approval date of an NDA protected by a patent.

New chemical and orphan drug exclusivities are marked in the Orange Book by the codes "NCE" and "ODE", respectively. Using these codes, we construct the indicator variables New chemical exclusivity and Orphan drug exclusivity measuring whether or not a patent protects a drug with new chemical or orphan drug exclusivity.

A drug receives pediatric exclusivity if its producer has conducted clinical trials and proved efficacy and safety of its drug for children. In this case, additional six month of exclusive marketing rights are added to all existing patents and exclusivities covering the drug (21 U.S.C. § 355a(b)). A

[7]The latest releases of the Orange Book and Drugs@FDA are available, respectively, at `https://www.fda.gov/Drugs/InformationOnDrugs/ucm129662.htm` and `https://www.fda.gov/Drugs/InformationOnDrugs/ucm079750.htm` (accessed April 2, 2020).

patent listed for an application with pediatric exclusivity has is recorded twice in the Orange Book: first, with the original patent number and its corresponding original expiration date, and second, with the original patent number followed by a "*PED" mark and a new expiration date six months later than the original one. Based on "*PED" designations, we assign value one to the indicator variable Pediatric exclusivity if the patent protects a drug with pediatric exclusivity. Using the Orange Book, we also compute the latest FDA exclusivity year for each patent, and identify the patent expiration date.

We identify priority reviewed drugs from the Drugs@FDA database. Priority reviewed drugs are assessed by the FDA faster as they, if approved, would represent significant improvements over available therapy, and thus might be particularly valuable. During our data period, such drugs were marked by a letter "P" or "P*" in the Therapeutic Potential column of the Drugs@FDA Application Tab. Based on these codes, we construct the indicator variable Priority measuring whether or not a patent protects a drug that was priority reviewed by the FDA.

Online A.1.2 Identifying Successful Patent Challenges: Abbreviated New Drug Application Approval Letters and Orange Book

We measure successful PIV challenges to new drug patents. In a PIV patent challenge, a generic firm seeks to enter prior to the expiration of a new drug patent by filing an Abbreviated New Drug Application (ANDA) to the FDA containing a certification that the new drug patent is invalid or noninfringed by the generic drug (21 U.S.C. § 355(j)(2)(A)(vii)(IV)). To construct the outcome variable, we first obtain the list of the approved ANDAs containing PIV certifications from the FDA through a FOIA request. To link this list with patent numbers, we seek the FDA's letters approving these ANDAs. Each originator drug is usually protected by multiple patents, and these approval letters specify all originator drug patents that have successfully been challenged. Some of the letters are readily available from the FDA web-pages through the Drugs@FDA search engine. To collect more letters, we submit FOIA requests to the FDA. The data collection process has been slow, since the FDA only accepts a few FOIAs per month.

We read each approval letter and record patent numbers together with corresponding PIV certifications mentioned in the letters. Using patent numbers, we then link the data on PIV patent challenges from ANDA approval letters to the data on all new drug patents listed in the Orange Book. Some patents are challenged by multiple generics, so we calculate the earliest PIV ANDA approval date for each patent from the Orange Book files. We then construct our outcome variable, PIV patent challenge, as an indicator which equals one if a new drug patent has successfully been challenged via PIV certification at least once.

PIV patent challenges by definition only concern patents before they expire. However, six patents in our sample are listed in the FDA approval letters as challenged through PIV certification even though the patents expired before the approval of the first generic drug. We classify them as non-challenged, but the results do not change if we assign them as challenged.

Missing Observations in PIV Challenged Patent Numbers. Some of the PIV challenged

patent numbers are missing from our dataset. The FDA provided us with a list of 1020 ANDAs containing PIV certifications and we have collected the 677 approval letters for these ANDAs. There are thus 343 ANDAs in our sample for which the exact numbers of challenged patents are missing.

While we continue to file the FOIA requests to the FDA to update our data, we believe that our estimation sample is comprehensive enough to allow for an accurate measurement of both our outcome variable, PIV patent challenge, and its timing: First, when filing the FOIA requests to the FDA, we *randomize* over the target ANDA approval letters. Since the FDA sends us exactly what we ask for, each month we receive approval letters of several randomly chosen ANDAs and add the challenged patent numbers from these letters to our sample. For some time now, our outcome variable of interest, PIV patent challenge, has remained virtually unchanged as we add missing patent numbers. Second, to measure PIV patent challenge correctly, we only have to observe one of potentially many successful challenges of a new drug patent.

Third, we aggregate the challenges to the active ingredient level, and measure whether or not we observe both 180-day generic exclusivity and a PIV challenged patent for each challenged active ingredient. That 180-day exclusivity is granted to the first filer of an ANDA and reliably measures the first successful PIV patent challenge of an active ingredient. ANDAs which have received 180-day exclusivity are marked by "PC" designation in the Orange Book. Our sample includes 1009 unique active ingredients and 150 of them are associated with ANDAs holding 180-day exclusivity. Out of those 150 active ingredients, we fail to observe a PIV challenged patent number only in 13 cases. Hence, the probability of PIV patent challenge of a drug is almost the same (0.15 versus 0.14) regardless of whether it is calculated based on the 180-day exclusivity or our granular patent challenge information. The probability of PIV patent challenge is also fairly similar at the active ingredient level and at the patent level (0.15 versus 0.17), despite the different units of observation (see Table 1 for the patent-level summary statistics). Fourth, when looking at those 150 PIV patent challenged active ingredients, the average year of the first PIV patent challenge is in practice the same regardless of whether we calculate it based on the earliest approval date of generic drugs with 180-day exclusivity or based on the earliest approval date of generic drugs with non-missing challenged patent numbers. We thus also appear to measure the timing of (the first) PIV patent challenge reliably.

