Patent Length, Innovation, and the Role of Technology Disclosure Externalities*

Fabio Bertolotti†

January 2022

Please click here for the most updated version

Abstract

The length of patents is a key policy tool for innovation and long-run growth, but causal empirical evidence on its effects is scarce. This paper empirically exploits an anticipated policy change that generated quasi-experimental variation in US patent length across technical fields. A difference-in-difference analysis identifies two empirical facts. First, news of a future patent term extension causes a drop of R&D and innovation until policy implementation. Second, the drop continues even after implementation of the new, longer, term. I provide direct empirical evidence that the latter effect is driven by a technology disclosure externality. The drop in innovation at news induces a lower disclosure of novel knowledge through patents, which hinders subsequent ability of other innovators to lean from recent advances when they research ideas for new projects in the same field. Once controlling for the externality, the direct effect of longer patent length on innovation is positive, and the short-run elasticity of innovation to patent length is around 3. Theoretically, the paper proposes a semi-endogenous growth model in which i) research and development activity is split in two distinct steps, and ii) research productivity increases with more frequent disclosure of new knowledge through patents. Thanks to the novel R&D structure, the model can replicate the empirical facts. A structural estimation implies a long-run elasticity of US innovation to patent length of +0.35. Normatively, a 28-years patent length—–longer than the 20-years status quo—–would maximize welfare in the absence of policy anticipation. However, due to disclosure externalities, anticipation would lower the benefits of the 28-years term, as output would decline in the short run, as seen in data.

Keywords: Patent length, Innovation, Externality
JEL classification codes: O31, O33, O34, O38, O41

---

*I am deeply indebted to my supervisors Ricardo Reis and Xavier Jaravel for the help and the insightful feedback, as well as Andrea Alati, Philippe Aghion, Filippo Biondi, Thomas Brzustowski, Francesco Caselli, Luca Citino, Jamie Coen, Maarten De Ridder, Gaia Dossi, Arnaud Dyevre, Martina Fazio, Nicola Fontana, Alessandro Gavazza, Jonathon Hazell, Ethan Ilzetzki, Per Krusell, Andrea Lamorgese, Andrea Lanteri, Edoardo Leonardi, Salvatore Lo Bello, Massimiliano Marcellino, Benjamin Moll, Virginia Minni, Francesco Nava, Micheal Peters, Akash Raja, Giorgio Saponaro, Mark Schankerman, John Van Reenen, Alwyn Young, and Martina Zanella for their precious comments. I thank Antoine Dechezlepretre and Dennis Verhoeven for helping me access the data and Hwan Lin for sharing his numerical algorithms. Also, I thank ESRC for financial support.

†London School of Economics and Centre for Macroeconomics, email: f.bertolotti@lse.ac.uk

1 Introduction

What is the effect of patent length on innovation and welfare? Historically, this question has attracted a lot of attention from economists and industry, and now a new debate has arisen regarding patents for Covid vaccines (see, e.g., Gross and Sampat (2021)). This centrality originates from the fact that patents and their length are a critical policy tool for promoting innovation and long-run productivity growth. Therefore, it is surprising that in all developed countries the patent term is set according to international standards or a simple rule of thumb, rather than being based on empirical economic and welfare considerations.[1] Part of the reason is that economic research has so far provided no clear answers regarding the optimal patent length and its effects.

The primary welfare trade-off has been clear since Nordhaus (1967): A longer patent length increases welfare by inducing more innovation and higher productivity, but it imposes a larger deadweight loss on consumers because monopolistic distortions last longer. Nevertheless, models of optimal patent length suggest policies that range from zero (Boldrin and Levine (2013)) to infinity (Gilbert and Shapiro (1990)). However, quantitative evaluation of the previous trade-off in the data is difficult: Empirical evidence on the effect of patent length on innovation, R&D, and welfare is limited, mainly due to the scarcity of variation in this specific policy tool over time and across countries (Budish, Roin and Williams (2016)). More general evidence regarding the strength of patent rights is wider but also inconsistent. Most papers find that patents have a positive impact on innovation (e.g., Budish, Roin and Williams (2015), Moscona (2021)), but Galasso and Schankerman (2015) show that patent protection may even actually harm technical advancement in specific sectors, by blocking cumulative innovations.

This paper studies the positive and normative implications of patent length by combining causal empirical evidence with a new quantitative model. Empirically, it causally estimates the dynamic impact of an anticipated patent term change on innovation, R&D effort, and welfare, and documents a technology disclosure externality based on the pace of existing projects' development to subsequent research productivity. The paper proceeds to develop and estimate a structural model of semi-endogenous growth with novel characteristics, which builds on the empirical facts and offers new theoretical insights. Finally, this paper is the first, to the best of my knowledge, to quantitatively evaluate the welfare trade-offs that originate from a change in patent length using a structural model that is tightly linked to the

[1] For example, the patent term in the US was introduced in 1790 and set to 14 years after the grant date, in line with English law. In turn, the English term was based on the expected training period of two sets of apprentices, as reported by Nordhaus (1969), and not on any welfare considerations. In 1861, the US patent term was changed to 17 years, and Nordhaus (1969) reports that this change was the result of a political compromise.

data.

The empirical analysis relies on a dynamic difference-in-difference (DiD) specification that exploits the quasi-experimental variation of effective patent length across technical fields (4-digit IPC classes).[2] This variation originates from a policy–the *Agreement on Trade Related Aspects of Intellectual Property Rights* (henceforth TRIPs), which is a chapter of the Uruguay Round Agreements of the *General Agreement on Tariffs and Trade* (henceforth GATT)–that changed the term of a US patent from 17 years after the grant date to 20 years after the application date, and thus caused it to conform to other jurisdictions. The identification relies on the fact that patents classified in different technical fields are examined by different units within the US Patent Office (USPTO), and these units differ in terms of congestion and the technical difficulty of examination. As a result, the average pending period–the time from application to grant–varied markedly across fields before the policy. Because the TRIPs shifted the starting date of the patent term from grant to application and because patent-related monopoly is enforceable only starting from the date of the grant, the interaction of the term change with heterogeneity in the average pending period generated considerable variation in effective protection time across fields. Several analyses support the exogeneity of the preexisting heterogeneity: It is unrelated to preexisting heterogeneous trends in innovation across fields, does not endogenously respond to the policy, and is orthogonal to other TRIPs-related changes in the US innovation environment. Therefore, policy-induced variation in patent length allows me to causally estimate the impact of effective patent length on R&D effort and innovation as measured by patents, citation-weighted patents, and patent value.

Anticipation is the other crucial aspect of the policy change. The law was formally passed in December 1994 and became fully effective in June 1995, but subsection 2.1 discusses at length–based on official documents, research papers, and newspaper articles–the fact that the US business community had known the content of the policy change since at least 1992, which I take to be the policy "news" date. Therefore, the empirical analysis relies on two shocks: A "news" shock at the end of 1992 and an implementation shock in June 1995.

My analysis highlights two key empirical facts. First, news of a future patent length increase induces a reduction of R&D and innovation before implementation of the new, longer term. Second, R&D and innovation continue to be negatively affected by the policy, even after implementation of the longer patent term. The data provide suggestive evidence that the latter negative effect is temporary. The paper proceeds to empirically investigate the economic mechanisms that drive the DiD estimates.

[2]Data sources for the main analysis are PATSTAT, the NBER Patent Database, and Kogan et al. (2017). The firm-level analysis also relies on COMPUSTAT and the sectoral analysis on the NBER CES Manufacturing Database.

The data support the argument that a dynamic technological spillover drives the post-implementation fall in innovation in technological fields that are subject to a patent term increase. Because innovators build on recent technologies to produce new ones, the drop in innovation upon news of a future patent term increase–the first fact–implies a lower ability to subsequently generate new projects that are technologically related to or inspired by recent advances in the same field.

Several empirical analyses corroborate this interpretation. First, the observed drop in innovation before implementation generates a stronger negative impact on post-implementation innovation in fields in which the degree of new patents' technological dependence on previous patents in the same field is higher. I measure technological dependence using patents' backward citations.[3] Second, I document that fields that experience a fall in innovation before implementation–as a consequence of the news of a future patent term increase–reduce the intensity with which they rely on previous technologies from the same field. The latter effect is observed to occur with a delay of approximately 4 years from the negative reaction of innovation to the policy news. I relate this delay to lags in knowledge diffusion based on the fact that patents are published only after being granted–which requires 2 years from application, on average–and to research gestation lags on new projects, which are documented by Pakes and Schankerman (1986) to be around 2 years. Because completed projects and patent documents carry greater and more detailed informational content than undeveloped ideas, ongoing projects' slower pace of development decelerates the diffusion of technical knowledge to other innovators, which reduces their ability to produce new technologically related ideas. Also, I show that this effect mainly occurs between firms rather than within firms, which motivates interpretation of the effect as a spillover.[4] The estimates allow to infer the elasticity of cumulative future innovation to current innovation shocks, which is 0.997.

I also investigate other structural forces that may explain the negative relationship between patent length and post-implementation innovation outcomes. The effect does *not* seem to be driven by *i*) an adjustment of patenting strategy–e.g., a defensive breakup of patent applications in fields losing protection–or *ii*) a fall in competition in fields gaining protection as a result of discouraging new entrants, lower competitive pressure on incumbents, or the blocking effect induced by the stronger patent rights available for current innovators.

Next, I turn to interpretation of the first empirical fact. Innovation and R&D fall upon

[3]The preferred measure of within-field technological dependence is the share of patents classified in a specific field that have at least one backward citation by the patent applicant to another patent classified in the same technical field.

[4]I also document that the intensity of direct citations from post-implementation innovations to pre-implementation innovations in the same technical field decreases in fields experiencing a positive patent term change. Finally, I use COMPUSTAT firm-level data to show that firms that are *ex ante* technologically close to other firms whose R&D effort falls due to the policy news invest less in R&D after policy implementation.

news of a future patent term increase. Economically, news of a future longer term implies that a patent application filed before implementation of the new regime provides relatively shorter protection than a patent application filed afterward. This opens the door to two opposing interpretations of the empirical fact. First, innovators like a longer patent length, and therefore they reduce their innovative effort upon hearing the news in order to finish projects and file applications after implementation of the change, when they can profit from the longer term. Second, innovators may reduce innovation upon hearing the news because they anticipate that a longer patent term will have a harmful effect on innovation after implementation. Previous literature links the potential negative effect of patent protection on innovation to *i*) new innovators' lower ability to compete away incumbents, whose monopoly is protected by stronger patent rights, or *ii*) to lower effort by incumbents, who face weaker competitive pressure. However, I find that the policy does not impact measures of competition, and the average quality of incumbents' patents does not suggest their use for defensive purposes. Hence, while I cannot definitively rule out this narrative, I do not find empirical support for it.

Therefore, I focus on the first interpretation: Innovators slow the development of their existing projects at the news because they want to obtain longer protection after implementation of the longer term. Assuming that completing ongoing projects rapidly is costly, in normal times innovators set the optimal pace of development by trading-off the benefits of gaining monopolistic profits sooner against the higher costs of being fast. The news of a future patent term extension increases the benefits of being relatively slower: Innovators save on costs and may file a patent application under the new, more favorable, regime. After the policy implementation, the policy-induced incentive to be slower on development ends, but we expect that the longer patent length affects the incentives to generate new projects. The latter effect of the policy is difficult to isolate from the reduced-form DiD estimates, because the technology disclosure spillover depresses innovation in fields that gain protection after the policy implementation.

To empirically test this channel, I revisit the main DiD analysis and control for the impact of the technological spillover, which is proxied by average flow of patents in the same field in the previous four years. The first empirical fact is unaffected: The news of a future patent term extension decreases patenting before the policy implementation. While the externality has no effect during the pre-implementation phase, it becomes the main driver of innovation in the post-implementation period. The policy-induced drop in patenting in a given field between the news and the policy implementation generates a sizable reduction of patenting in the same field after implementation. Once the effect of the spillover is controlled for, the estimated coefficients quarter-specific *direct* effect of a change in patent length becomes *positive* after implementation. This confirms that the externality can fully account for the reduced-form negative relationship between patent length and innovation after implementation. Moreover, the estimate post-implementation effects imply that the short-run elasticity of patenting to patent length is 3.

Overall, the paper's empirical evidence highlights the crucial role of *anticipation* and technological *externalities* to determine the response of innovation to changes in patent length. Moreover, interpretation of the results stresses the importance of separately considering the roles of existing projects' development and research for new projects in the innovation process. The latter responds to long-run incentives to produce new ideas, while the former captures short-run intertemporal incentives to transform ideas into products. This theoretical distinction is crucial–together with anticipation and technology disclosure externalities–to correctly predict the dynamic response of innovation to the policy quasi-experiment observed in the data. Consistently, I show that standard models of R&D-based endogenous growth, where research and development are implicitly collapsed in a single activity, would not replicate the causal empirical evidence of the first part of the paper.

Therefore, I propose a new structural model to quantify the normative trade-offs away from the status quo, and to estimate key elasticities such as the elasticity of long-run innovation to patent length. I start from the variety-expansion semi-endogenous growth model of Jones (1995), adapted to allow finite patent length along the lines of Lin and Shampine (2018), and I formalize a new structure for the innovation process. The first novelty is that research and development are explicitly modeled as distinct activities.[5] Research is competitive and produces new ideas. An idea is abstract, and it can be privately stored by the firm that generates it. Subsequently, firms invest in development–i.e., in turning the stock of ideas into actual products with specific characteristics and detailed technological content. If firms are successful in the development process, they obtain a patent that grants the rights of exclusive economic exploitation of the new product for the finite patent length of T periods. Because abstract ideas cannot be patented, innovators can obtain patents only after development activity has ended. The second novelty is that the productivity of research activity is a (positive) function of the average pace of development in the economy–i.e., of how rapidly abstract ideas become products and generate detailed patent documents from which other innovators can learn. This introduces an externality from the average pace of development to subsequent research productivity, and it formalizes the technology disclosure spillover documented in the data.

Thanks to the new structure of R&D, the model can replicate the empirical facts. First,

[5]The mathematical structure is analogous to Comin and Gertler (2006), but its interpretation is different, which I show to be crucial to replicate the empirical facts.

upon news of a future patent term increase, firms slow their development of existing ideas into patented products–i.e., intermediate capital varieties. In normal times, firms trade-off the convex costs of the pace of development against the desire to more quickly obtain a patent and the related monopolistic profits. News of a future patent lengthening increases the value of an existing project, because it discounts the higher value of a *future* patent. However, the news only mildly affects the value of a *current* patent, because the longer patent term is not yet in effect. Since the benefits of obtaining the patent sooner decrease but cost convexity is unaffected, the optimal pace of development falls until implementation. Second, after policy implementation, the technology disclosure externality translates the slower pace of development at the news into lower subsequent research productivity, with a delay related to knowledge diffusion and the development lags documented in the data. This temporarily hinders research investment and causes the observed temporary drop in innovation, thus matching the second empirical fact. However, research productivity gradually recovers, because after implementation of the policy the pace of development returns to its pre-news level. In the long run, innovation and total R&D effort increase: A longer patent length leads to higher patent and project values. The latter increases research effort and promotes the creation of ideas for new varieties.

I estimate the structural parameters of the model using generalized method of moments to match the causally identified evidence on innovation and R&D effort. The model's parameters imply a mild convexity of development intensity costs but severely decreasing returns to research investment. The structural model also allows me to estimate that a 1% permanent increase in patent length increases long-run innovation flow by 0.35%.