Online A.1.3 Measuring Effective Patent Length, Grant Lags, PTAs, and Some Patent Characteristics: USPTO PatEx and Orange Book

Effective Patent Length. Our primary measure of the effective length of patent i is

$$\text{Effective length}_i = \text{Expiration date}_i - \max\{\text{Grant date}_i, \text{Drug approval date}_i\},$$

in which Expiration date_i is the date when patent i expired, $\text{Drug approval date}_i$ is the date when the FDA approved the first of potentially many new drugs protected by patent i, and Grant date_i is the date when the USPTO issued patent i. We identify patent grant dates from the Application

Data Tab of the USPTO PatEx, and the FDA approval and patent expiration dates from the product and patent files of the Orange Book. The expiration dates take into account various patent term extensions and adjustments that compensate for regulatory delays at the FDA and at the USPTO separately.

Patent Grant Lags. The grant lag of patent i is calculated as

$$\text{Grant lag}_i = \text{Grant date}_i - \text{Filing date}_i,$$

in which Filing date$_i$ is the date on which USPTO received the application that was subsequently issued as patent i. Like grant dates, we identify patent filing dates and application numbers from the Application Data Tab of PatEx.

PTAs. We use patent application numbers to collect PTAs from the Patent Term Adjustment Tab of PatEx. Compensations for patent term forgone due to the USPTO regulatory delays are labeled as patent term extensions for patents filed after the adoption of TRIPS on June 8, 1995 but before the implementation of AIPA, on May 29, 2000, after which the compensations are labeled as patent term adjustments.[8] For brevity and following the variable label in the PatEx files, we call both of these term modifications PTAs.

Other Patent Characteristics from PatEx. We assign the value one to the continuing patent indicator if a patent has been filed as a continuation, a continuation-in-part or a divisional application. We identify this information using application numbers from the Continuity Data Tab of PatEx where continuation, continuation-in-part and divisional applications are recorded as "CON", "CIP" and "DIV", respectively.

We use the Application Data Tab of PatEx to retrieve the USPC numbers and examiners' Art Unit codes. We identify the Technology Centers from the first two digits of the Art Unit codes. As the number of patents and their examiners in some Technology Centers is small, we combine patents prosecuted in the Technology Centers that have issued less than 100 new drug patents into one group (with 188 patents).

Online A.1.4 Measuring Claim Scope: USPTO Patent Claims Research Dataset

Our measures of patent claim scope are the numbers of Markush groups (as proxied by the phrase "selected from"), conjunctions "or" and words in the first independent claim, and the number of independent claims in each patent. The data on these measures of claim scope is available from the USPTO Patent Claims Research Dataset and we link it to our main dataset using patent numbers.

We instrument these measures of claim scope with the "examiner's historical average" calculated according to equation (3). To construct the instruments, we first identify the examiners of new drug patents from the Application Tab of PatEx. We then find all patents reviewed by these examiners and link the data using the patent number to the USPTO Patent Claims Research Dataset. Finally, we identify the scope outcomes for all granted patents reviewed by these examiners.

[8] See, e.g., `https://www.uspto.gov/web/offices/pac/mpep/s2720.html` and `https://www.uspto.gov/web/offices/pac/mpep/s2730.html` (accessed April 3, 2020).

Online A.1.5 Measuring Method and Active Ingredient Patents: Google Patent

We combine text recognition algorithms and manual verification to identify method and active ingredient patents. Our classification approach employs the texts of the abstracts and first claims of patents. We collected the full texts of patent abstracts and the first claims using Google Patent Search Engine and data scrapping algorithms (written in Python using `BeautifulSoup` and `urlib2` packages.)

Method patents covering drugs mostly pertain to the method of use of a drug or the ultimate intended effect of a patented chemical. Typically the first claim of a method patent begins with a word "method" or "process". For example, the first claim of the USPTO patent number *4870105* begins with: "A method for inhibiting gastrointestinal absorption ...". Hence, we assign the the method patent indicator equal to one, if the first claim of the patent begins with the words "method" or "process". We manually verified texts of multiple patents to ensure the accuracy of the assignment.

We assign the value one to the active ingredient indicator if the first claim of a patent pertains a chemical formula. We classify the first claim as a chemical formula when its first four symbols are a combination of single letters, dashes, commas or digits, for example, "N,N" or "R-R". The active ingredient indicator also equals one if the first three words of the first claim indicate a chemical compound and its variations. Word combinations used in chemical compound patents are, e.g., "composition of matter", "compound", "chemical compound", "antiviral compound", "amine", and "peptide". We also identify patents claiming solid forms (crystallines) of compounds, derivatives of compounds, and some other cases based as active ingredient patents. In identifying the active ingredient patents, we used Python package `re` and relied on regular expressions.

Manual checks suggest that the patents in our sample not classified as methods nor active ingredients mostly pertain to drug delivery systems, devices, or formulations.

Online A.1.6 Measuring Backward and Forward Citations: USPTO PatFT

We collect the data on forward and backward citations from PatFT using a Python algorithm. We define the number of backward citations as the number of patent documents (including foreign) listed on under the headline "References Cited" of our sample patent, and the number of forward citations as the number of the U.S. issued patents mentioning our sample in their "References Cited" list. Forward citation data was collected from PatFT using "Referenced By" retrieval tool in October 2017.

Online A.1.7 Measuring Patent Family: EPO Open Patent Services

Various definitions of patent family are used in the literature. One commonly used definition of patent family (see, e.g., Lanjouw and Schankerman, 2004; Sampat and Williams, 2019) is the number of distinct countries where the same patent has been filed. We follow this definition and construct patent family size as follows: We first collect the "DOCDB simple patent families" of the new drug

patents in our sample from the Open Patent Services of the EPO in October 2017 [9]. To collect the data, we used an algorithm and the Application Programming Interface granted to authors by the EPO. (The algorithm is written in Python using `epo_ops` and `xml` packages).