Finally, I use the estimated model to quantitatively evaluate the normative trade-offs and the output consequences of patent term changes. The first trade-off is analogous to Nordhaus (1967) and relates to the steady state. At the current policy, long-run productivity gains from a longer patent term largely outweigh the welfare cost of larger monopoly distortions. Therefore, a patent length much longer than the current status quo of 20 years would induce higher output and consumption in the long run. However, the second trade-off arises from the transitional dynamics generated by an unanticipated implementation of a new patent length. Productivity gains from a longer patent length require upfront R&D investment, which must be financed by reducing consumption in the short run. This happens because technological improvements are slow to achieve, and the presence of development lags drives this mechanism. The short-run consumption loss renders a longer patent length less desirable than its value from a steady-state perspective: I estimate that a patent term of 28 years would maximize the time-zero utility of the representative agent in the absence of policy anticipation.[6] Unanticipated implementation of the new term would increase consumption and output by 0.5% and 1.6% of the status quo levels, respectively.[7] Lastly, I consider what the effect of implementation of the 28-year patent term would be in the presence of anticipation. As documented in the empirical part of the paper, the combined action of policy anticipation, intertemporal development incentives, and powerful spillovers imply that the anticipated implementation of the longer term generates perverse negative effects on output and innovation in the short-run. Consistently, I find that all output gains would be dissipated with an anticipation of 5 months, and that with a 1-year anticipation the economy would suffer an output loss of 1%.

Structure of the paper After relating these findings to the literature, the remainder of the paper is organized as follows. Section 2 illustrates the policy, the identification strategy, and the data. Section 3 shows the main empirical facts. Section 4 empirically investigates the economic mechanisms driving the results and identifies the technology disclosure externality. Section 5 presents the model, and Section 6 describes its structural estimation. Section 7 quantifies the normative trade-offs and studies the welfare impact of changes in patent length under different settings. Section 8 concludes.[8]

1.1 Connection to the literature

The scarcity of empirical evidence on the effect of patent length on innovation and R&D is likely due to the lack of variation in policies. The most recent empirical paper investigating this issue is Budish, Roin and Williams (2015), which documents that in the pharmaceutical sector R&D is disproportionately directed towards treatments with shorter clinical trials, which imply a longer effective protection time. However, the paper cannot disentangle the relative importance of the policy variable, i.e. the finite patent term, vs. firms' preference for projects in which the return from investment is quicker. Other papers examine more comprehensive measures of patent protection strength (Lerner (2009), Moser (2005), Moser and Voena (2012), Sakakibara and Branstetter (2001), Schankerman and Schuett (2017), Moscona (2021)), but not patent length specifically. This paper instead uses one source of variation due to a major policy change and exploits the heterogeneity in its impact across fields.

[6] The baseline estimate is 28 years, but this figure falls to 23 if the model includes the possibility that patents would block subsequent innovations with some probability. Both figures would be longer than the current 20 years status quo.

[7] Sectoral data on productivity (total factor productivity, henceforth TFP) and prices confirm the quantification of welfare trade-offs implied by the model. As in the model, the implied static pass-through of TFP gains into higher consumer welfare is high at around 0.83. In addition, I find that consistent with the model, productivity and welfare gains are slow to achieve.

[8] Appendix A and D describe the data, Appendix B and E report additional empirical results, and Appendix C and F include further theoretical results.

Abrams (2009) is the paper closest to my empirical analysis. It uses the same quasi-experimental variation in patent length to examine its effects on innovation outcomes. However, the two papers differ in terms of the assumed timing of the policy and in terms of econometric specification, thus coming to divergent results. Abrams (2009) *i*) assumes that the policy was unanticipated until its formal signing in December 1994, *ii*) employs a two-periods DiD specification that compares innovation outcomes across fields in a narrow window of data (6, 12, or 24 months) before and after the implementation shock of June 1995, and *iii*) includes in the specification a field-specific linear trend to remove heterogeneous long-run innovation patterns independent from the policy. On the contrary, this paper provides documental evidence that the content of the policy was known by the business sector before the formal implementation of the policy and it examines the impact of both news and implementation shocks. Considering potential anticipation is crucial: If innovation reacts to the news, the level of innovation just before implementation–taken as the baseline of the DiD–would be itself affected by the policy change. Hence, a two-periods DiD specification comparing outcomes just before and after implementation would give confounded estimates of the impact of the patent term change, because it could not separately isolate the news effect. Section 3 shows that this is the case. Therefore, I take as a reference level for the DiD exercise the field-specific level of innovation before the policy *news*, and I employ a multi-period DiD specification that allows to capture quarter-specific effects of the treatment on the outcome variable (in deviations from the pre-news baseline level). The rich specification allows to formally document the absence of correlation between the treatment and heterogeneous trends in innovation across fields (pre-trends) before the policy news.[9] In addition, it allows to flexibly and separately capture the effect of the news shock and of the implementation shock on innovation. Finally, relative to Abrams (2009), this paper also explores the impact of patent length on R&D, and it identifies a new spillover in the data. Theoretically, it develops a structural semi-endogenous growth model with novel features, which is used to investigate the normative implications of patent length.

The model extends Jones (1995) and it models finite patent length along the lines of Lin and Shampine (2018). However, it radically changes the key characteristics of the endogenous R&D block, distinctly modelling research and development activities and embedding a new spillover from the latter to the former. The 2-stages structure of the innovation process is mathematically similar to Comin and Gertler (2006), but the interpretation of the two stages

[9]The formal analysis of pre-trends highlights that the estimation of the regression in deviations from a field-specific linear trend fitted on a narrow window of data–as in Abrams (2009)–violates the parallel trends assumption underlying the DiD exercise. Appendix B.1 provides detailed evidence on how the setting used in the present paper is less prone to obtaining confounded estimates, and it shows how the results of the two papers can be reconciled once Abrams (2009)'s specification is modified to satisfy the parallel trends assumption.

is different. I discuss in section 5 that this difference is crucial for rationalizing the empirical evidence. In addition, differently from the several theoretical papers that try to examine the normative implications of patent length (Nordhaus (1967), Gilbert and Shapiro (1990), Klemperer (1990), Futagamia and Iwaisako (2007)), this paper tightly links the model to causally-identified empirical evidence. This allows to credibly estimate the key structural parameters of the model and to conduct normative analysis.

Finally, the paper contributes to the large empirical and theoretical literature on innovation related spillovers, which include: Knowledge accumulation spillovers, at the core of Romer (1990), and recently re-examined by Bloom et al. (2020) and Aghion and Jaravel (2015); spillovers from basic to applied research (Akcigit, Hanley and Serrano-Velarde (2020)); geographic spillovers (Moretti (2020), Lychagin et al. (2016)); externalities at the inventor level (Bell et al. (2019), Akcigit et al. (2020)); and spillovers in the technological space (Bloom, Schankerman and Van Reenen (2013), Moretti, Steinwender and Van Reenen (2019)). This paper documents and models a novel externality channel from development intensity to subsequent research productivity.

2 Description of policy variation and data

This section describes the key terminology used in the paper, the sources of policy variation employed for identification, and the sources of data.

The *application date* is the day in which an applicant files an application for a patent at the US Patent Office (USPTO). The *grant date* is the day in which the USPTO issues the patent, by publishing a document that grants to the applicant(s) the rights of exclusive economic exploitation of the invention. The *pending period* is the time elapsing between the application date and the grant date. By *statutory patent term*, I mean the time between the first day in which the legal patent term formally starts to elapse and the last day of legal protection. By *effective patent term* or *effective protection time*, I mean the time between the first and the last day in which the patent owner can *in practice* enforce its monopoly power on the invention.

2.1 The TRIPs and its news

The estimation of the impact of effective patent length on innovation and R&D exploits a change in the US patent term from 17 years after the grant date to 20 years after the application date. The adjustment was motivated by the adoption in the US system of the *The Agreement on Trade-Related Aspects of Intellectual Property Rights* (TRIPs), the intellectual property chapter of the Uruguay Round of negotiations of the GATT, spanning from 1986 to 1994. Among other things, the TRIPs implied for the US patent system the adoption of the standard patent length already in place in virtually any other developed country.[10] The formal ratification occurred through the Uruguay Round Agreements Act (URAA) of December 8, 1994 and the new regime became fully effective on June 8, 1995.

Previous literature (Abrams (2009)) assumes that the policy change was fully unanticipated and it motivates this assumption by arguing that the adoption of the Uruguay Round agreements remained uncertain until the URAA was signed, as documented by several newspaper articles. However, several sources and official documents provide evidence that the precise terms of the future patent term change were known long before the formal adoption of the URAA. First, the US business sector was directly involved in the negotiation process since the start of the Uruguay Round, in 1986. Morgese (2009) and Matthews (2002) report that the US Advisory Committee on Trade Policy and Negotiations (ACTPN) included members of the business sector such as the CEOs of IBM and Pfizer, and crucially contributed to shape the position of the US delegation within the TRIPs group. Second, the adjustment of the US patent term was explicitly mentioned for the first time at the end of 1991, in a final draft for the whole Uruguay Round circulated by the GATT Director-General.[11] In addition, Montalvo (1996) examines the steps that finally led to the patent term change included in the TRIPs and reports that *"the first step towards this revolutionary change to domestic patent law occurred in August 1992, when the Advisory Committee on Patent Law Reform issued a report to the Secretary of Commerce recommending adoption of a twenty-year term beginning from the filing date of the first complete United States application"*.[12] The cited report was jointly signed by a number of representatives of the business community and explicitly referred to the patent term change mentioned by the 1991 draft.[13] Finally, early academic articles on law journals examined several aspects of the TRIPs draft (Reichman (1993), Martin and Amster (1994), Doane (1994)) and the potential patent term change was also mentioned by the press.[14] Therefore, the business and the legal communities surely knew about the content of the negotiations and could anticipate the nature of the potential policy changes.

As to the uncertainty about the adoption of the policy, one of the main obstacles toward

[10]The Uruguay Round Agreements Act (URAA) introduced three main changes to US patent law. The first was the patent term change examined in this paper. The second was a no discrimination rule for foreign inventors. The third was the introduction of provisional applications. More generally, the TRIPs broadened the patentable subject matter in developing countries and increased protection for developed countries innovators. Subsection 3.1.3 discusses at length the potential confounding effects arising from these concomitant changes.

[11]*GATT doc. MTN.TNC/W/FA, Draft Final Act Embodying the Results of the Uruguay Round of Multilateral Trade Negotiations, 20/12/91*

[12]*The Implementation of the Uruguay Round Agreement on Trade-Related Aspects of Intellectual Property - the TRIPs Agreement: Hearings on S.2368 and H.R. 4894 before the Subcomm. on Patents, Copyrights and Trademarks of the Senate Judiciary Comm. and the Subcomm. on Intellectual Property and Judicial Administration of the House Judiciary Comm., 103rd Cong., 2d Sess.*

[13]Representatives of IBM, 3M, Procter&Gamble, Motorola, Garret&Dunner, among the others.

[14]*Panel Proposes Patent Changes*, New York Times, Late Edition (East Coast); New York, N.Y. 15 Sep 1992.

the positive conclusion of the negotiations was an unsettled dispute between the US and European countries on agricultural trade. However, most of the disagreement was resolved in November 1992, in a deal informally known as the *Blair House Accord*, which paved the way for the signing of the final agreement on April 15, 1994.[15]

Overall, this narrative evidence suggests that the terms of the policy change were known before the URAA (December 8, 1994), even though uncertainty surrounded its final adoption until it was signed. This implies that the TRIPs potentially induced two distinct shocks: A news shock and an implementation shock. Ignoring the news shock–i.e., assuming no policy anticipation and restricting the effect of the news to be null–can confound the difference-in-difference (DiD) estimates of the impact of policy implementation. Indeed, if the news has an impact, the pre-implementation level of innovation used as a reference baseline for the DiD exercise is itself endogenous to the policy.[16] In the paper, I employ a multi-period DiD specification that *i*) takes the level of innovation just before the *news* as the reference baseline for the DiD, and that *ii*) estimates the quarter-specific effect of the policy. This allows innovation to potentially respond to information available before the URAA was signed, and distinguishes the impact of news from the effect of implementation. This is the most conservative econometric choice. If innovation does not react to news until the formal vote of the Congress, the multi-period specification would simply estimate a null effect before the URAA. Otherwise, the impact of the news shock would be reflected in the quarterly DiD estimates. Finally, I take as a reference date for policy news November 1992, the date of the *Blair House Accord*.[17]

2.2 Variation in patent length across technical fields

Hereafter, a *technical field* or *technological field* is defined as one of the 621 4-digit International Patent Classification (IPC) subclasses in my sample.[18] The identification of the impact of changes in patent length on innovation outcomes exploits two sources of variation in the data: *i*) Cross-sectional variation of the average pending period by technical field, and *ii*) policy-induced time variation in the statutory patent term. As to the first source of variation,

[15]This is reported by Morgese (2009) and at `https://en.wikipedia.org/wiki/Uruguay_Round`, where it reads: *"The round was supposed to end in December 1990, but the US and EU disagreed on how to reform agricultural trade and decided to extend the talks. Finally, In November 1992, the US and EU settled most of their differences in a deal known informally as "the Blair House accord", and on 15 April 1994, the deal was signed* [...]*"*

[16]Appendix B.1 shows that this is the case in Abrams (2009) and that a sample extension would mitigate the concerns.

[17]All the results are unaffected by an additional anticipation of one or two quarters.

[18]For example, the 4-digit IPC "A23D" is "Edible Oils or Fats, e.g. Margarines Shortenings, Cooking Oils". It is included in the 3-digit IPC "A23", "Food or Foodstuffs; Their Treatment, not covered by other classes" and in the 1-digit IPC "A", "Human Necessities". It further includes two 8-digit IPCs: "A23D 7/00", "Edible oil or fat compositions containing an aqueous phase, e.g. margarines", and "A23D 9/00", "Other edible oils or fats, e.g. shortenings, cooking oils".

it partly originates from the fact that patents classified in different fields are examined by different technical units, which differ in terms of congestion–due to, e.g., staffing issues or intensity of foreign filings–and the difficulty of technical examination.[19] The second source of variation originates from the fact that the URAA changed the statutory patent term from 17 to 20 years *and* moved the reference for determining patent expiration from the grant date to the application date.

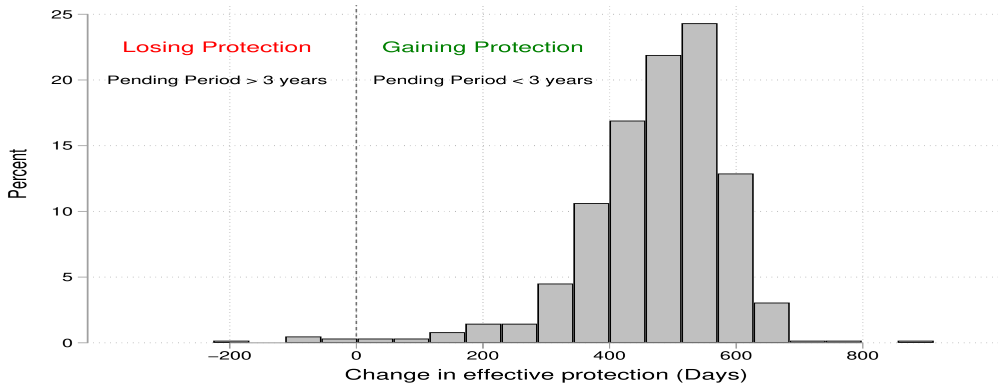
Because applicants can fully enforce their monopoly over the invention only after the grant date, the interaction of the legal change in the statutory term–from 17 years after grant to 20 years after application–*and* of technical field-level heterogeneity in the average pending period implied a policy-induced change in average effective protection time that varied across technical fields.[20] Fields with an average pending period shorter than 3 years obtained an average increase in effective patent length, and the increase was longer the shorter the average pendency duration. Figure 1 shows the distribution of the change in effective protection time, the treatment variable, across technical fields before the policy news. Most of the fields experienced an expected average increase in protection but heterogeneity was considerable, and a few fields faced an expected reduction.