A "DOCDB simple patent family" may contain multiple patents from one patent office since it, e.g., includes continuing patents in addition to their parent patents. To solve the problem, we count the unique countries based on the country codes, which are the two letters preceding patent numbers retrievable from the Open Patent Services (e.g., Canadian patents contain a prefix CA, Japanese JP, Finnish FI, *etc.*). Thus, even if the DOCDB simple patent family lists several patents with the country code JP, we count Japan in our family size variable only once.

Online A.2 Patent Length Estimations: Robustness

Online A.2.1 Event Study

We estimate the following event study specification using the same controls as in the DiD specification in column (3) of Table 3:

$$\begin{aligned}\mathbb{1}[\text{PIV entry}_{it}] \;=\; & \alpha + \beta_1 \mathbb{1}[\text{Grant lag}_i \geq 3\text{years}] \\ & + \sum_{l=-5, l\neq -1}^{4} \mu_l \mathbb{1}[\text{Grant lag}_i \geq 3\text{years}] \times \mathbb{1}[t - 1995 \in \{2l, 2l+1\}] \\ & + \mu_{-6} \mathbb{1}[\text{Grant lag}_i \geq 3\text{years}] \times \mathbb{1}[t - 1995 < -10] \\ & + \mu_5 \mathbb{1}[\text{Grant lag}_i \geq 3\text{years}] \times \mathbb{1}[t - 1995 > 9] + \gamma' \mathbf{X_i} + \delta_t + \varepsilon_{it}.\end{aligned} \tag{A1}$$

In equation (A1), $\mathbb{1}[t - 1995 \in \{2l, 2l+1\}]$ is an indicator variable equaling one if the difference between the filing year t of patent i and the implementation year of TRIPS is either $2l$ or $2l+1$ years. To estimate the event study coefficients μ_l around TRIPS, we thus bin observations into two-year periods to improve the precision of estimates (the average number of patents per filing year is only 103 with a standard deviation of 74). We also bin the largest leads and the largest lags, respectively, as shown by the third and fourth rows of equation (A1). The first lead $l = -1$, i.e., patents filed one and two years before TRIPS, is excluded as a normalization. The coefficients μ_l for $l \leq -2$, $l = 0, 1, 2$ and $l \geq 3$ capture, respectively, possible pre-trends in the outcome, the effects of TRIPS, and the the effect of AIPA.

We report the event study estimates in Figure A1. The x-axis shows the number of years between patent filing and the implementation of TRIPS in 1995. Five years after TRIPS corresponds to the implementation of AIPA in 2000. The event study estimates are consistent with the DiD estimates but more imprecise: they suggest a decrease in the probability of PIV patent challenge between TRIPS and AIPA on average. There appears to be no statistically significant pre-trend in the outcome.

[9] https://www.epo.org/searching-for-patents/helpful-resources/first-time-here/patent-families/docdb.html (accessed April 24, 2020)

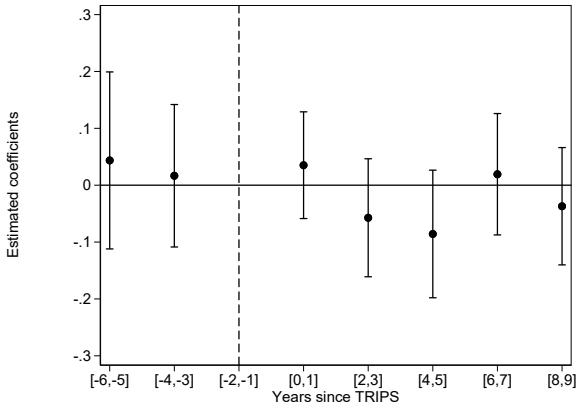


Figure A1: Event Study Estimates.

Notes: This figure shows event study coefficient estimates and 90 percent confidence intervals of the effects of TRIPS and AIPA on the probability of successful PIV patent challenge. The coefficients are estimated for each two-year period, rather than each year, due to a small number of patents per filing year (103 on average with a standard deviation of 74). The x-axis shows the number of years between patent filing and the implementation of TRIPS in 1995. Five years after TRIPS corresponds to the implementation of AIPA in 2000. The horizontal line represents the omitted coefficient for one and two years before TRIPS. The specification includes the same controls as the one in column (3) of Table 3. Standard errors are clustered at the level of patents protecting the same active ingredient.

Online A.2.2 Other Impacts of TRIPS and AIPA

Besides the changes affecting effective patent length described in Section 4.1, TRIPS and AIPA introduced other changes to the U.S. patent law. We next analyze some of these other changes that could affect the credibility or interpretation of our results concerning the effect of effective patent length on the probability of PIV patent challenge.

Provisional Patent Applications. TRIPS introduced a provisional patent application as a simplified version of a regular utility patent application. Provisional applications include no claims nor description of prior art, are not subject to examination, and automatically expire after one year (35 U.S.C. § 111 (b)). The purpose of a provisional application is to give priority rights on invention without starting the patent clock: If a regular utility patent application is filed for the same invention before the expiration of its provisional application, the priority date is the provisional application date but the patent term is calculated from the utility patent application date.

A threat to our identification might arise if the use of provisional applications varies across patents depending on their prosecution time. To address this concern, we first create the provisional application indicator which equals one if a patent claims the priority date of a provisional application. We identify these patents based on the "PRO" designations listed in the Continuity Data Tab of PatEx. We then estimate the DiD model of equation (2) using the provisional application indicator as the dependent variable.

The results reported in column (1) of Table A1 suggest no systematic relationship between the probability of a patent claiming the priority date of a provisional application and its grant lag in the

different policy regimes. Another robustness check confirms this conclusion: when we estimate the specification in equation (2) using a sample excluding all patents with the provisional application indicator equal to one, the effects of TRIPS and AIPA, reported in column (1) of Table A2 in the next subsection, become stronger and more precisely estimated than in Table 3.

Table A1: Patent Characteristics and Patent Law Changes by Patent Grant Lag.

Outcome	Provi-sional (1)	Conti-nuing (2)	log(Markush groups+1) (3)	Active ingredient (4)	log(Public length) (5)	log(Family size) (6)	log(Backward citations+1) (7)
Grant lag ≥3 years	0.000	0.076	0.030	0.001	-0.081	-0.174	0.385
	(0.000)	(0.030)	(0.038)	(0.033)	(0.013)	(0.091)	(0.078)
Grant lag ≥3 years, Post-TRIPS	0.024	-0.138	-0.032	-0.012	-0.088	0.013	0.140
	(0.030)	(0.055)	(0.054)	(0.041)	(0.023)	(0.141)	(0.123)
Grant lag ≥3 years, Post-AIPA	0.019	-0.193	-0.036	-0.063	0.294	0.216	-0.078
	(0.045)	(0.052)	(0.058)	(0.037)	(0.026)	(0.140)	(0.133)
Mean dep. variable	0.214	0.589	0.247	0.226	2.777	1.905	2.760
Observations	3517	3517	3485	3517	3517	3511	3517
Filing year FE	×	×	×	×	×	×	×

Notes: This table reports estimates of the effects of TRIPS and AIPA on patent characteristics. Columns (1)–(7) show coefficients from an OLS regression of the Provisional application indicator, the Continuing patent indicator, log(Markush groups+1), the Active ingredient patent indicator, log(Public length), log(Patent family size), and log(Backward citations+1), respectively, on three different indicators for patents with at least a three-year grant lag and filing year fixed effects (FEs). Standard errors, in parentheses, are clustered at the level of patents protecting the same active ingredient.

Continuing Patents. In the case of continuing patents, which represent close to 60 percent of our sample (see Table 1), TRIPS had an impact on effective patent length besides those effects discussed in Section 4.1. Before TRIPS, the terms of continuing patents and other patents were calculated in the same way, implying that continuing patents typically expired later than their (earlier-filed) parent patents. After TRIPS, continuing patents expire simultaneously with their parent patent (35 U.S.C. §154 (a)(2)). Thus, TRIPS shortened the effective patent length of continuing patents more than that of other patents, and this shortening is stronger for continuing patents prosecuted over three years. As a result, long patent prosecution times might incentivize originators to opt for separate patent applications instead of continuing applications after TRIPS. Supporting this idea, patents with long prosecution lags are less likely to be continuing after TRIPS, compared to other patents in our sample, as shown in column (2) of Table A1.

A smaller share of continuing patents in the post-TRIPS period of our sample has several implications: First, it implies a stronger variation in effective patent length stemming from the variation in patent grant lags. Second, since continuing patent applications might be used as a tool to make patent protection stronger (see, e.g., Lemley and Moore 2004), it raises the concern that our baseline estimates of the effect of TRIPS on patents with long prosecution lags would partially reflect narrower patent protection. To address this concern, we estimate the DiD model of equation (2) using a sample of continuing patents. The results, reported in column (2) of Table A2, support

our main result of the positive impact of longer effective patent terms on the probability of PIV patent challenge: the effects of TRIPS and AIPA on the patents prosecuted at least three years are stronger than in any other specification. The results in columns (3), (4), and (7), discussed at the end of the subsection, also mitigate the concern that potentially narrower patent protection in the post-TRIPS period would be driving our main results.

Third, taken together, the results in columns (2) of Tables A1 and A2 suggest that TRIPS made continuing patents with long grant lags less valuable for originators even if it simultaneously reduced the threat of PIV patent challenges to these patents. We interpret these results as reflecting the strong negative effect of TRIPS on the effective length of continuing patents with long grant lags. Then, by revealed preference, the results also support our policy conclusion of Section 7 that $f(p_G^*) > 0$ is more likely than $f(p_G^*) < 0$: if $f(p_G^*)$ were negative, we should have observed the proportion of continuing patents to increase in the post-TRIPS period.

Disclosure of Patent Applications. AIPA affected disclosure of patent applications (see, e.g., Johnson and Popp 2003). Prior to AIPA, patent applications were not published before they were issued. AIPA harmonized disclosure in the United States with international standards according to which a pending patent application is published 18 months after its filing date. For patents with long grant lags, the resulting loss in secrecy can be substantial, whereas patents granted (and, by implication, published) within 18 months loose little. Earlier disclosure of a patent application may lengthen the effective time for patent challenging, and hence affect the interpretation of our results concerning the effect of AIPA.

To evaluate the effects of AIPA on patent information disclosure in our setting, we measure the new drug patent length beginning from its disclosure. We define the public length of patent i filed in year t as

$$\text{Public length}_{it} = \begin{cases} \text{Expiration date}_i - \text{Grant date}_i, \\ \qquad \text{if } \mathbb{1}[\text{Grant lag}_i \geq 18 \text{ months}] \times \mathbb{1}[\text{Post-AIPA}_i] = 0 \\ \text{Expiration date}_i - \text{Filing date}_i + 18 \text{ months}, \\ \qquad \text{if } \mathbb{1}[\text{Grant lag}_i \geq 18 \text{ months}] \times \mathbb{1}[\text{Post-AIPA}_i] = 1, \end{cases}$$

in which the indicator variable $\mathbb{1}[\text{Grant lag}_i \geq 18 \text{ months}]$ gets, analogous to $\mathbb{1}[\text{Grant lag}_i \geq 3 \text{ years}]$ of equation (2), value one if the grant lag of patent i is at least 18 months.

Our calculation provides a crude measure of the effect of AIPA on public patent length. For instance, the calculation ignores the disclosure of the U.S. inventors' international applications after 18 months prior to AIPA, the exception to the post-AIPA publication requirement concerning applicants who waive the possibility of international patenting (35 U.S.C. § 122 (b)(2)(B)(i)), and the requirement that the filing date of patent i should be measured from its earliest filing date in the post-AIPA period (35 U.S.C. § 122 (b)(1)(A)).