Figure 2 further illustrates with an example how the policy was changing effective patent length. The top panel represents how effective protection time varied in a technical field where the average pending period was longer than 3 years before the policy change. The *pre*-implementation average effective protection time was equal to the statutory term of 17 years from the grant date, represented by the blue pre arrow. After the policy implementation, however, the average effective protection time covered the time between the grant date and the end of the statutory term of 20 years from application, i.e., the red post arrow in Figure 2 top panel. Because the average pending period was longer than 3 years–and assuming that it did not respond to the policy, which I show to be the case in subsection 3.1–the effective patent term was shorter after the implementation of the policy than before. Therefore, the TRIPs induced an expected average patent length reduction in these fields. The bottom panel of Figure 2 shows that the situation is opposite for technical fields where the average pending

[19]One may be concerned with the endogeneity of the pending period to innovation outcomes. Indeed, Lemus and Marshall (2018) argue that the pending period is partly endogenous, because applicants can also influence it by strategically choosing to delay or accelerate their replies to the inquiries of the patent office. However, Table A.4 of Appendix A.2 shows that the average pending period by field is not correlated with the growth rate of patenting before the policy news, but it is correlated with proxies of congestion, such as the share of foreign second filings, and of examination difficulty, i.e. the average pending period at the European Patent Office. In addition, subsections 3.1.2 and 3.1.3 present several analyses showing that the cross-field variation exploited for identification is exogenous, allowing a causal interpretation of the results.

[20]Patents grant an exclusive monopoly power on an invention and legal protections against its violations from the grant date or from the publication of the application. Under current regulation, publication occurs after 18 months from the application date, but before 2000 patents were published only at grant. Therefore, monopoly power was enforceable only from the grant date onward.

Figure 1: **Distribution of the expected change in effective protection time**



The histogram shows the distribution of the treatment variable (average change in effective protection time) across technical fiels (4-digit IPC classes). The treatment is computed, by technical field, subtracting the average pending time of patent applications first filed at USPTO and obtaining the grant between January 1^{st} 1990 and May 31^{st}, 1992 to 365×3 days, the statutory change. Details on the construction are in subsection 2.3.1.

period was shorter than 3 years before the policy and where, therefore, the TRIPs induced a longer effective protection time.

2.3 Data

Empirical evidence is provided at three levels of analysis: At the technical field level, at the firm level, and at the NAICS 6-digit industry level.[21] I mostly rely on PATSTAT for the technical field-level analysis, on the NBER Patent database and COMPUSTAT for the firm-level evidence, and on the NBER CES manufacturing database for the sectoral results. The next two subsections describe the construction of the continuous treatment variable by technical field and at the firm-level. Appendix A reports summary statistics and Appendix D contains the construction details for all the other variables.

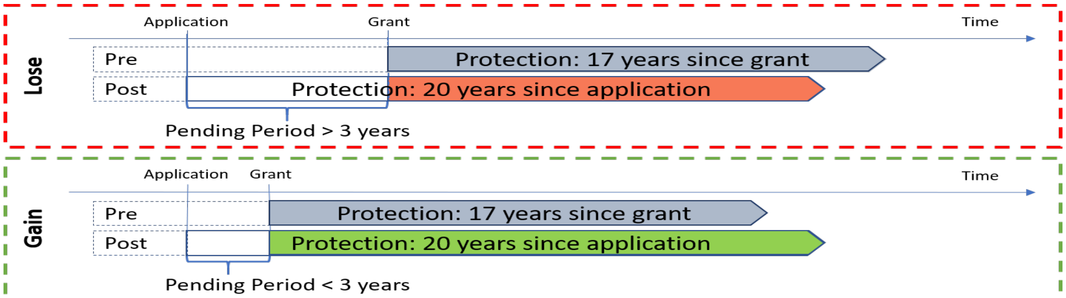
2.3.1 Expected change in protection time - Technical field-level

The continuous treatment variable by technical field is the policy-induced change in effective protection time faced by applicants filing a patent classified in a specific technological field. As explained in subsection 2.2, it can be measured as the difference, in number of days, between 3 years–the policy-induced statutory change–and the average pending period by field. The latter is measured as the average number of days between the grant date and the application date across all patents classified in a specific field.[22] To compute such measure, I use patent data from PATSTAT and focus on granted patents *i*) whose earliest application

[21] All the details and all the results of the industry-level analysis are not reported in the paper, but can be found in Appendix B.5

[22] If a patent covers multiple technical fields (4-digit IPCs), I use that patent for computing the average pending period for all those fields.

Figure 2: **The change in patent length induced by the TRIPs**



The graph shows the policy-related change in the expected effective patent length in a technical field losing protection (top panel) and in another gaining protection (bottom panel). Monopoly is enforceable only starting from the grant date. The key difference between the fields is the average pending period, i.e. the average time between the application and the grant. In the losing field, this is longer than 3 years. Therefore, the old-regime effective protection time of 17 years after the grant (blue pre arrow) exceeds the effective protection time of 20 years from application date (new statutory term) *minus* the pending period (>3 years) (red post arrow), effective with the new regime after June 8, 1995. In the field gaining, the opposite occurs, because the average pending period is shorter than 3 years.

is filed at the USPTO and *ii*) whose grant date is between January 1^{st}, 1990 and May 31^{st}, 1992. The first restriction is imposed to capture the examination time of a novel US patent and not of those inventions already examined at the USPTO or elsewhere (and applying for legal protection in the US as well).[23] The second restriction is imposed to estimate the average pending period in a window of time unaffected by the policy change and recent enough to be representative of applicants' expectations. Subsection 3.1.3 performs a number of checks to show that the treatment is unlikely to be endogenous, and that it has good predictive power on the effective change in protection time experienced by applicants after the policy change.[24]

2.3.2 Expected change in protection time - Firm-level

The expected treatment at the firm level is a firm-specific weighted average of the technical field-level treatments. Weights are built using the NBER-COMPUSTAT matched dataset and computed, for each firm, as the share of patents granted in a given field over the period 1971-1991. These weights should represent the average ex-ante exposure of firms to technical fields that are differently impacted by the policy change. I stop at 1991 to exclude the year of the policy news.

[23]Results, not shown, are robust when including applications at the USPTO where the application date is subsequent to the priority date of the patent.

[24]First, I show that the effective pending period experienced by applicants after the policy change does not react to the treatment variable built using the ex-ante pending period. Second, I perform an analysis where the treatment variable employed in the main DiD regression is built using the realized, ex-post pending period. In this case, the latter regressor is instrumented by the treatment computed using the ex-ante average pending period, described in this subsection. The instrument is strong and all the results are identical to the OLS specification that employs the treatment based on the ex-ante pending period. Lastly, I check that the raw correlation between the ex-ante and the ex-post average pending period is generally above 0.6.

3 Estimating the effect of changes in patent length

This section shows the estimation results of the impact of effective patent length on innovation and R&D effort. Subsection 3.1 presents evidence by technical field, subsection 3.2 focuses on firm-level evidence, and subsection 3.3 summarizes the key findings.

3.1 Analysis by technical field

3.1.1 Specification of the DiD regressions

The main analysis employs a quarterly panel dataset at the technical field-level and the baseline difference-in-difference (DiD) regression is

$$Y_{j,t} = \alpha_j + \sum_{k=1985Q1}^{2000Q4} \gamma_k \mathbf{1}_{(t=k)} + \sum_{k=1985Q1}^{2000Q4} \beta_k \mathbf{1}_{(t=k)} T_j + \varepsilon_{j,t} \tag{1}$$

where $Y_{j,t}$ is the technical field-j and quarter-t dependent variable, α_j are technical field fixed effects, $\mathbf{1}_{(t=k)}$ are quarter-specific dummy variables, with the γ_k coefficients capturing the effect of any quarter-specific factor common to all technical fields and unrelated to treatment. T_j is the technical field-specific treatment variable, i.e. the expected average change in effective protection time described in subsection 2.3.1, and $\varepsilon_{j,t}$ is the error term. The β_k's are the coefficients of interest, which capture the quarter-specific marginal effect of one more day of expected protection time on $Y_{j,t}$. The quarter-specific dummy referring to the quarter before the policy news, i.e. 1992Q3, is always omitted. Therefore, the β_k's should be interpreted as deviations of the outcome from its 1992Q3 baseline level.

Specifically, β_k $k \leq 1992Q2$ capture the correlation of the treatment variable with the outcome of interest *before* the treatment should have any effects. The latter set of coefficients being close to zero provides evidence that the treatment is not related to unobserved, preexisting heterogeneous trends of the outcome across technological fields, supporting the exogeneity assumption. In the absence of pre-trends, any candidate confounders must take the form of an unobserved shock *contemporaneous* to treatment. Subsection 3.1.3 performs several additional analyses suggesting the absence of these confounders, and supporting a causal interpretation of the results. The post-news, pre-implementation coefficients β_k $1992Q4 \leq k \leq 1995Q2$ capture the marginal impact of a future 1-day increase in patent length $T_j = 1$. Finally, β_k $k \geq 1995Q2$ capture the reduced form impact of the implementation of a 1-day anticipated increase in T_j.

3.1.2 Results

The first outcome of interest is $P_{j,t}$, i.e. the number of patents whose application is filed in quarter-t and technical field-j and that are subsequently granted.[25] Importantly, patents are counted in the quarter they are applied for and not in the quarter they are granted, thus removing any effect of pending period adjustments or examination backlogs on the time series of the outcome variable by field. The preferred specification features a dependent variable in levels and not, e.g., in natural logarithms. First, because the implications of endogenous growth theory suggest that patent length should affect the *level* of aggregate innovation or productivity and not its *growth* rate.[26] Second, because patenting by technical field in the later 80's and early 90's seems to be more consistent with arithmetic growth than with exponential growth.[27] However, Appendix B.2.13 shows that results are robust to *i*) transformations of the outcome variable, including logs; *ii*) sample restrictions aimed at reducing skewedness of the patenting variable across fields; and *iii*) estimation of a negative binomial model for count data.

Figure 3 plots the β_k coefficients - and their 95% confidence bands, with standard errors clustered by technical field. The first result is that the β_k $k \leq 1992Q2$ coefficients are very close to zero and, therefore, the parallel trends assumption seems to be satisfied. As a result, endogeneity problems caused by preexisting unobservables correlated with the treatment seem to be ruled out.[28] I postpone to subsection 3.1.3 additional investigations of potential confounders contemporaneous to the treatment.

The second result is that the coefficients estimated between the news date 1992Q4 and the

[25] I use granted patents and not all patent applications because granted patents are thought to capture actual innovation better than applications, and because data on all US applications are available only after year 2000. Therefore, data on pending or abandoned patent applications filed during the sample period may be partial or unavailable. In addition, while previous literature found a negative correlation between pending time and quality or grant probability at the patent-level, I find a mildly positive correlation between the average pending period and the average patent quality–as measured by the average forward citations per patent–at the technical field level.

[26] An extreme implication of assuming that patent length affects the growth rate of innovation rather than the level would be that a country with a patent term marginally longer than the one of another all-else identical country would grow at a persistently higher rate. As a consequence, in the limit, the relative productivity advantage of the country with marginally longer patent length would be infinite.

[27] This observation comes from fitting a quadratic time trend with field fixed effects on quarterly patenting data by field, in levels and in natural logarithms. The coefficient of the linear term is positive and statistically significant in both cases, but the coefficient of the quadratic term is negative and statistically different from 0 only in the log case.

[28] One of the endogeneity concerns could be caused by the interplay of the pending period and the point of the life-cycle of different fields at the moment of the policy. For example, fields where innovation was growing very fast ex-ante may have had both a longer average pending period–because of congestion–and a slower growth rate ex-post–because of a natural decay unrelated to the policy. This would result in lower innovation in fields with shorter treatment, without such lower innovation being *caused* by the treatment. Crucially, however, this would also be reflected by pre-trends coefficients being different from 0, and capturing the ex-ante correlation between the treatment and heterogeneous innovation pace.

implementation date 1995Q2 are negative: The news of a future patent term increase induces a fall in innovation. The magnitude of the estimated effects is small right after news, but gradually grows as implementation gets closer. Consider a 1-month (30-days) increase of patent length. One year after news and two years before implementation, the policy change is related to -0.5 quarterly patents per technical field, which is approximately -1.5% of the baseline level of 1992Q3. The effect becomes much bigger two years after news and approximately one year before implementation: The same policy generates -1.4 quarterly patents per technical field, i.e. -4.4% of the baseline. Excluding the pre-implementation quarters 1995Q1 and 1995Q2, the average impact during 1992Q4-1994Q4 is -0.93 quarterly patents per technical field, which is approximately -3% of the baseline level in 1992Q3. The effects in 1995Q1 and 1995Q2 are excluded from the computation of the average effect because the policy–officially signed in December 1994–allowed applicants to choose whichever policy regime was most favorable until the final implementation of the new regime in June 1995. Therefore, there is strong evidence of bunching there. A 30-days longer future protection decreased patenting by around 25% in 1995Q2, which is 4 times bigger than the effect estimated for 1995Q1 and 8 times bigger than the average effect.[29]

The third result is that the β_k's estimates remain negative even after policy implementation. Therefore, the reduced-form effect of the implementation of a longer patent length on innovation is negative. The magnitude of the effect is also strong. On average, the implementation of a 1-month (30-days) patent term change generates a drop of -3.7 quarterly patents by technical field, which is around 12% of the baseline. I will refer to this effect as post-implementation *persistence*.[30]

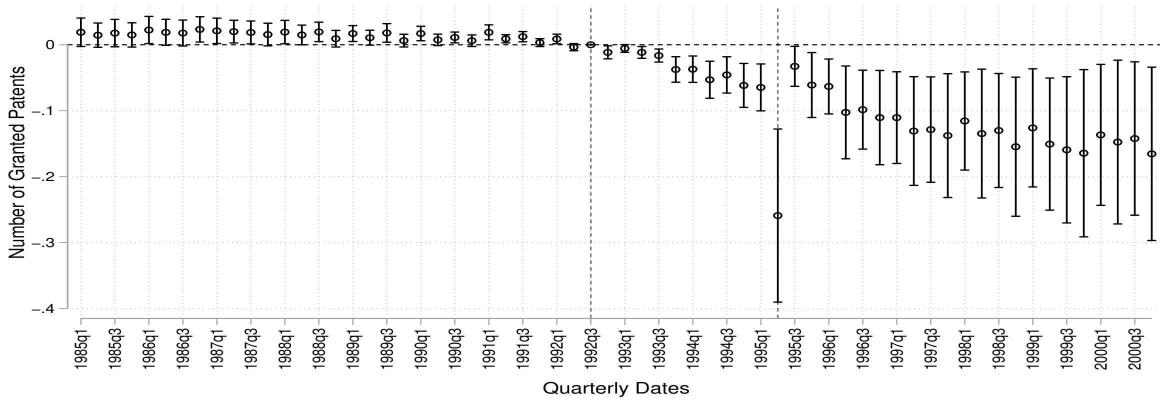
Quality-adjusted measures Results are qualitatively similar using alternative measures of innovation that capture the scientific value of patents–weighting them by the number of forward citations received within 5 years from grant–or the private economic value of patents.[31] Figure 4 reports the β_k coefficients of specification (1) taking as the dependent variable citations-weighted patents. The dynamics of the effects are analogous, but the magnitude is stronger. Excluding the pre-implementation quarters 1995Q1 and 1995Q2, a 1-month (30-days) future increase in patent length generates an average fall of 10.5 citations-weighted patents per quarter and field, i.e., around 6.6% of the 1992Q3 baseline. After the implemen-

[29] Alternative specifications reported in Appendix B.2.13–e.g., a negative binomial model for count data or a linear specification with dependent variable in natural logarithms–deliver similar results.

[30] In an heterogeneity analysis–whose results are not reported–I discretize T_j by quartile and I run a DiD regression comparing neighbor quartiles. The estimated coefficients are similar across quartiles but for the case of the third and the fourth, where the estimates are closer to 0.