We then estimate the DiD model of equation (2) using $\log(\text{Public length}_i)$ as the dependent variable. Column (5) of Table A1 reports the results. Compared to the estimate reported in column

(6) of Table 3, the coefficient estimate of the term $\mathbb{1}[\text{Grant lag}_i \geq 3 \text{ years}] \times \mathbb{1}[\text{Post-AIPA}_i]$ is now larger. While this result may arise from the imprecise measurement of public length, it may also indicate that an increase in public length partially contributes to the estimated positive effect of the longer effective patent length on the probability of PIV patent challenge. Overall, however, estimates of the effects of TRIPS and AIPA depending on grant lags reported in column (6) of Table 3 and column (5) of Table A1 are similar.

Other Changes due to TRIPS and AIPA. Finally, we explore whether our results concerning the differential impacts of TRIPS and AIPA on the probability of PIV patent challenge depending on grant lags could be explained by simultaneous differential changes in patent scope or value induced by various patent law modifications of TRIPS and AIPA. In this respect, estimates reported columns (3), (4), (6), and (7) of Table A1 are comforting: we observe no clear impacts of TRIPS and AIPA on the relationship between grant lags and patent scope or value, as measured by the number of Markush groups in the first independent claim, the proportion of active ingredient patents, patent family size, and the number of backward citations.

Online A.2.3 Other Subsamples and Specifications

Table A2 provides results from the estimation of the DiD model of equation (2) using different subsamples and specifications. We report estimates using a sample that excludes patents claiming the priority date of a provisional application in column (1) and using a sample that only includes continuing patents in column (2).

The results in column (3) are generated by a sample that excludes all other new drug patents sharing the common parent application except for the latest one. We identify the excluded patents using information on parent applications from the Continuity Data Tab of PatEx. In column (4) the results arise from a sample that only includes patents protecting drugs for which we observe FDA exclusivity. These patents might be more valuable or, as our results of Table (2) indicate, more difficult to challenge than patents protecting drugs for which we observe no FDA exclusivity. We observe no exclusivity if a patent protected drug did not have one or if its exclusivity expired before 2001. Also, when constructing fixed effects for the latest FDA exclusivity expiry year, we group together patents protecting drugs without observed exclusivity.

Column (5) reports estimates from a model that, analogous to survival models, includes fixed effects to control for the length of exposure of a patent to PIV patent challenges. In this specification we, in addition to a full set of other controls, include an indicator measuring the time (in years) between the grant date of a patent and its expiration date or the earliest successful PIV patent challenge date. whichever is earlier. Finally, we estimate a model allowing for richer interactions between grant lags and patent policy reforms, and report the results in column (6).

The signs and magnitudes of the coefficient estimates reported in Table A2 are similar to those coefficients reported in columns (1)–(5) of Table 3. If anything, the effects of TRIPS and AIPA are now stronger and more precisely estimated. Coefficients of the additional interaction terms involving patents with a grant lag exceeding five years reported in Column (6) are imprecisely

estimated, possibly because we have only 417 such patents in our sample.

Table A2: Probability of Successful PIV Patent Challenge and Patent Law Changes by Patent Grant Lag: Robustness Analysis.

	(1)	(2)	(3)	(4)	(5)	(6)
Grant lag ≥3 years	0.025	0.030	0.016	0.032	0.040	
	(0.025)	(0.029)	(0.029)	(0.031)	(0.024)	
Grant lag ≥3 years,						
Post-TRIPS	-0.123	-0.180	-0.126	-0.118	-0.132	
	(0.041)	(0.052)	(0.044)	(0.046)	(0.035)	
Grant lag ≥3 years,						
Post-AIPA	0.089	0.154	0.105	0.087	0.090	
	(0.044)	(0.051)	(0.041)	(0.039)	(0.030)	
Grant lag 3-5 years						0.023
						(0.027)
Grant lag 3-5 years,						
Post-TRIPS						-0.102
						(0.041)
Grant lag 3-5 years,						
Post-AIPA						0.075
						(0.039)
Grant lag > 5 years						0.033
						(0.039)
Grant lag > 5 years,						
Post-TRIPS						-0.117
						(0.071)
Grant lag > 5 years,						
Post-AIPA						0.065
						(0.063)
Mean dependent variable	0.179	0.176	0.167	0.180	0.173	0.173
Observations	2731	2054	2627	2668	3478	3483
R-squared	0.233	0.268	0.210	0.262	0.385	0.234
Exposure length FE					×	
Filing year FE	×	×	×	×	×	×
Drug controls	×	×	×	×	×	×
Exclusivity expiration year FE	×	×	×	×	×	×
Patent controls	×	×	×	×	×	×
USPC FE	×	×	×	×	×	×
Sample	No prior provisional	Continuing	Latest filed	FDA exclusivity	Full	Full

Notes: This table reports estimates of the effects of TRIPS and AIPA on the probability of successful PIV patent challenge. Columns (1)–(5) show coefficients from an OLS regression of the PIV challenge indicator on three and column (6) on five indicators for patents with at least a three-year grant lag and controls. FE stands for fixed effects. Drug controls include the indicators New chemical exclusivity, Orphan drug exclusivity, Pediatric exclusivity, Priority review, Capsule, Injectable, and Tablet. Patent controls include log(Markush groups+1), log(Backward citations+1), log(Forward citations+1), log(Patent family size), and the indicators Active ingredient patent, Method patent, and Continuing patent. In columns (1)–(4) the samples exclude patents claiming the priority date of a provisional application, first-filed patents covering the same invention, all but the latest-filed patents covering the same invention, and patents protecting drugs without (observed) FDA exclusivity, respectively. Standard errors, in parentheses, are clustered at the level of patents protecting the same active ingredient.