[31] I build a quarterly measure of private economic value of patents by technical field aggregating the individual-patent estimates by Kogan et al. (2017). The results for this outcome are reported in Appendix B.2.1. Alternatively, I present results for claims-weighted patents in Appendix B.2.2

Figure 3: **Marginal effect of 1 more day of protection on granted patents**



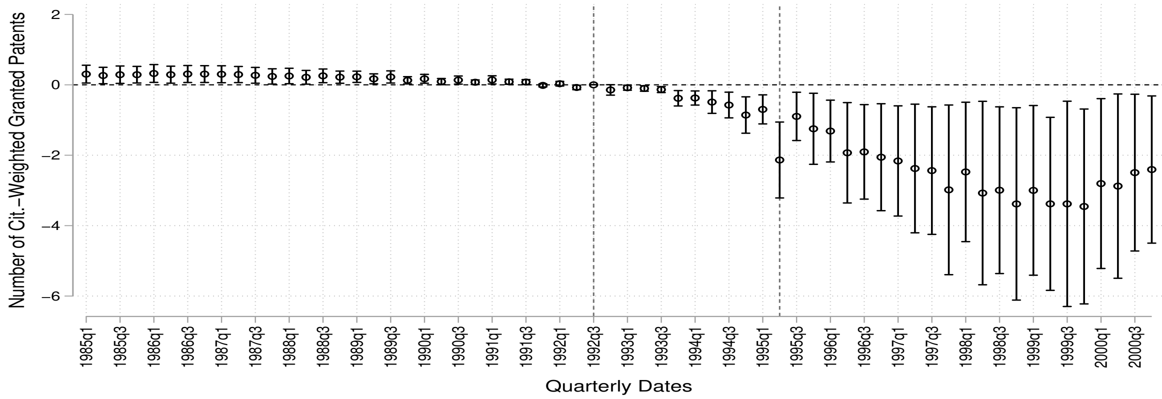
The plot shows the β_k coefficients of specification (1) having as dependent variable quarter-t and field-j number of granted patents. Standard errors are clustered by technical field and 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

tation of the policy, the average quarterly drop amounts to 75 citations-weighted patents, i.e., 47% of the baseline.

R&D effort Next, I investigate the impact of the policy change on R&D effort, which is relevant to support the argument that the policy-driven reaction of patenting outcomes originates from actual changes in innovative activity rather than from mere changes in patenting strategies. A direct measure of R&D expenditure at the technical field-level is not available, so I proxy R&D by the number of inventors that are listed on at least one patent filed in a given quarter and technical field, avoiding multiple counting of inventors listed on more than one patent.[32] Indeed, an important item in firms' R&D expenditure is the wage bill of inventors: While I cannot observe the wages, I can count the number of people having actively worked on a patent of a given field and quarter, which should give a sense of how the research workforce employed by firms evolves over time. This approach is also employed by other papers in the innovation literature. Appendix B.2.3 (Figure B.7) reports the results of the dynamic DiD specification (1) having as a dependent variable the number of unique inventors. The response is qualitatively similar to the one observed for innovation variables and, quantitatively, the pre-implementation β_k's estimates imply that a 30-days future increase of the patent term generates an average drop of 5% of inventors per-quarter and technical field. Given the potential limitations of the field-specific measure of R&D effort, subsection 3.2 also provides firm-level evidence that a direct measure of R&D expenditure, taken from

[32]To build this inventors count, I use the STAN harmonized inventor's identifiers from the EPO Worldwide Bibliographic Database reported in PATSTAT.

Figure 4: **Marginal effect of 1 more day of protection on citations-weighted patents**



The plot shows the β_k coefficients of specification (1) having as dependent variable quarter-t and field-j 5-years citations-weighted patents. Standard errors are clustered by technical field and 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

COMPUSTAT, moves consistently with previous evidence.

Post-implementation persistence The economic mechanisms underlying the reduced-from DiD results remain an open point, which section 4 will empirically investigate. However, as a preliminary step, I try to investigate whether the negative post-implementation impact of the policy gradually weakens over time. I estimate (1) on a sample extended to 2010Q4, and I find that the $\hat{\beta}_k$ coefficients tend to revert back to 0 as the policy implementation gets farther in time.[33] Given the constraints posed by the policy and by the econometric framework, it is difficult to draw definitive conclusions on this aspect.[34] However, I interpret this as suggestive evidence that the effect is transitory and that it features mean-reversion toward zero.

3.1.3 Addressing endogeneity concerns

The inclusion of field fixed effects in specification (1) should capture all field-specific time-invariant unobservable confounders. In addition, the absence of pre-trends in the data suggests that the treatment is unrelated to omitted factors that generate ex-ante differential growth paths across fields, potentially driving differential innovation ex-post. The pending period and the treatment are correlated with the *level* of innovation across fields, but not with the ex-ante *trend*.[35] Therefore, potential confounders must be time-varying factors that

[33]Results are reported in Appendix B.2.5. The linear specification of Appendix E.1.5 with the outcome variable in natural logs further supports this hypothesis.

[34]The farther away data go from the policy implementation, the more likely it is that some unobserved factors correlated with the treatment undermine a causal interpretation of the results.

[35]Heterogeneity in the pending period across technical fields mainly originates from the different degree of complexity of inventions and, hence, of the examination process. Also, patents in different fields are examined

correlate with the treatment and that are *contemporaneous* to the policy change.[36]

The first concern is that applicants' incentives to respond more quickly to the patent office during the examination process may have changed endogenously with the policy.[37] This would generate an innovation response correlated with this reaction rather than with the treatment itself. To address this issue, I first run specification (1) having as the dependent variable the average pending period in quarter-t and field-j. Appendix B.2.4 (Figure B.8) reports the results. I find that: *i*) The treatment variable is not related to a different trend of the average pending period before the policy change and that *ii*) the pending period does not show substantial *level*- or *trend*-discontinuities around the treatment date. Therefore, applicants' incentives to be more responsive to the patent office's inquiries do not seem to be differentially correlated with the ex-ante pending period across fields. As a second check, Appendix B.2.7 also shows the results of an IV specification where the treatment variable is the policy-induced change in protection time actually experienced by applicants in every quarter and field. Such treatment variable is computed along the lines of subsection 2.3.1 but replacing the ex-ante pending period with the actual quarter- and field-specific pending period observed in the data. The latter is instrumented with the T_j of specification (1), interacted with quarterly dummy variables. The use of the actual pending period should improve the representativeness and the accuracy of the main regressor, and the use of variation induced by the ex-ante treatment should mitigate endogeneity concerns. Results are almost identical to subsection 3.1.2, and the first-stage regressions confirm that the ex-ante treatment is a statistically strong predictor of the ex-post effective change in protection time.[38,39]

A second relevant concern comes from the fact that the TRIPs, beyond homogenizing the US patent term to the rest of the world, strengthened intellectual property protection in many developing countries, favoring the access of American and European innovators. If the benefits from these changes were heterogeneous across fields and somehow related to the ex-ante pending period by field, then the evidence of subsection 3.1.2 would fail to cap-

by different technical units at the USPTO, which may face staffing issues that are persistent over time. Table A.4 shows the correlation of the average pending period with several field-specific variables, finding support for heterogeneity being related to congestion and technical examination difficulty.

[36] Appendix B.2.10 also shows that results are identical when including in specification (1) the interactions of quarter- and 3-digit IPC class-specific fixed effects, which remove any confounders specific to the broader 3-digit field and quarter. This further reduces the set of potential confounders to care about.

[37] Lemus and Marshall (2018) find that in the pharmaceutical sector applicants' responsiveness increased in response to the TRIPs, making the pending period at least in part endogenous. Appendix B.2.14 shows that the results of subsection 3.1.2 are fully consistent when excluding technical fields related to the pharmaceutical and biotechnologies.

[38] The raw correlation between T_j and its ex-post quarterly realization is generally between 0.5 and 0.6.

[39] Appendix B.2.8 shows that the response of innovation outcomes is stronger in magnitude in fields where the policy-induced treatment could be inferred ex-ante with more precision, i.e., where the standard deviation of the ex-ante average pending period is smaller. This supports that the observed response is generated by the treatment of interest and not by other factors.

ture a causal effect of patent term on innovation.[40] To address this concern, I exploit the fact that European countries had similar benefits as the US from the TRIPs-implied strengthening of IP worldwide, and that the only crucial difference for patent policy between the two regions was the US patent term adjustment, which remained unchanged in Europe.[41] Despite the fact that the US patent term change may have affected those European firms also patenting in the US, aggregate innovation in Europe should be relatively less sensitive to US policies than US innovation. In practice, I build the same quarterly measures of patenting by technical field for patents filed at the European Patent Office, and I run a triple-difference regression that implicitly removes from the US patenting variable the technical field-specific baseline effect observed in Europe. This triple-difference specification absorbs all technical field-specific time-varying unobserved factors that are common to the US and Europe and that happen independently from the change in the patent term. Figure 5 plots the triple difference coefficients and shows that the findings emerging from the baseline specification are largely unaffected. The coefficients represent the marginal effect of one more day of patent protection on US field-specific quarterly patenting relative to EPO, both in deviations from the respective 1992Q3 baseline level.[42]

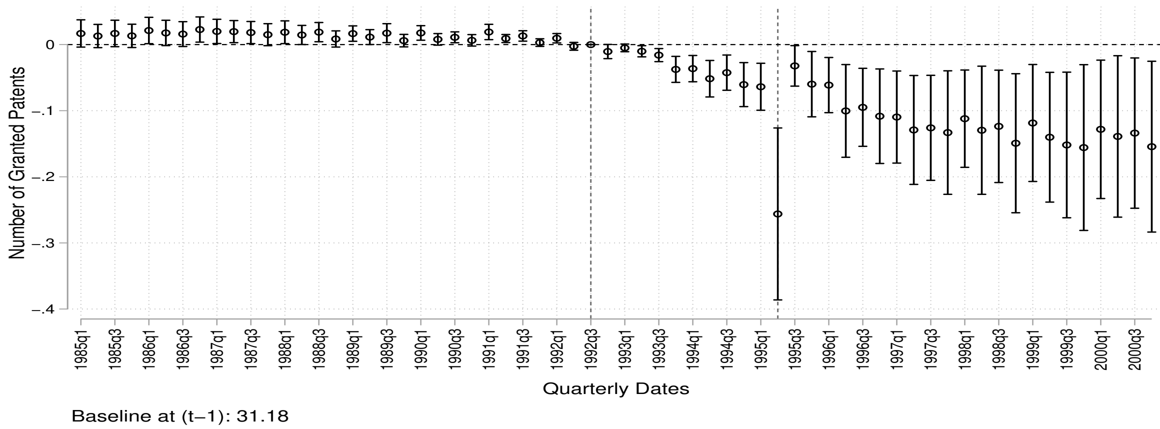
Motivated by both the potential manipulation of the pending period by applicants and the potential presence of additional unobserved shocks contemporaneous to the policy event, I finally implement an IV analysis where the US treatment variable is instrumented by *i*) the technical field-specific share of second filings by foreign applicants before the news, and *ii*) the technical field-specific pending period at the European Patent Office. The first external instrument should capture pre-existing congestion of the examination offices unrelated

[40]Kyle and McGahan (2012) argue that US pharmaceutical firms increased their innovation investment after the TRIPs, mostly due to their ability to enforce patents in new developing markets. Appendix B.2.14 shows that the results of subsection 3.1.2 are fully consistent when excluding technical fields related to the pharmaceutical and biotechnologies.

[41]The URAA introduced three main changes to US patent law. The first was the patent term change examined in this paper. In European countries the patent term was already 20 years since the application date. The second change was a no discrimination rule for foreign inventors, which held for European countries too. The third change was the introduction of provisional applications, i.e. preliminary applications that could be filed at the USPTO before filing the actual application for examination. This normative change did not have a counterpart at the European Patent Office. If this induced a substitution from normal applications to provisional ones and this shift were correlated with the treatment variable, triple-diff results would potentially be bias too. Unfortunately, I cannot directly check this correlation in the data, because PATSTAT does not allow to assign technical fields classification to provisional applications. However, a specific constraint posed on provisional applications rules out any major concerns, because the latter type of applications must be turned into an actual application within 12 months, otherwise they are considered definitely abandoned. Therefore, should the above-mentioned substitution happen, it would simply induce a slight re-timing of the observed post-implementation persistence rather than an actual bias to the results.

[42]Appendix B.2.10 also shows that results are identical when including in specification (1) the interactions of quarter- and 3-digit IPC class-specific fixed effects. The latter should remove the impact of any effects of the TRIPs that is unrelated to the patent term change and that is specific to the broader 3-digit field and quarter.

Figure 5: **Marginal effect of 1 more day of protection on granted patents - Triple difference specification**



The plot shows the β_k coefficients of specification $P_{r,j,t} = \psi_r + \alpha_j + \kappa\mathbf{1}_{(r=US)}T_j + \sum_k \gamma_k \mathbf{1}_{(t=k)} + \sum_k \eta_k \mathbf{1}_{(t=k)}\mathbf{1}_{(r=US)} + \sum_k \theta_k \mathbf{1}_{(t=k)}T_j + \sum_k \beta_k \mathbf{1}_{(t=k)}T_j\mathbf{1}_{(r=US)} + \varepsilon_{j,t}$, $k = 1985Q1, ..., 2000Q4$. The dependent variable is region-r's, quarter-t, and field-j number of granted patents. T_j is the field-specific treatment described in Section 2.3.1, and $\mathbf{1}_{(r=US)}$ is a dummy variable taking value 1 if the region considered is the US. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

to manipulation by domestic applicants. The second external instrument should proxy the technical difficulty of examination of patents in a specific class.[43] Results are reported in Appendix B.2.9 and are fully consistent with those of subsection 3.1.2.

3.1.4 Additional evidence by technical field

A first additional concern is that the observed reaction of patenting may come from a change in the patenting strategy of innovators rather than being an actual innovation response. For example, in fields losing protection time, applicants may try to break down patents into smaller ones or file patents with lower quality. However, this does not seem to be the case. Subsection 3.1.2 already showed that quality-adjusted patent measures such as citations-weighted patents respond in a way that is similar to the simple patent count. This is true also for claims-weighted patents or their economic value (Appendices B.2.2 and B.2.1). Appendices E.1.1, E.1.2, and E.1.3 also show that the average number of citations per patent mildly increases in fields where the policy generates more innovation, and that the number of claims per patent, their average originality, and generality do not change in response to the policy.

[43]Both instruments are computed using patents granted before the policy news in 1992Q4, in order to minimize potential endogeneity concerns. The strategy is valid as long as time-varying, field-specific unobservable confounders–correlated with both US patenting and the ex-ante US pending period for first filings–are orthogonal to pending period at the EPO and congestion at the USPTO generated by foreign filings.

A second additional concern is that the estimated effects are not specifically driven by the maximum patent term change but by other factors related to the average pending period. To address this concern, I try to show that the response of innovation is stronger in fields that are expected to be more sensitive to the maximum patent term. In fact, technological fields show substantial heterogeneity in the rate of payment of patent maintenance fees, which are due at 3.5, 7.5 and 11.5 years after the grant to keep patent protection active. Fields with a higher share of patents for which the last fee is paid–thus extending protection to the maximum term–are expected to be more sensitive to the maximum patent term changes. Therefore, I run a triple-difference specification where the treatment variable T_j is interacted with the technical field-specific percentage of patents for which the last maintenance fee, at 11.5 years, is paid. If the estimated negative effects from the simple DiD are stronger in magnitude in more sensitive fields, the triple-difference coefficient should be negative. Appendix B.2.11 shows that this is the case, suggesting that the response of innovation outcomes is precisely related to the policy-induced patent term change. In addition, Appendix B.2.8 shows that the magnitude of the response is stronger in fields where the inference on the average pending period by field could be more precise, i.e., where the within-field heterogeneity in the pending period was smaller ex-ante.