Online A.3 Patent Scope Estimations: Robustness

Table A3 reports estimates of various claim scope measures from the OLS model of equation (1). Column (1) shows an estimate using the specification of column (3) in Table (2). We then replace the number of Markush groups in the first independent claim with alternative measures of patent claim scope: the number of the conjunction "or" in the first independent claim (column (2)), the number of words in the first independent claim (column (3)), the number of independent claims (column (4)). Only the coefficient estimate in column (2) suggests a statistically significant negative relationship between the probability of PIV patent challenge and patent scope. When compared with our IV estimates from Table (5), there appears to be an upward bias in most of these OLS estimates.

Tables (A4)–(A6) report the results from our exploration of 2SLS estimates of patent scope measures. Tables (A4) and (A5) show that the results reported in Table (5) are robust, respectively, to the exclusion of control variables, and to the inclusion of the interaction of Technology Center fixed effects with filing year fixed effects, which controls for Technology Center specific trends. Table (A6) reports estimates from a regression in which we replace Technology Center fixed effects with Art Unit fixed effects. The number of patents per Art Unit is small, only 21 on average. Even if we group Art Units with less than 10 patents into one group of 266 patents, coefficients of the scope measures become less precisely estimated (although remaining similar in magnitude) than in Table (5). The only scope measure generating statistically significant effects is the number of words in the first independent claim.

Our results are robust to using the following alternative measures of claim scope: the number of claims, and the numbers of characters and the phrase "consisting of" in the first independent claim, with the last measure being an alternative proxy for a Markush group (see Section 2.4). We have also experimented with a binary variable of whether or not there is at least one Markush group in the first independent claim without changing the results. We omit these regression results for brevity.

Table A3: Probability of Successful PIV Patent Challenge and Patent Scope: OLS Estimates.

	(1)	(2)	(3)	(4)
log(Markush groups+1)	-0.013			
	(0.012)			
log(Conjunctions "or"+1)		-0.020		
		(0.007)		
log(Words)			-0.001	
			(0.007)	
log(Independent claims)				-0.009
				(0.008)
Mean dependent variable	0.173	0.173	0.173	0.173
Observations	3483	3483	3483	3485
R-squared	0.242	0.241	0.243	0.241
Technology Center FE	×	×	×	×
Filing year FE	×	×	×	×
Drug controls	×	×	×	×
Exclusivity expiration year FE	×	×	×	×
Patent controls	×	×	×	×
USPC FE	×	×	×	×

Notes: This table reports coefficients from OLS regressions of the probability of successful PIV patent challenge on scope measures and controls. FE stands for fixed effects. Drug controls include the indicators New chemical exclusivity, Orphan drug exclusivity, Pediatric exclusivity, Priority review, Capsule, Injectable, and Tablet. Patent controls include log(Effective length), log(Backward citations+1), log(Forward citations+1), log(Patent family size), and the indicators Active ingredient patent, Method patent, and Continuing patent. Standard errors, in parentheses, are clustered at the level of patents protecting the same active ingredient.

Table A4: Probability of Successful PIV Patent Challenge and Patent Scope: Robustness of IV Estimates to Exclusion of Control Variables.

	(1)	(2)	(3)	(4)
log(Markush groups+1)	-0.244			
	(0.071)			
log(Conjunctions "or"+1)		-0.095		
		(0.030)		
log(Words)			-0.484	
			(0.233)	
log(Independent claims)				-0.154
				(0.123)
Mean dependent variable	0.173	0.173	0.173	0.173
Observations	3447	3447	3447	3449
First-stage F-statistic	41.446	127.960	5.417	15.264
Technology Center FE	×	×	×	×
Filing year FE	×	×	×	×
USPC FE	×	×	×	×

Notes: This table reports coefficients from 2SLS regressions of the probability of successful PIV patent challenge on instrumented scope measures and controls. The instruments are the cumulative averages of the scope measures of all patents reviewed by the examiner of a new drug patent, until one year preceding the filing year of the new drug patent. The first stage F-statistic tests is on the excluded instruments. FE stands for fixed effects, respectively. Robust standard errors are in parentheses.

Table A5: Probability of Successful PIV Patent Challenge and Patent Scope: Robustness of IV Estimates to Technology Center Specific Trends.

	(1)	(2)	(3)	(4)
log(Markush groups+1)	-0.218			
	(0.091)			
log(Conjunctions "or"+1)		-0.101		
		(0.052)		
log(Words)			-0.219	
			(0.091)	
log(Independent claims)				0.062
				(0.132)
Mean dependent variable	0.173	0.173	0.173	0.173
Observations	3445	3445	3445	3447
First-stage F-statistic	24.520	43.897	14.769	10.592
Technology Center FE × Filing year FE	×	×	×	×
Drug controls	×	×	×	×
Exclusivity expiration year FE	×	×	×	×
Patent controls	×	×	×	×
USPC FE	×	×	×	×

Notes: This table reports coefficients from 2SLS regressions of the probability of successful PIV patent challenge on instrumented scope measures and controls. The instruments are the cumulative averages of the scope measures of all patents reviewed by the examiner of a new drug patent, until one year preceding the filing year of the new drug patent. The first stage F-statistic tests is on the excluded instruments. FE stands for fixed effects. Drug controls include the indicators New chemical exclusivity, Orphan drug exclusivity, Pediatric exclusivity, Priority review, Capsule, Injectable, and Tablet. Patent controls include log(Effective length), log(Backward citations+1), log(Forward citations+1), log(Patent family size), and the indicators Active ingredient patent, Method patent, and Continuing patent. Robust standard errors are in parentheses.

Table A6: Probability of Successful PIV Patent Challenge and Patent Scope: Robustness of IV Estimates to Art Unit Fixed Effects.