Finally, I check that the effects are not driven by the specific time window when the policy change is taking place. I carry out a placebo analysis using 1982Q4 rather than 1992Q4 as the treatment date. In this case, the estimated policy effect is null.[44]

3.2 Firm-level evidence

This subsection studies the effect of the patent term change on firm-level patenting and a *direct measure* R&D expenditure, using a sample of COMPUSTAT firms matched to patents in the NBER Patent database. The specification of the firm-level regression is

$$\begin{aligned} ln(1+Y_{i,t}) = \alpha_i + \sum_j \eta_{1,j} sic_j + \sum_j \sum_{k=1987}^{2000} \eta_{2,j,k} sic_j \mathbf{1}_{(t=k)} + \sum_{age \in A} \delta_{age} + \\ + \theta ln(1+S_{i,t}) + \sum_{k=1987}^{2000} \gamma_k \mathbf{1}_{(t=k)} + \sum_{k=1987}^{2000} \beta_k \mathbf{1}_{(t=k)} T_i + \varepsilon_{i,t} \end{aligned} \tag{2}$$

where T_i is the firm-level treatment variable described in subsection 2.3.2, α_i are firm fixed effects, sic_j are 2-digit SIC industry fixed effects that take value 1 if firm i is active in SIC j and 0 otherwise, δ_{age} are firm-age fixed effects, $S_{i,t}$ are firm i's sales in year t, and $\mathbf{1}_{(t=k)}$ are yearly dummies. The 1991 dummy is omitted. The coefficients of interest are the β_k, and standard errors are clustered by 2-digit SIC industries. Specification (2) is run for several dependent

[44] Appendix B.2.12 shows the results

variables. I report in the paper the results for $R\&D_{i,t}$, i.e. R&D expenditures of firm i in year t (Figure 6).[45] Appendix B.3.2 reports the results for granted patents (Figure B.26); the number of citations-weighted granted patents (Figure B.27); and the private economic value of patents (Figure B.28).

Firm-level evidence on innovation outcomes and direct R&D expenditure measures is fully consistent with the patterns documented in subsection 3.1.2. On average, a 30-days future increase of patent length decreases yearly patenting by 2.6% at the firm level before implementation. This estimate is close to the field-level effect. After the implementation, the impact of the same policy change implies a decrease of yearly firm-level patenting of 2.1%. Appendix B.3.4 shows the results obtained by estimating a negative binomial model with patent count as outcome variable. Evidence is fully consistent with the linear specification (2).[46] The findings are qualitatively similar when studying citations-weighted patents and private economic value of patents.[47] Figure 6 shows the effect on firms' R&D expenditure. The news of a future 30-days patent term increase implies an average reduction of yearly R&D by 1.9% before implementation of the new policy. The effect remains negative but smaller in size (-1.4%) after implementation.[48] This confirms that the policy change generated an actual response of R&D effort, consistent with the observed changes in innovation outcomes.

Specification (2) exploits *cross*-firms variation in expected patent protection, determined by the interaction of firms' ex-ante technological exposure with the field-specific average patent term change. In a complementary analysis, I exploit variation of patent term *across* technical fields *within* the firm. I build a yearly panel dataset where the cross-sectional unit is a firm $\times$ technical field, which makes it possible to analyze how innovation activity is reallocated *within* the firm, *across* technical fields experiencing heterogeneous patent term changes. Controlling for the ex-ante firm's technological position, I find that all innovation outcomes respond to the policy in a way that is consistent with the evidence documented so far.[49]

3.3 Key takeaways

To sum up, I have so far established two facts. First, news of a future patent term increase induces a slowdown in innovation and R&D effort before policy implementation. Second, a

[45] R&D is the variable `xrd` in COMPUSTAT.

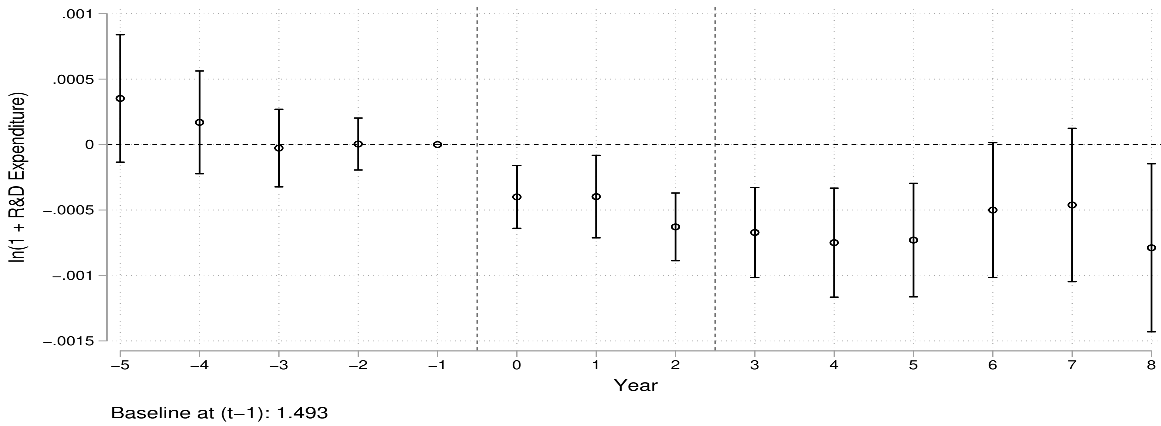
[46] The estimation of a negative binomial model is only possible for patents count, as the other firm-level outcomes are continuous.

[47] Results available in Appendices B.3.2 and B.3.3.

[48] The magnitudes are smaller than in the technical field analysis. The reason can be twofold. First, aggregate innovation is affected by entry but firm-level results are just based on firms already innovating before 1991. Second, firms' income statement R&D better captures investment flows than patent-read R&D, which partly reflects past effort.

[49] I refer to Appendix B.3.5 for all the details and the results of this analysis.

Figure 6: **Marginal effect of 1 more day of protection on firm-level R&D**



The plot shows the β_k coefficients of regression (2) having as dependent variable $ln(1+R\&D_{i,t})$, where $R\&D_{i,t}$ is year-t and firm-i R&D expenditure. Standard errors are clustered by 2-digit SIC industry. 95% confidence bands are plotted. The first vertical line lies just before the news year (1992) and the second vertical line lies just before the implementation year (1995).

negative reduced-form relationship between a positive patent term change and innovation and R&D outcomes continues even after policy implementation. The data suggest that the second negative effect gradually weakens over time.

4 Investigation of the economic mechanisms

This section empirically investigates the mechanisms underlying the results. First, subsection 4.1 discusses several forces that may drive the post-implementation negative effect. Subsection 4.2 shows empirical evidence of the presence of a technology disclosure spillover whereby the drop in innovation at news of a future patent term extension negatively impacts subsequent innovative activity after policy implementation. Subsection 4.3 investigates alternative candidate drivers. In subsection 4.4, I interpret the reaction of innovation and R&D to the news shock, building on the results of subsections 4.2 and 4.3. Moreover 4.4 estimates the elasticity of future innovation to current innovation shocks. Finally, subsection 4.5 provides a unified narrative for the empirical facts of section 3, highlighting the key role played by policy anticipation and technological spillovers.

4.1 Post-implementation: Competing interpretations

I investigate several economic mechanisms that may rationalize the negative effect of patent length on innovation observed after policy implementation. The data support that the main driver of the observed relationship is a technology disclosure externality that directly links the fall in innovation after policy implementation to the drop in innovative activity induced by news of a future patent term extension. New inventions are often inspired by recent technical advances and technologically rely on them. In fields where news of a future patent term increase generates a drop in innovation, there is a lower flow of recent improvements on which innovators can build to start new technologically related projects. If dependence of new patents on past patents from the same field is sufficiently strong, this intertemporal link may generate powerful negative effects on post-implementation innovation and explain the negative reduced-form DiD estimates.

I also scrutinize other mechanisms. The first alternative is that the observed relationship is driven by an adjustment of patenting strategies–e.g., a defensive breakup of patent applications in fields losing protection–unrelated to real innovation outcomes. However, the response of average patent quality to the policy does not support this hypothesis. The second alternative is that a positive patent term change decreases innovation by worsening the competitive environment. A patent lengthening could be a bad news for new innovators, because longer patent rights provide incumbents with stronger protection from competition and could even help anti-competitive strategies that aim at blocking entrants. Both forces would discourage innovative effort by new entrants. In addition, they could induce incumbents to innovate less due to lower competitive pressure and the opportunity to rely on existing patents longer. However, I provide evidence that competition measures–such as the entry intensity of new applicants or the concentration of patents among different innovators in a field–are unaffected by changes in patent length. In addition, movements in the average quality of patents granted to incumbents–both in absolute terms and relative to new applicants–do not support the anti-competitive use of patent rights to foreclose competition. Finally, incumbents do not seem to innovate less because they rely on existing patents longer: The renewal rate of patents at different maintenance stages is unaffected by the policy change. Therefore, competition does not seem to directly react to patent term changes, which suggests that it is not the main driver of the post-implementation negative DiD estimates.

While other unexplored forces may also be at work, I document that the proposed technological spillover can fully account for the policy-driven drop in innovation after policy implementation. Therefore, I take it as the privileged explanation of post-implementation *persistence*. Subsection 4.2 details all the empirical analyses that document its action, and subsection 4.3 shows the evidence on alternative channels.

4.2 Post-implementation: Evidence of an dynamic technological spillover

If the proposed technological spillover drives the policy-driven drop in innovation after policy implementation, the following testable hypotheses should hold in the data:

H1 *If the post-implementation effect is driven by a technological spillover, its magnitude should be stronger in technical fields in which within-field technological dependence is higher.*

H2 *If the post-implementation effect is driven by a technological spillover, the degree of within-field technological dependence should fall in fields in which innovation drops before implementation.*

H3 *If the post-implementation effect is driven by a technological spillover, post-implementation R&D investment should be lower for firms more exposed to technological areas experiencing a slow-down of R&D before implementation.*

These hypotheses are tested and confirmed in the following subsections.

4.2.1 Testing H1 and H2: Backward citations by technical field

To test **H1** and **H2**, I perform an analysis by technical field and use patent backward citations to measure the strength of within-field technological dependence. I proxy the intensity with which new technologies in a given technical field rely on previous technologies from the same field by the share of patents filed in a given field that have at least one applicants' backward citation to some patents classified in the same technical field.[50,51] This within-field backward citations intensity measure is computed for each field both at the quarterly level and as an average across patents granted before the policy news. The latter is denoted by S_j and, to test **H1**, it is interacted with the treatment T_j in the following triple difference specification

$$\begin{aligned} C_{j,t} = \alpha_j + \kappa S_j + \sum_{k=1985Q1}^{2000Q4} \gamma_k \mathbf{1}_{(t=k)} + \sum_{k=1985Q1}^{2000Q4} \eta_k \mathbf{1}_{(t=k)} S_j \\ + \sum_{k=1985Q1}^{2000Q4} \theta_k \mathbf{1}_{(t=k)} T_j + \sum_{k=1985Q1}^{2000Q4} \beta_k \mathbf{1}_{(t=k)} T_j S_j + \varepsilon_{j,t} \end{aligned} \tag{3}$$

where the dependent variable is citations-weighted patents.[52] If **H1** holds true, the β_k's for the post-implementation period are expected to be negative, as the post-implementation negative effect should be stronger in fields with a stronger reliance on previous technologies from the same field. Figure 7 confirms that this is the case, supporting **H1**.

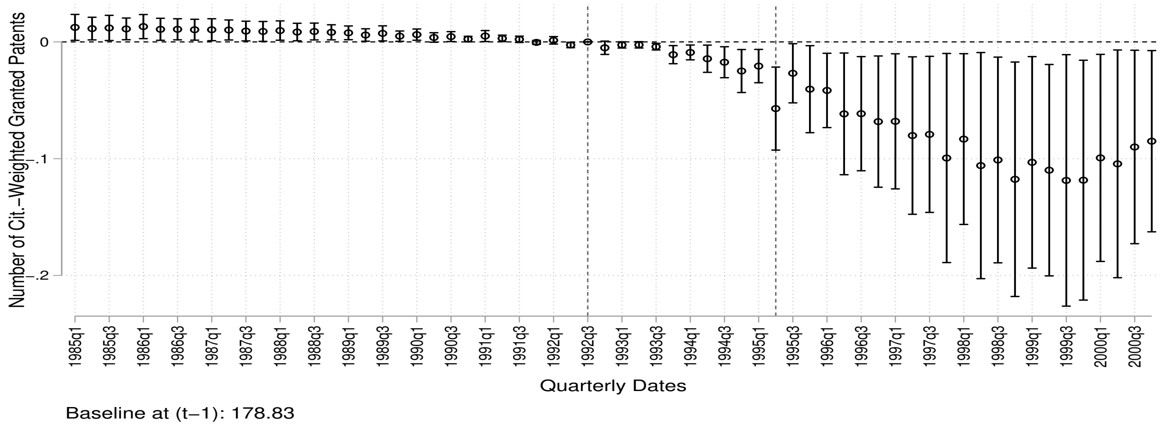
In addition, for **H2** to hold, measures of technological dependence should directly respond to the policy after its implementation: If policy-induced, post-implementation, innovations are directly linked to pre-implementation patents, within-field backward citations

[50]The preferred backward citations intensity measure considers only backward citations originally made by the applicant because the latter are thought to better represent genuine knowledge flows compared to other backward citations added ex-post by examiners to comply with legal requirements. Results are similar using all types of backward citations.

[51]As backward citations to prior art are a legal requirement, they may also reflect crowdedness of the technical field. Therefore, I focus on the share of patents backward citing prior art from the same field rather than the number of backward citations, which should be more prone to reflect crowdedness. In addition, the heterogeneous size of fields is captured by the technical field fixed effects in the econometric specification.

[52]Appendix B.4.6 report analogous evidence for granted patents.

Figure 7: **Heterogeneity analysis for the same-field citation intensity - Citation-weighted patents**



The plot shows the β_k coefficients of specification (3), having as dependent variable the number of citations-weighted patents filed in quarter-t and field-j. Clustered 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

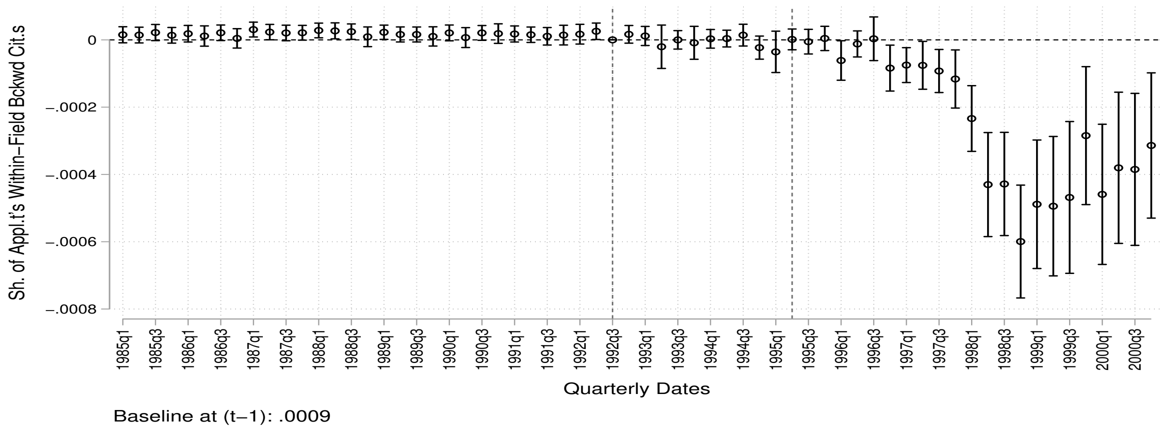
and their *intensity* should reflect it. Figure 8 plots the estimated β_k coefficients of specification (1) having as the dependent variable the quarterly and technical field-specific share of patents with at least one applicant-made backward citation to a patent from the same field. The negative estimated coefficients in the post-implementation phase support **H2** and provide suggestive evidence of the timing of the technology disclosure spillover, whose effect takes action approximately 4 years after the initial of the policy news on innovation.