	(1)	(2)	(3)	(4)
log(Markush groups + 1)	-0.257			
	(0.214)			
log(Conjunctions "or"+1)		-0.073		
		(0.065)		
log(Words)			-0.190	
			(0.077)	
log(Independent claims)				0.117
				(0.172)
Mean dependent variable	0.173	0.173	0.173	0.173
Observations	3445	3445	3445	3447
First-stage F-statistic	6.202	30.447	18.815	6.669
Art Unit FE	×	×	×	×
Filing year FE	×	×	×	×
Drug controls	×	×	×	×
Exclusivity expiration year FE	×	×	×	×
Patent controls	×	×	×	×
USPC FE	×	×	×	×

Notes: This table reports coefficients from 2SLS regressions of the probability of successful PIV patent challenge on instrumented scope measures and controls. The instruments are the cumulative averages of the scope measures of all patents reviewed by the examiner of a new drug patent, until one year preceding the filing year of the new drug patent. The first-stage F-statistic tests is on the excluded instruments. FE stands for fixed effects. Drug controls include the indicators New chemical exclusivity, Orphan drug exclusivity, Pediatric exclusivity, Priority review, Capsule, Injectable, and Tablet. Patent controls include log(Effective length), log(Backward citations+1), log(Forward citations+1), log(Patent family size), and the indicators Active ingredient patent, Method patent, and Continuing patent. Robust standard errors are in parentheses.

Online A.4 Results Without Assumption 1

As the main text characterize results for the case $\partial^2 C_G/(\partial p_G \partial b) > 0$, we here focus on the case $\partial^2 C_G/(\partial p_G \partial b) \leq 0$. The proofs of Proposition 1 and 2 imply that if $\partial^2 C_G/\partial p_G \partial b = 0$, changes in patent scope have no impact on the firms' incentives and, if $\partial^2 C_G/(\partial p_G \partial b) < 0$, an increase in patent scope *increases* incentives for patent challenges and *reduces* incentives for new drug development. The results of Proposition 1 and 2 concerning the effect of patent length remain unaffected.

We next analyse the optimal structure of patent policy for $\partial^2 C_G/(\partial p_G \partial b) \leq 0$, implying that $\epsilon_b \leq 0$. We first consider the case $\partial^2 C_G/(\partial p_G \partial b) < 0$, before moving to the case $\partial^2 C_G/(\partial p_G \partial b) = 0$.

When $\partial^2 C_G/(\partial p_G \partial b) < 0$, the counterpart to Proposition 3 can be expressed as follows:

Proposition A1: *Assume that* $\partial^2 C_G/(\partial p_G \partial b) < 0$. *Then, i) if* $\epsilon_b < \phi(p_G^*) < 0$, *reducing patent length and increasing patent scope is efficient; ii) If* $\phi(p_G^*) > 0$, *reducing both patent length and scope is efficient; iii) If* $\phi(p_G^*) < \epsilon_b$, *reducing patent scope and increasing patent length is efficient.*

Proof: The proof follows the proof of Proposition 3. The efficient structure of patent policy is still characterized by the signs of equations (23) and (26). As before, equation (26) tells us the optimal direction of patent scope and equation (23) tells us the direction in which patent length needs to be adjusted so as to compensate the change in patent scope.

But now $\partial^2 C_G/(\partial p_G \partial b) < 0$ and, consequently, the proof of Proposition 1 implies that $\partial p_G/\partial b > 0$. Thus, equation (23) shows that the sign of $\partial T/\partial b$ is given by the sign of $f(p_G^*)$.

Note next from equation (26), that if $\phi(p_G^*) > 0$, then $dV^S/db < 0$ because $\epsilon_b < 0$. If $\phi(p_G^*) < 0$, the first term in the square brackets of equation (26) is positive. Then the term in the square brackets is positive if $\epsilon_b < \phi(p_G^*)\phi(p_G^*)$, implying $dV^S/db < 0$.

To sum up: i) If $\epsilon_b < \phi(p_G^*) < 0$, $\partial T/\partial b < 0$ and $dV^S/db > 0$. Increasing patent scope and reducing length is efficient; ii) If $\phi(p_G^*) > 0$, $\partial T/\partial b > 0$ and $dV^S/db < 0$. Reducing both patent scope and length is efficient; iii) If $\phi(p_G^*) < \epsilon_b$, $\partial T/\partial b < 0$ and $dV^S/db < 0$. Reducing patent scope and increasing length is efficient. □

Recall that if $\partial^2 C_G/(\partial p_G \partial b) < 0$, an increase in patent scope has counterintuitive effects: Even if an increase in patent scope continues to make patent challenging more expensive, it has a *positive* effect on incentives for patent challenging and, consequently, a *negative* impact on incentives to develop new drugs. With this observation, the explanation of Proposition A1 is analogous to the one of Proposition 3.

Finally, let us consider the case $\partial^2 C_G/(\partial p_G \partial b) = 0$, implying that $\epsilon_b = 0$. To simplify the analysis, we use a standard relationship between welfare flow and market structure and define $w_N := N\pi_N + cs_N$ in which $N\pi_N$ and cs_N are industry profits and consumer surplus (when $N \in \{0, 1, 2\}$ drugs compete in the market), respectively. Furthermore, define $T' := \arg\max p_B^*(T)$ and assume for the moment that the conditions stipulated in footnote 6, $\partial\phi/\partial p_G < 0$ and $\lim_{T\to\infty} \phi(p_G^*(T)) < 0$, hold. Then T' is a finite and strictly positive unique solution to $\phi(p_G^*(T')) = 0$. Under these assumptions, we get the following result:

Proposition A2: *Assume that* $\partial^2 C_G/(\partial p_G \partial b) = 0$. *Then, patents should have a minimum scope. The optimal patent duration is given by* $T^* := \arg\max_T W(T) = p_B^*(T) V^S(T) - C_B(p_B^*(T))$. *in which* $T^* < T'$.