I interpret this delay as the combination of *i*) a knowledge diffusion lag and of *ii*) R&D gestation lags. As to the first factor, most of the technical details of novel (patented) innovations are visible only after the publication of the patent document that grants protection.[53] Before 2000, patent documents were published upon the grant, and the average time between the creation of the innovation–proxied by the date of application for the patent– and the grant date was on average 2 years. Starting from the publication date, other innovators can learn about the details of recent technological advances and potentially begin new projects inspired by those. R&D gestation lags precisely refer to the average time that innovators need to finish a project, filing a patent application on the inventive output. Pakes and Schankerman (1986) estimate research gestation lags to be 2 years on average. Therefore, 4 years coincides with the time needed to new innovators to learn from recent technical advances and to produce new innovations based on the latter.

Finally, in support of **H2**, I also find that the total number of applicant-made within-field

[53] As a legal requirement, patents must provide a detailed description of the technical characteristics of the invention.

Figure 8: **Marginal effect of 1 more day of protection on the within-field backward citations intensity**



The plot shows the β_k coefficients of specification (1) having as dependent variable the share of patents classified in field j and filed in quarter t that has at least one applicant-made backward citation to patents classified in the same field j. Clustered 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

backward citations moves consistently with Figure 8. I also provide evidence of *direct* technological links between post-implementation and pre-implementation patents in the same field.[54] Post-implementation patents show lower within-field backward citations rates to patents filed before implementation exactly in those technical fields with a relative patent protection gain–i.e., fields experiencing an innovation slowdown in the pre-implementation period.

4.2.2 Does the technology link act within-firm or between-firms?

The evidence in favor of **H1** and **H2** suggests that the main driver of post-implementation persistence is an intertemporal technological link among innovations in the same technological field. An open question is whether such link mainly occurs within the firm or between firms. In the former case, it should be interpreted as an internality, while in the latter case it should be considered as a spillover. To distinguish between the two cases, I carry out a decomposition of the aggregate change in innovation induced by the policy during the post-implementation period 1996-1999. I isolate three components: The direct effect of the implementation of the new patent term, the between-firms spillover, and the within-firm internality. The strategy leverages on two ingredients.[55] The first is the assumption that, in the pre-implementation phase, the policy-driven technology disclosure channel is muted for all the firms, which allows to isolate the policy-induced innovation shock. The

[54]Results for both analyses are reported in Appendix B.4.7 and Appendix B.4.8.
[55]I leave to Appendix B.4.9 all the details about the steps involved in these calculations.

second ingredient is the fact that the within-firm component is necessarily muted for entrant firms, i.e. innovation produced by entrant firms in the period in which they enter cannot be driven by their own past innovations. I find that the between-firms component of the decomposition accounts for more than 99% of the policy-induced change in innovation in the post-implementation period. This component largely outweighs the contribution of the within-firm internality, which is close to zero. Therefore, the decomposition exercise strongly supports that post-implementation effects are driven by an *externality* across innovators.

4.2.3 Testing H3: R&D investment in the post-implementation period

Lastly, I perform a firm-level analysis to explore the validity of **H3**. The aim is to examine whether firms that are ex-ante technologically close to other firms whose R&D expenditure falls in 1992-1995 due to the policy, spend less on R&D in the 1995-1999 post-implementation period. I make two key assumptions for this exercise: *i*) The spillover does not affect the pre-implementation response of R&D, which is driven by the firm-level treatment only; *ii*) the spillover affects firms' R&D investment with a delay. I would find support for **H3** if the lagged spillover measure positively impacts firm-level R&D in the post-implementation period, conditionally on the firm-level private treatment. The analysis proceeds in steps. The firm-level sample is aggregated over the 4 periods 1986-1988, 1989-1991, 1992-1995, and 1995-1999, that correspond to control (-2), pre-news (-1), news (0), and post-implementation (1) periods, respectively. Under the first assumption, it is possible to adapt the difference-in-difference specification (2) to estimate the effect of T_i on firm-level R&D in the pre-implementation periods (-2),(-1), and (0) only, disregarding the lagged spillover effect. The post-implementation period ($p = 1$), when the spillover should be in action, is excluded from estimation. The fitted values of the treatment-induced R&D for firm i in period p are computed as $\hat{R}_{i,p} = \exp\{ln(\widehat{1 + R\&D_{i,p}})\} - 1$. In order to build a firm-specific measure of spillover, I follow the literature and I compute, for every pair of firms (i, j), Jaffe (1986)'s measure of technological distance, denoted by $d_{i,j}$.[56] The externality measure for firm i in period p is

$$E_{i,p} = \sum_{j \neq i} d_{i,j} R\&D_{j,p}$$

and it can be computed also using the fitted R&D measure $\hat{R}_{i,p}$ described above, in which case it is denoted by $\hat{E}_{i,p}$. Finally, it is possible to run the regression of interest, estimating the period-specific effect of the *lagged* spillover measure and of the firm-specific treatment

[56]The formula for the technological distance measure is $d_{i,j} = \frac{f_i f_j'}{\sqrt{(f_i f_i')(f_j f_j')}}$, where f_i is a vector that reports the number of patents obtained by firm i in a given class over the period 1971-1991.

on firm-level R&D spending. The specification is

$$\begin{aligned} \tilde{R\&D}_{i,p} &= \alpha_i + \sum_j \eta_{1,j} sic_j + \sum_j \sum_{k=-1}^{1} \eta_{2,j,k} sic_j \mathbf{1}_{(p=k)} + \sum_{age \in A} \delta_{age} + \\ &+ \sum_{k=0}^{1} \gamma_k \mathbf{1}_{(p=k)} + \sum_{k=0}^{1} \beta_k \mathbf{1}_{(p=k)} T_i + \sum_{k=0}^{1} \delta_k \mathbf{1}_{(p=k)} \tilde{E}_{i,p-1} + \zeta \tilde{E}_{i,p-1} + \varepsilon_{i,p} \end{aligned} \tag{4}$$

where $\tilde{R\&D}_{i,p} = ln(1+R\&D_{i,p})$ and $\tilde{E}_{i,p-1} = ln(1+E_{i,p-1})$. Table 1 reports the results. The first column shows the OLS estimates and the second column refers to the IV specification where $ln(1 + E_{i,p-1})$ and its interactions are instrumented by the fitted value version of the externality measure, i.e. $ln(1 + \hat{E}_{i,p-1})$. The other columns report the estimates of the first stage regressions of $ln(1 + E_{i,p-1})$ alone (column (5)), and interacted with the 1989-1991 and 1992-1995 dummies (columns (3) and (4), respectively). The F-statistic of the excluded instruments exceeds 30 in all first-stage regressions. In column (2), the delayed spillover variable is not statistically different from 0 both in the baseline term and when interacted with the 1989-1991 period dummy, but is positive and statistically significant when interacted with the post-treatment period dummy. This confirms that the positive spillover contributes to post-implementation effects and acts only after implementation. The firm-specific treatment remains negative and economically sizable in the pre-implementation phase but drops to zero in the post-implementation period. Overall this evidence supports **H3**.

4.3 Evidence on other competing explanations

This subsection investigates additional structural forces that may act as the drivers of post-implementation persistence. The first is manipulation of patenting strategies, and the second is a deterioration of competition in the technological field, discouraging entrant innovators and inducing incumbents to be lazier. However, neither seems to find support in the data.

4.3.1 Manipulation of patenting strategies and quality

A patent term change may produce a change in firms' patenting strategies. For example, a reduction of patent protection may induce a strategic breakup and staggered filing of lower-quality patents on the same invention, with the aim of protecting the innovation for longer than a single patent would do. This would artificially inflate the patent count in fields losing protection without any change to actual innovation. However, Section 3 provides evidence against this hypothesis. First, the average quality and the average value of patent do not change in response to the policy. Second, measures of R&D investment react to the policy consistently with aggregate innovation outcomes, confirming the real nature of the effect.

Table 1: **Firm-level evidence on the delayed investment spillover**

	(1) OLS	(2) IV	(3) FS1	(4) FS2	(5) FS3
d_{92-95} × Firm Treat.	-0.00057***	-0.00040	-0.00327***	0.00032***	-0.00024***
	(0.00018)	(0.00029)	(0.00061)	(0.00009)	(0.00004)
d_{96-99} × Firm Treat.	-0.00034	0.00000	-0.00026***	-0.00243***	-0.00046***
	(0.00033)	(0.00065)	(0.00009)	(0.00058)	(0.00011)
d_{92-95} × R&D Ext.$_{(t-1)}$	0.02197	0.02361			
	(0.01584)	(0.01625)			
d_{96-99} × R&D Ext.$_{(t-1)}$	0.02657	0.03511*			
	(0.02177)	(0.01968)			
R&D Ext.$_{(t-1)}$	0.23727	0.75417			
	(0.18213)	(0.82114)			
$d_{92-95} \times \widehat{R\&DExt.}_{(t-1)}$			1.03960***	-0.00431	0.00312
			(0.02265)	(0.00741)	(0.00335)
$d_{96-99} \times \widehat{R\&DExt.}_{(t-1)}$			-0.01582***	1.03829***	-0.00863
			(0.00498)	(0.01942)	(0.00855)
$\widehat{R\&DExt.}_{(t-1)}$			-0.38329**	1.49015***	0.25297***
			(0.18539)	(0.29041)	(0.05465)
Firm F.E.	Y	Y	Y	Y	Y
Period F.E.	Y	Y	Y	Y	Y
Age F.E.	Y	Y	Y	Y	Y
Industry×Period F.E.	Y	Y	Y	Y	Y
Observations	5132	5132	5132	5132	5132

Column (1) reports the OLS estimates of the specification (4). Column (2) reports the results of IV estimation of the same specification where the externality variable and its interaction terms are instrumented with the externality measure computed using the fitted value from a regression of firm-level R&D on the firm-specific change in protection based on the 1986-1988, 1989-1991, and 1992-1995 periods. Columns (3), (4), and (5) report the first stage regressions coefficients. Statistical significance levels: $^{*}(p < 0.10)$,$^{**}(p < 0.05)$,$^{***}(p < 0.01)$

4.3.2 A longer patent length may worsen competition and harm new innovators

Competitive pressure is expected to influence the desirability of patent protection for innovators and, in turn, patent length may affect innovation outcomes by changing the competitive environment.[57] Stronger patent rights provide incumbent firms with better protection from the competition of new entrants and may help anti-competitive practices aimed at foreclosing new innovators. Moreover, longer patents may also induce incumbents to innovate less: With lower pressure from entrants, they may simply be able to economically exploit existing patents for longer. Therefore, a patent lengthening may harm innovation both in the short run and in the long run, because it reduces entrants' and, potentially, incumbents' incentives to invest in R&D. I test these hypotheses empirically.

First, I test whether competition decreases due to an increase in patent length. Concentration of patents among different innovators and the entry rate of new applicants by technical field proxy competition. I build these measures using PATSTAT. As a measure of concentration by quarter and technical field, I use the Herfindahl-Hirschman Index based on the quarterly flow of granted patents. As to entry, I determine, by technical field and quarter, the percentage of patents granted to applicants that never filed a patent in the field before. Disambiguation of applicants is performed using STAN harmonized applicant's identifiers from the EPO Worldwide Bibliographic Database available in PATSTAT.[58] Finally, I run specification (1) having as dependent variables quarter- and field-specific measures of concentration and entry. Entrants contribute to the post-implementation effect proportionally to incumbents, so that the policy has no effect on the entry *rate*. Concentration, as measured by the HHI, also does not endogenously respond to the policy. The estimated β_k's are close to zero after both the policy news and implementation.[59] This evidence does not speak in favor of a deterioration of the competitive environment.[60]

Second, I show in Appendix B.4.5 that the average quality of patents filed by incumbent innovators–as measured by the average number of forward citations per patent–does *not* show major drops due to the implementation of a longer patent length. This is true in absolute terms and even more evidently relative to the average quality of patents granted to entrants. This evidence suggests that incumbents are *not* profiting from the longer term to

[57] In the extreme case of a firm with a guaranteed monopoly, the length of patent protection would not matter at all because there is no competitor that can imitate an innovation and erode profits.

[58] All the details on variables' construction are reported in Appendix D

[59] Appendix B.4.3 and B.4.4 report the results.

[60] In Appendix B.4.1 and B.4.2, I also test whether the patent term change has a stronger impact in more competitive fields. I run a triple-difference specification where the policy-induced change in patent length by technical field is interacted with either the average concentration or the average entry rate in the field *before* the policy news. I find support for the fact that patent length affects innovation by more in more concentrated fields. However, I do not find any significantly stronger impact in fields where the entry rate is higher.

foreclose competition.

Finally, I test whether a longer patent length increases the time for which innovators endogenously decide to keep their patents active. As argued above, both incumbents and potential entrants may have the incentive to reduce innovation. If the effect has equal strength across the two groups, competition measures such as entry rate would be unaffected in equilibrium. While it seems a rather special case, it cannot be ruled out a priori. Therefore, I turn to the test of an alternative implication of lower competition: A worsened competitive environment should induce less creative destruction from lower (potential) entry and, therefore, a longer average life of existing innovations. I use specification (1) to test whether a longer patent length increases the renewal rates of patents up the the maximum patent length or up to 11.5 years from grant. As in the previous cases, I do not find support for this hypothesis in the data.[61]

4.4 Pre-implementation: Competing interpretations and evidence

Previous subsections established that the negative post-implementation effect of patent length on innovation is directly related to the drop of innovation and the R&D fall upon news of a future patent term increase, i.e., the first empirical fact. Next, I turn to the interpretation of the latter. Economically, news of a future lengthening implies that the duration of patent protection obtainable by filing a patent application before implementation of the new policy is relatively shorter than what will be obtainable under the new regime. This observation opens the door to two opposing interpretations of the empirical fact.

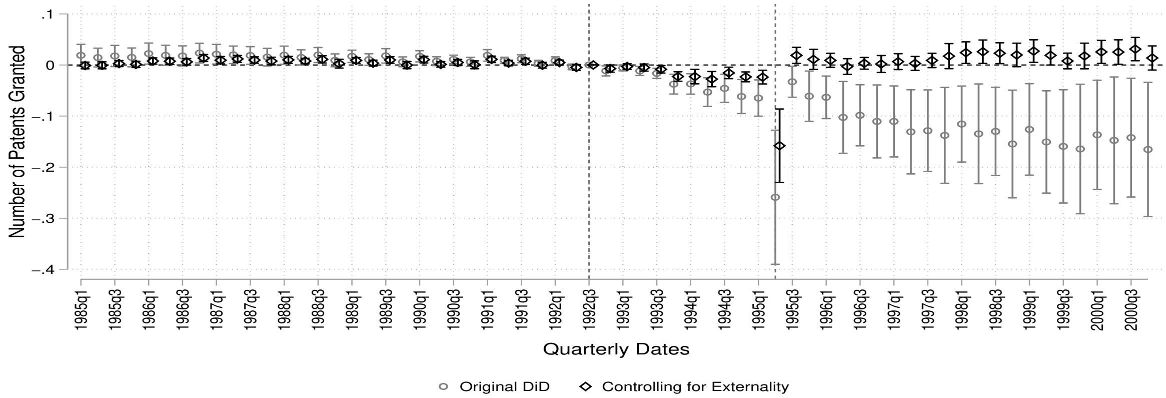
First, innovators reduce the pace of innovation at news of a future patent term increase because they want to profit of the relatively longer protection available after implementation. Indeed, if innovators prefer a longer patent length to a shorter one, they may want to reduce the speed at which they complete existing projects, and file the related patent applications after the implementation of the new regime.

Second, innovators reduce the pace of innovation at news of a future patent term increase because they anticipate some negative effect of the policy after its implementation. As discussed in subsection 4.3.2, new innovators may dislike a longer patent length because they may be discouraged or foreclosed by the stronger protection available to incumbents. At the same time, incumbents may reduce innovation because, due to lower competitive pressure, they can rely on existing patents for longer. However, subsection 4.3.2 showed that this narrative does not seem to find support in the data. While the empirical analyses cannot definitively rule out this interpretation, they suggest to look for alternative explanations.

Therefore, I turn to the first interpretation, and I provide suggestive evidence that supports it. I revisit the main DiD analysis in light of the findings of subsection 4.2, and I disen-

[61] Results are shown in Appendix E.1.4.

Figure 9: **Direct marginal effect of 1 more day of protection on granted patents**



The plot shows in black the β_k coefficients of the augmented DiD specification (5) and in gray the β_k coefficients of specification (1), having as dependent variable quarter-t and field-j number of granted patents. Standard errors are clustered by technical field and 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

tangle the direct effect of the patent term change from the impact of the technology disclosure externality. I run the specification

$$P_{j,t} = \alpha_j + \sum_k \gamma_k \mathbf{1}_{(t=k)} + \sum_k \beta_k \mathbf{1}_{(t=k)} T_j + \sum_k \psi_k \mathbf{1}_{(t=k)} \underbrace{\bar{P}_{j,k-16:k-1}}_{\equiv (1/16)\sum_{q=k-16}^{k-1} P_{j,q}} + \varepsilon_{j,t} \tag{5}$$

which is a version of (1) enriched to control for the quarter-specific impact of $\bar{P}_{j,k-16:k-1}$, which is defined as the average flow of patents in field j in the previous 16 quarters. This term is meant to capture the effect of the technology disclosure externality through the flow of recently granted patents. The chosen timing of 16 quarters, i.e., 4 years, is motivated by the evidence of subsection 4.2.

The coefficients of interest β_k capture in this case the *direct* effect of the news or implementation of a 1-day patent term increase, after controlling for the impact of the technology disclosure externality. Figure 9 plots in black the β_k of (5) and reports in gray the treatment effects from the original DiD specification (1). There are two takeaways. First, the β_k coefficients remain negative between news and implementation of a positive change in patent length, i.e., the direct effect of news of a future patent term is to depress innovation. This confirms the first empirical fact of section 3. Second, the β_k coefficient estimates of (5) turn to positive after policy implementation. The action of the technology disclosure externality fully accounts for relative drop in innovation after policy implementation—i.e., the second empirical fact. Instead, the *direct* effect of implementing a longer patent length is *positive* on innovation, which suggests that innovators indeed like longer patent protection. Quantitatively, the average of the post-implementation coefficients imply that a one-month increase in patent length increases innovation in the short run by 1.5% of the baseline 1992Q3 level. This corresponds to an elasticity of patenting to patent length of 3.

Figures B.43 and B.44 in Appendix B.4.10 report the results of specification (5) having as dependent variable citations-weighted patents and quarter- and field-specific measure of R&D effort, respectively. All the results are fully consistent with those in Figure 9.[62] In addition, Appendix B.4.11 reports additional results from a version of (5) estimated on aggregated sample sub-periods and exploiting an instrumental variable strategy that aims at ruling out potential confounders included in the lagged externality term. Results are consistent with Figure 9. Moreover, the coefficient estimate of the effect of the externality term on patenting in the post-implementation period allows to infer that the elasticity of post-implementation innovation to a 1% change in innovation in the "news" period is around 1. This is a synthetic, reduced-form measure of the strength of the technology disclosure externality and it comparable to the 0.5 elasticity of innovation to spillovers in the technological domain estimated by Bloom, Schankerman and Van Reenen (2013).[63]

4.5 Takeaways, interpretation, and key elasticity estimates

I interpret below the facts of section 3 in light of the evidence of subsections 4.2 and 4.4. While other economic interpretations of the results are possible, I focus on the one that seems to be most supported by the data.

First empirical fact At news of a future patent term increase, innovators reduce the pace at which they invest to complete *existing* projects because they want to file a patent under the new regime, which grants longer protection. The documented drop in R&D spending shows that the decline of innovation is not driven by a pure re-timing of patent applications but rather is generated by changes in actual innovation effort. Suppose that firms can decide the pace at which they develop abstract ideas into finished products–i.e., run projects–and that being faster becomes increasingly costly. In normal times, innovators would trade-off the benefits of getting monopolistic profits sooner against the higher costs of a faster pace of development. However, news of a future patent term extension makes a slower pace relatively more desirable. Firms can reduce their costs and may be able to file a patent application under the new regime, conditional on the future success of the project. Therefore, the response to the news shock is driven by *development* investment on *existing* projects. After

[62] I also assess the relevance of potential cross-field technological externalities, but I find that the latter seem to be negligible in the present context once controlling for within-field effects. Results are not reported, but are available upon request.

[63] Appendix B.4.12 reports all the details about the derivation of the elasticity of post-implementation innovation to a 1% innovation change in the pre-implementation period.

policy implementation, the new patent length is set, the temporary incentives to slowdown development terminate, and the previous trade-off returns to normal times.

Second empirical fact At implementation of the patent term extension, two counteracting forces affect innovation and total R&D. On the one hand, a technological spillover depresses the creation of *new* projects in fields where innovation has fallen due to policy news. On the other hand, the empirical evidence suggests that the direct effect of longer patent length is to promote more innovation, as it provides positive incentives to invest in new ideas. Therefore, despite this positive direct effect of the policy variable, the reduced-form impact of the shocks is crucially affected by anticipation and powerful technology disclosure externalities. As to anticipation, innovation falls at news of a future longer patent length, consistently with innovators seeking longer patent protection. As to technological spillovers, they translate the initial drop into a protracted decline even after implementation of the longer term. The timing of the impact of the spillover and the characteristics of patent documents are suggestive about the mechanisms through which the intertemporal link works. Indeed, patent documents convey to the public greater, better, and more precise information than undeveloped ideas or ongoing innovation projects. In this respect, a slower pace of development of ideas into products–and, therefore, into patent documents describing them–reduces the ability of innovators to learn from recent technological advances and to generate *new* related research ideas. Therefore, technology disclosure externalities and the *research* of ideas for *new* projects are the key drivers of the second empirical fact.

Key elasticities Given the interpretation of the first empirical fact, the DiD estimates of the pre-implementation period highlight a *positive* elasticity of innovation to patent term changes available at a future point in time. The estimates of section 3 imply that the elasticity of patenting to a 1% increase of patent length available in 2 years is 3.1. This figure increases to 9.1 considering a patent term increase available in just 1 year.[64]. In addition, the estimates of Figure 9 suggest that the short-run elasticity of patenting to patent length is 3.[65] Finally, appendix B.4.12 discussed that the elasticity of future innovation to current innovation shocks is 0.997.

5 Model

The analysis of the positive and normative consequences of patent term changes that differ from the observed quasi-experimental setting require a structural model. The latter also

[64]The elasticities are computed by using the quarterly estimates of subsection 3.1.2. First, I infer the effect of a 1-day future increase in patent length on quarterly patenting, expressed in percentage deviation from the average level of innovation in a technical field before the policy news. Second, I divide it by $1/(365 \times 17)$, which is 1-day expressed as a percentage of the status quo effective patent length of 17 years.

[65]I will use the structural model of section 5 to make more precise inference about the long-run elasticity of innovation to patent length.

allows me to shed light on aspects of the innovation process that cannot be directly captured by the data, such as the elasticity of long-run innovation to patent length. Unfortunately, while the evidence of Section 3 has an intuitive economic explanation, existing models of R&D-based endogenous growth would generate a counterfactual response of innovation and R&D effort to the policy shocks.

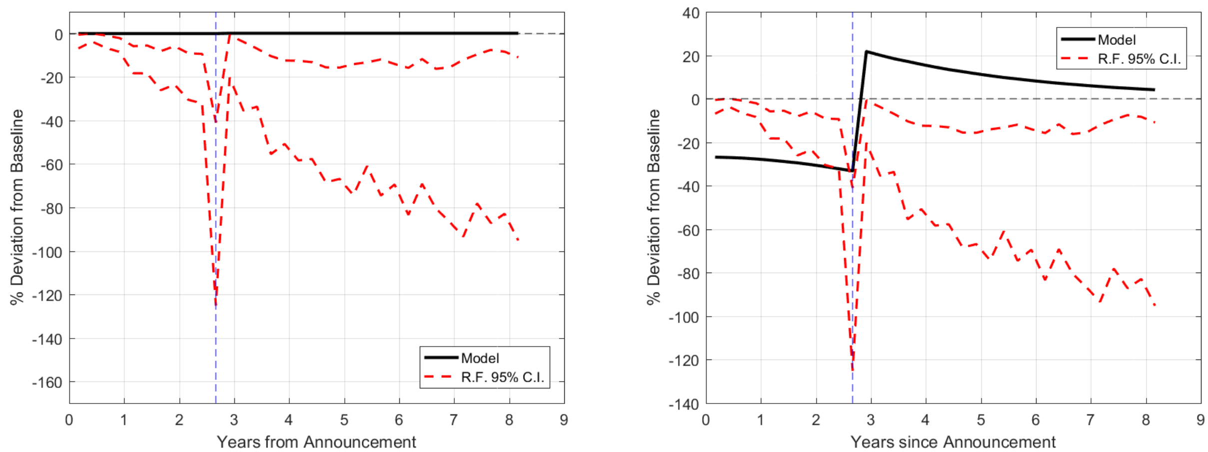
Therefore, I build a structural model of innovation with novel features, which reflect the description of the mechanisms of subsection 4.5. I start by discussing the performance of existing models in subsection 5.1, and present the key novel ingredients of my framework. I motivate the modelling choices and explain their consequences. Subsection 5.2 presents the standard parts of the model. Firms employ labor and intermediate capital varieties to competitively produce a final good. The expansion of the number of intermediate varieties drives endogenous productivity growth, which occurs through R&D. Innovators use raw physical capital to monopolistically produce patented varieties. The monopoly ends when the finite patent T on the technology expires. Therefore, only a fraction of total varieties is monopolistic, and this affects the distortions in the model. Subsection 5.3 describes research and development, which are formally distinct and constitute the novelties of the model. Firms invest in research to generate new ideas. Subsequently, they turn ideas into new intermediate capital varieties through development activity. Patents on novel varieties can be filed only at the end of development. Subsections 5.4 and 5.5 define the competitive equilibrium and describe the key mechanisms of the model, respectively. Appendix C.1 contains additional details and derivations.

5.1 Performance of existing theories and new ingredients

In standard models of R&D-based endogenous growth, research and development constitute a single activity. The output of R&D is new products and there is no separate role for abstract ideas. Implicitly, innovators are successful if they both get a new idea through research and develop the idea into a product within the same period. Otherwise the idea is lost and the process re-starts from scratch next period. This is in clear contrast with the discussion of subsection 4.5 on the separate role of existing projects' development and research of new ideas to explain the two empirical facts and, consistently, has important implications for how in standard models innovation and R&D respond to an anticipated patent term increase analogous to the one observed in the data. The left panel of Figure 10 shows the reaction of innovation in the variety-expansion model of Jones (1995), adapted to feature finite patent length.[66] The figure compares the model-based response to the reaction implied by the em-

[66]This is the basic setup on which I nest the new ingredients of my model. Appendix F.1 outlines the full description of Jones (1995), adapted to feature finite patent length. For the policy simulation, the parameters of are set to the estimates of Section 6. However, the specific choice of the parameters does not affect the qualitative

Figure 10: **Policy in a model without new ingredients *vs.* empirical estimates**



(a) R&D in a single stage (Jones (1995))

(b) R&D distinct, but no spillover

The plots show the response of innovation (patents) to a 100-days patent term increase anticipated by 2 years and 8 months. The news shock occurs at time $t = 0$. The red dashed lines are the 95% confidence intervals implied by the reduced-form estimates of Section 3. In the left panel, the black solid line shows the response of Jones (1995) semi-endogenous growth model, adapted to features a finite patent length. Appendix F.1 provides a full description. In the right panel, the black solid line shows the response of a model where research and development are distinctly modelled, but there is no spillover from the latter to the former. This model is equivalent to the one of Section 5, setting $\chi = 0$. In both model-implies responses, the system is assumed to be at the pre-policy change steady state at $t = 0$.

pirical estimates of section 3. In the model (black solid line), innovation increases mildly at the news and fully adjusts to its new higher level only at implementation of the patent length increase.[67] This is in stark contrast with the sizable drop of innovation observed in the data (red dashed lines).

According to the evidence-based narrative of subsection 4.5, the drop of innovation at the news is driven by a slowdown of the pace of development of the stock of ideas into patented products. The policy creates an incentive to keep the idea and to try to develop it into a product–filing a patent application that protects the developed technology–once the longer patent length will be in place. However, for this mechanism to work, ideas must be formally distinct from products and must be storable. As argued above, this is not the case in standard models: Research–that generates ideas–and development–that transforms ideas into products–implicitly occur as a single activity. If not developed within the period,

implications of the analysis.

[67] The same would occur in Romer (1990)'s model. Instead, in Schumpeterian growth models where the average length of protection is modelled through a Poisson arrival rate of imitators, news of a future increase of patent length–i.e., a fall in the arrival rate of imitators–implies a lower discounting of future, post-implementation, profits. At news, this immediately boosts the value of an innovation, causing a rise–rather than a fall–of R&D effort and innovation.

ideas vanish and have no value. Therefore, overall R&D directly responds only to changes in the value of patented products. Since the latter fully adjusts to its higher value only upon implementation of longer patent length, R&D investment and innovation behave similarly.

Hence, the first novel ingredient of my structural model is that research and development are explicitly treated as distinct activities. Innovation occurs in two steps. First, research generates novel ideas, which can be stored by the firm that came up with them. Second, the firm invests to develop the stock of ideas into intermediate capital varieties, i.e., new technologies that increase the productivity of the economy. Since abstract ideas cannot be patented, innovators can file patent applications only at the end of development, which terminates the innovation process.[68] The 2-stages structure formalizes the narrative of subsection 4.5 and allows to replicate the empirical reaction of inventive activity at news of a future patent term increase. The *current* value of a new variety does not increase sizably at news. It adjusts only at implementation–i.e., when the longer patent length becomes actually available. On the contrary, the value of an undeveloped idea (project) increases immediately at news, because it discounts the higher value of a *future* patented variety. This induces innovators to optimally *reduce* the pace of development to profit of the longer patent length.[69] Therefore, patents–which are filed at the end of development activity–fall at news, and so does total R&D through a marked fall of the resources spent on development.[70]

The 2-stages structure of the innovation process is mathematically similar to Comin and Gertler (2006). However, the two settings radically differ in terms of content of the two steps. In Comin and Gertler (2006), the first step encompasses both research *and* development, which constitute again a single activity that generates patented intermediate varieties. The second step is adoption, i.e., inclusion of technologies produced by the first step into the consumption good. The modelling of R&D as a single activity implies that the response of the model to the policy news is similar to Jones (1995), because the values of R&D and adoption respond symmetrically to patent term changes.[71]

The second novel ingredient of my model–which constitutes a further departure from Comin and Gertler (2006)–is that the average pace of development positively affects new

[68] The separation of research and development also reflects the fact that, in the real world, the two activities are distinct, and that the development of an idea into a finished intermediate product may take several years to firms.

[69] This implies that my theoretical setup nests Jones (1995) model if the development stage is suppressed by assuming that there are no developments costs.

[70] In the model, research investment increases at news. Research produces new ideas and, as argued above, the value of an idea increases at news, promoting investment in research. However, the drop in development is quantitatively stronger and depresses total R&D.