Proof: We may write the total (ex ante) welfare from a new drug as

$$W(b,T) = p_B^*(b,T) V^S(b,T) - C_B(p_B^*(b,T)) \tag{A2}$$

in which $p_B^*(b,T)$ and $V^S(b,T)$ are given by equations (9) and (11), respectively, and $C_B(p_B(b,T))$ is the originator's cost of developing a new drug.

Differentiate next $W(b,T)$ from equation (A2) with respect to b. Since $\partial^2 C_G/(\partial p_G \partial b) = 0$, $\partial p_B^*/\partial b = 0$. As a result, the sign of $\partial W(b,T)/\partial b$ is given by the sign of $\partial V^S/\partial b$. Since $\partial p_G^*/\partial b = 0$, equation (24) implies that $\partial V^S/\partial b = -\partial C_G/\partial b < 0$. Thus, it is optimal to have as narrow patents as possible.

Since $\partial p_B^*/\partial b = 0$, equation (23) implies that $\partial T/\partial b = 0$. As a result, T is the only relevant dimension of the patent policy, and the socially optimal T is given by $T^* := \arg\max_T W(T)$. We next characterize the circumstances under which $T^* < \overline{T}$.

With the the help of equation (9), the first-order condition for the optimal T can be written as

$$\frac{\partial W}{\partial T} = \frac{\partial p_B^*(T)}{\partial T}\left[V^S(T) - V^P(T)\right] + p_B^*(T)\frac{\partial V^S(T)}{\partial T} = 0. \tag{A3}$$

Let us first prove that $V^S(T) - V^P(T) > 0$, i.e., that the social value of a new drug is larger than its private value. If $V^P(T) \geq V^S(T)$, the issue of the patent policy design would be moot. However, since the social value of the new drug includes the costs of patent challenging, the question of whether $V^S(T) > V^P(T)$ holds is not trivial.

Using $x_N := \widetilde{x}_N/r$, $x = w, \pi$, in subtracting equation (8) from equation (11) yields

$$\begin{aligned} & V^S(T) - V^P(T) \qquad\qquad (A4)\\ = \; & \left(1 - e^{-rT}\right)\left[\left(1 - p_G^*(T)\right)(w_1 - \pi_1) + p_G^*(T)(w_2 - \pi_2)\right] + e^{-rT}(w_2 - \pi_2) - C_G(p_G^*(T)). \end{aligned}$$

From equation (6) we observe that $\Pi_G(p_G^*(T)) \geq 0$ implies that $\left(1 - e^{-rT}\right)\pi_2 p_G^*(T) \geq C_G(p_G^*(T))$. Approximating the right-hand side of equation (A4) downwards by substituting $\left(1 - e^{-rT}\right)\pi_2 p_G^*(T)$ for $C_G(p_G(T)^*)$ gives

$$\begin{aligned} & V^S(T) - V^P(T) \\ \geq \; & \left(1 - e^{-rT}\right)\left[\left(1 - p_G^*(T)\right)(w_1 - \pi_1) + p_G^*(T)(w_2 - 2\pi_2)\right] + e^{-rT}(w_2 - \pi_2) > 0, \end{aligned}$$

in which the last inequality follows from $w_N = N\pi_N + cs_N$.

Next, we evaluate $\partial W/\partial T$ at $T = T'$. Then, $\phi(p_G^*(T')) = 0$, and equation (25) implies that $\partial V^S(T')/\partial T < 0$. From the proof of Proposition 2 we also observe that when $\phi(p_G^*(T')) = 0$, $\partial p_B^*(T')/\partial T = 0$. As a result, equation (A3) shows that $\partial W(T')/\partial T < 0$.

Assume then that i) $\partial^2 W/\partial T^2 < 0$ for all T so that the problem is well-behaving. Assume further that ii) $\partial W(0)/\partial T > 0$. A sufficient condition for ii) is $\partial p_B^*(0)/\partial T > p_B^*(0)$. To see this, note first from equation (7) that $p_G^*(0) = 0$. Then, when evaluating $\partial W/\partial T$ at $T = 0$ by using equations (25) and (A4) we get

$$\frac{\partial W(0)}{\partial T} = \frac{\partial p_B^*(0)}{\partial T}(w_2 - \pi_2) - p_B^*(0)\, r\,(w_2 - w_1),$$

in which, as indicated by equations, (8), (9), and (19), $\partial p_B^*(0)/\partial T > 0$ and $p_B^*(0) > 0$. Clearly, $w_2 - \pi_2 > r\,(w_2 - w_1)$.

If both conditions i) and ii) hold, then there exists exactly one solution for equation (A3) in the range where $T \in (0, T')$ and this solution characterizes the maximum. Note for completeness, that if condition ii) fails to hold, but condition i) holds, then the optimal policy is to set $T^* = 0$. If condition i) fails to hold, then there may be multiple solutions to equation (A3). If condition ii) nonetheless holds, there must at least be one local maximum in the range where $T \in (0, T')$. □

The explanation of Proposition A2 is the following: Since $\epsilon_b = 0$, changes in patent scope have no impact on incentives for new drug development, and an increase in patent scope only increases patent challenging costs with no welfare benefits. As a result, it is optimal to have narrow patents.

When changes in patent scope have no impact on incentives for new drug development, patent length becomes the only relevant patent policy tool. The social planner faces the classic Nordhausian patent length design problem with the twist that an increase in patent length increases incentives to challenge new drug patents, creating wasteful costs of patent challenging. Under some plausible restrictions on functional forms, the optimal length lies in the range $(0, T')$ in which patent length has a positive impact on incentives to develop new drugs and adverse impact on social welfare for a given level of drug development incentives.

Note that the proof of Proposition A2 is based on the assumptions guaranteeing that a finite T' solving $\phi(p_G^*(T')) = 0$ exists. However, if no such T' exists, then Proposition A2 holds trivially since in that case $T' \to \infty$.