[71] At the news of a future patent term increase, the value of adoption does not change immediately, as the period over which profits can be collected–i.e., patent length–does not change until implementation. Therefore, the value of R&D output–i.e., the value of subsequent adoption net of adoption costs–also does not change. As the two values do not respond to the policy news, investment in both adoption and R&D does not change.

ideas' research productivity. A faster average pace of existing projects' development increases the ability of other researchers to generate new related ideas. Consistently with the empirical evidence of section 4, this effect is modelled as an externality and innovators do not consider this positive effect when choosing how fast to turn their own ideas into intermediate varieties and patents. As shown by the right panel of Figure 10, the spillover is crucial to reproduce the post-implementation effect of the policy.[72] The slowdown of the pace of development at news of a future patent term extension reduces the availability of patent documents describing recent technological advances and, therefore, hinders knowledge diffusion. As productivity of research is lower, a lower investment in research depresses total R&D. In addition, innovation falls because of a lack of new ideas to turn into varieties. I move now to a more formal description of my theoretical setup.

5.2 Standard parts of the model: Consumers and producers

5.2.1 Consumers

The representative consumer has linear utility $u(c(t)) = c(t)$ in per-capita consumption, discounts the future at rate ρ, saves in real assets at a real rate of return $r(t)$, and inelastically supplies labor.[73] Aggregate labor supply coincides with population $L(t)$, which exogenously grows at rate n. Appendix C.1.1 formally sets up the representative agent's maximization problem and shows that the Euler equation is $r(t) = \rho \;\; \forall t$.

5.2.2 Competitive production of the final good

There is a competitively produced final good in the economy. The representative firm solves

$$\max_{L(t),\{X_{(i,t)}\}_i} \left\{ (h(t)L(t))^{1-\alpha} \int_0^{V(t)} X(i,t)^{\alpha} di - w(t)L(t) - \int_0^{V(t)} z(i,t)X(i,t)di \right\} \tag{6}$$

where $Y(t) = (h(t)L(t))^{1-\alpha} \int_0^{V(t)} X(i,t)^{\alpha} di$ is the production function of the final good. $h(t)$ is an exogenous productivity term, $L(t)$ is labor, and $X(i,t)$ is the amount of the intermediate capital variety i used in production. The wage rate is $w(t)$, and the price of each capital variety is $z(i,t)$. Final good producers generate the demand for labor and for each of the $V(t)$ intermediate capital varieties available in the economy. The expansion of $V(t)$ drives endogenous productivity growth in the model. New varieties originate from firms' investment in

[72] The right panel of Figure 10 shows the response of innovation to an anticipated patent term increase implied by a model where research and development activities are distinct but with no spillover, and compares it to the response implied by the empirical estimates of Section 3. This model is equivalent to the model of Section 5 with the externality parameter $\chi = 0$.

[73] The economy features multiple real assets, such as physical capital and firms' stocks. No arbitrage conditions ensure that, in the absence of uncertainty in this economy, the real rate of return is equal across assets. Appendix C.1.1 formally describes the representative agent's utility maximization problem and precisely defines total assets.

research and development. Successful firms become monopolistic producers of the intermediate variety until the patent that protects their monopoly expires after the finite term of T periods.

5.2.3 Production of intermediate capital varieties

Therefore, production of intermediate varieties can be either monopolistic or competitive. Firms produce intermediate capital varieties through a linear technology, by hiring raw capital $K(t)$ from the households at a competitive rate $(r(t) + \delta)$. δ is the depreciation rate of physical capital. Monopolistic producers maximize profits, taking the inverse demand for each variety as given. Their optimization problem is

$$\begin{aligned} &\max_{X(i,t),z(i,t)} \Big\{ z(i,t)X(i,t) - (r(t)+\delta)X(i,t) \Big\} \\ &\text{s.t.} \quad z(i,t) = \alpha h(t)^{1-\alpha} L(t)^{1-\alpha} X^{\alpha-1}(i,t) \end{aligned} \tag{7}$$

where the constraint equals the price $z(i,t)$ of the intermediate variety to final good producers' inverse demand.

When the maximum patent term T expires, the production of intermediate varieties becomes perfectly competitive. The maximization problem of competitive intermediate producers is analogous to (7), but the constraint becomes $z(i,t) = r(t) + \delta$, as competition drives the price to marginal cost.

5.2.4 Evolution of the share of monopolistic varieties

At any instant t, the creation of new varieties protected by patents and the expiration of old varieties generated at $t - T$–T being the finite patent length–determine the evolution of the share $\zeta(t)$ of total varieties that are monopolistic. Starting from the definition of $\zeta(t) \equiv \frac{V_p(t)}{V(t)}$– where $V_p(t)$ is the total number of monopolistic varieties–and time-differentiating both sides, it is possible to write the evolution of $\zeta(t)$ as

$$\dot{\zeta}(t) = (1-\zeta(t))\frac{\dot{V}(t)}{V(t)} - (1+\psi)\frac{\dot{V}(t-T)}{V(t)} e^{-\int_{t-T}^{t} \lambda(t')\mathrm{d}t'} \tag{8}$$

Appendix C.1 reports the detailed derivation of (8). Intuitively, the first addend captures the net contribution of new varieties produced at time t–i.e., $\dot{V}(t)$–to the growth of patent-protected varieties. The second added represents the fall of monopolistic varieties due to the expiration of patent protection on the gross mass of intermediates generated at $t - T$–i.e., $(1+\psi)\dot{V}(t-T)$. In the last term, $e^{-\int_{t-T}^{t} \lambda(t')\mathrm{d}t'}$ represents the fraction of these $t - T$ varieties which have survived the process of creative destruction until t.

Indeed, the model assumes that new innovations destroy existing varieties at an endogenous instantaneous probability $\lambda(t) \equiv \psi\frac{\dot{V}(t)}{V(t)}$. This is a reduced-form way to formalize the destructive effect of entry on on existing monopolies. I assume that creative destruction is proportional to the aggregate innovation rate and unaffected by T, in line with the results of Section 4 on competition.[74]

5.2.5 Capital market clearing condition and resource constraint

Clearance of capital market requires that the total stock of capital is equal to the quantity used for the production of the $V(t)$ existing varieties, i.e., $K(t) = \int_0^{V(t)} X(i,t)di$. In addition, the evolution of aggregate capital is governed by $\dot{K}(t) = I_K(t) - \delta K(t)$: Capital grows with new investment $I_K(t)$ by households and falls with depreciation $\delta K(t)$.

The resource constraint of the economy is

$$Y(t) = C(t) + I_K(t) + I_R(t) + \mu\upsilon(t)\iota_D(t)^\theta N(t) \tag{9}$$

$Y(t)$ is the total production of the final good determined by problem (6), $C(t)$ is aggregate consumption, $I_K(t)$ is investment in physical capital, $I_R(t)$ is investment in research activity, and $\mu\iota_D(t)^\theta\upsilon(t)N(t)$ is aggregate investment in development. The latter is the product of development costs on each project ($\mu\iota_D(t)^\theta\upsilon(t)$, with μ being a scale parameter, $\iota_D(t)^\theta$ the convex costs of development pace, and $\upsilon(t)$ the value of a patent) and the total number of projects $N(t)$. Investment in research and development is described in subsection 5.3.

5.3 Research and development

5.3.1 Research activity

Firms that want to obtain a patent on a new intermediate variety must separately succeed in both research *and* development. Research is competitive and firms invest to discover new ideas/projects, whose stock in the economy is $N(t)$ and whose value is $P(t)$. Crucially, an idea is storable by the firms that came up with it, and its value $P(t)$ is positive because the firm can *exclusively* develop the idea into an intermediate variety. All firms investing in research are identical and solve

$$\max_{I_R(t)} \left\{ P(t)\Big[E(t)^\chi V(t)^{\phi_1} I_R(t)^{\phi_2}\Big] - I_R(t) \right\} \tag{10}$$

where $E(t)^\chi V(t)^{\phi_1} I_R(t)^{\phi_2}$ is the production function of projects, and $I_R(t)$ is aggregate research investment in units of the final good.[75] The production function assumes that new

[74] Creative destruction is a standard effect in the literature. The model could also feature imitation: Old varieties would not disappear but their profits would be pushed to zero. This is omitted from the model because it does not add any relevant theoretical insights. Appendix F.3 shows the positive and normative implications of the model assuming that creative destruction depends linearly or quadratically on T.

[75] In an extension of the model used for welfare analysis, the production function of projects is transformed into $(1-\zeta(t))^{\phi_1\eta} E(t)^\chi V(t)^{\phi_1} I_R(t)^{\phi_2}$, where $(1-\zeta(t))^{\phi_1\eta}$ is a distortion term that represents the probability of *not*

ideas increase with research investment $I_R(t)$–subject to decreasing returns governed by ϕ_2– and in the number of existing intermediate capital varieties $V(t)$. $V(t)^{\phi_1}$, with $\phi_1 < 1$, captures the *standing on the shoulders of giants* effect, i.e., the contribution of the existing *stock* of knowledge to the creation of new ideas. This effect is standard in the endogenous growth literature. The parameters are constrained so that $\phi_1 + \phi_2 < 1$.

The term $E(t)^{\chi}$ formalizes the new technology disclosure externality documented in subsection 4.2. It is defined as $E(t) \equiv d^{-1} \int_{t-d}^{t} N(s)^{-1} \int_0^{N(s)} \iota_D(j,s) dj ds$, where, d is the (maximum) delay with which the externality acts. As explained in the next subsection, $\iota_D(j,s)$ is the flow probability that project $j \in [0, N(s)]$ is successfully developed at instant s. Higher ι_D means a shorter average duration of development activity. Therefore, $E(t)$ is the average pace of development during previous d years, with $d = 4$ to match the evidence of section 4. As the average pace of development is faster, the diffusion of *novel* technical knowledge to other innovators through patent documents is more rapid. Being able to learn from *recent* advances more rapidly positively affects the ability of all other innovators to generate new ideas, beyond the standard "standing on the shoulders of giants effect" captured by the *stock* of varieties $V(t)$.

5.3.2 Development activity

Next, firms must choose the optimal pace of development, i.e. the speed at which they try to turn existing ideas of value $P(t)$ in patented intermediate varieties of value $\upsilon(t)$. Following Lin and Shampine (2018), I assume that monopoly power on a variety can last *at most* T years.[76] Hence, the value of a new patent issued at t is

$$\upsilon(t) = \int_t^{t+T} e^{-\int_t^s (r(t')+\lambda(t'))dt'} \pi(s) ds \tag{11}$$

where $\pi(s)$ is the flow of profits at instant s–and it is the solution to the maximization problem (7)–$r(t')$ is the real interest rate, and $\lambda(t')$ is the instantaneous probability of endogenous creative destruction. Therefore, $\upsilon(t)$ is the expected net present discounted value of profits from a patent-protected monopoly on a variety[77]. The development problem is independent across ideas/projects and its value function is

$$r(t)P(t) - \dot{P}(t) = \max_{\iota_D(t)} \left\{ \iota_D(t) \Big[\upsilon(t) - P(t) \Big] - \mu \iota_D(t)^{\theta} \upsilon(t) \right\} \tag{12}$$

having an idea blocked by an existing monopoly. It is decreasing in the share of varieties that are monopolistic and in the parameter η, which captures the severity of the distortion.

[76]This gives rise to forward and delayed terms in the system of differential equations that solves the model, and I build on Lin and Shampine (2018)'s relaxation algorithm for the solution.

[77]From (11), we can see that, while the maximum statutory patent length is T, the effective one can be shorter. The expected patent duration along the balanced growth path (b.g.p) equilibrium is $T^e \equiv \frac{1}{\lambda^*}(1 - e^{-\lambda^* T})$, where λ^* is the endogenous rate of creative destruction along the b.g.p.

where $\iota_D(t)$ the pace of development chosen by firms–i.e., the instantaneous probability of turning the project into a product. The total investment needed to achieve any given $\iota_D(t)$ is $I_D(t) = \mu\iota_D(t)^\theta \upsilon(t)$ units of the final good. μ is a scale parameter, and $\iota_D(t)^\theta$–with $\theta > 1$–captures the convex costs of development pace. The term in square brackets shows that a successful firm generates an intermediate capital variety of value $\upsilon(t)$, but loses the value $P(t)$ of the idea/project, which is terminated. Crucially, when firms choose the optimal $\iota_D(t)$, they do not internalize the positive effect that a faster pace of development has on subsequent productivity of research.

5.3.3 Evolution of projects and varieties

Equation (13) governs the evolution of varieties $V(t)$

$$\dot{V}(t) = \iota_D(t)N(t) - \psi\dot{V}(t) \tag{13}$$

The first addend represents the increase in varieties due to ideas/projects turned into varieties by development activity. Since all projects have a symmetric instantaneous probability of success $\iota_D(t)$ in equilibrium, the mass of projects turned into patents at instant t is given by this probability *times* the total number of projects $N(t)$, i.e., $\iota_D(t)N(t)$. The second addend represents creative destruction, with $\psi\dot{V}(t)$ varieties destroyed by new ones.

The evolution of ideas/projects $N(t)$ follows the law of motion

$$\dot{N}(t) = \left[d^{-1}\int_{t-d}^{t} \iota_D(s)ds\right]^\chi V(t)^{\phi_1} I_R(t)^{\phi_2} - \iota_D(t)N(t) \tag{14}$$

The first addend represents the mass of new ideas/projects generated by research, i.e., the research production function of problem (10). $\left[d^{-1}\int_{t-d}^{t} \iota_D^*(s)ds\right]^\chi$ replaces the externality term $E(t)^\chi$, using its definition and the fact that, in equilibrium, all projects are symmetric. The previous paragraph already discussed the second added, which is the mass of projects turned into varieties at instant t.

5.4 Definition of the competitive equilibrium

A competitive equilibrium equilibrium for this economy is a sequence of quantities $\{V^*(t), N^*(t), \{X^*(i,t)\}_{i=0}^{V^*(t)}, \{\iota_D^*(j,t)\}_{j=0}^{N^*(t)}, I_R^*(t), I_K^*(t), C^*(t), K^*(t), \pi^*(t), \zeta^*(t)\}_{t=0}^{\infty}$, *prices* $\{r^*(t), w^*(t), \{z^*(i,t)\}_{i=0}^{V^*(t)}\}_{t=0}^{\infty}$, *and values* $\{P^*(t), \upsilon^*(t)\}_{t=0}^{\infty}$, *such that, given the exogenous evolution of* $\{h(t), L(t)\}_{t=0}^{\infty}$, ***i)*** $r^*(t) = \rho$ ***ii)*** $C^*(t)$ *and* $I_K^*(t)$ *solve consumer's utility maximization problem;* ***iii)*** $L(t)$ *and* $\{X^*(i,t)\}_{i=0}^{V^*(t)}$ *solve problem (6);* ***iv)*** $X^*(i,t)$ *and* $z^*(i,t)$ *solve problem (7)* $\forall i \in [0, V^*(t)]$; ***v)*** $I_R^*(t)$ *solves problem (10);* ***vi)*** $\iota_D^*(t)$ *solves problem (12) for all* $j \in [0, N^*(t)]$; ***vii)*** $\upsilon^*(t)$ *satisfies equation (11);* ***viii)*** $\pi^*(t) = (\rho+\delta)^{-\frac{\alpha}{1-\alpha}}\left(\alpha^{\frac{1+\alpha}{1-\alpha}} - \alpha^{\frac{2}{1-\alpha}}\right)h(t)L(t)$ *from the solution of* *problem* (7); ***ix)*** $P^*(t)$ *satisfies*

$$P^*(t) = \int_t^\infty e^{-\int_t^s [\rho + \iota_D^*(t')]dt'} [\iota_D^*(s) - \mu \iota_D^*(s)^\theta] \upsilon^*(s) ds \tag{15}$$

x) $\zeta^*(t)$ *satisfies equation* (8), ***xi)*** $K^*(t)$ *satisfies* $K^*(t) = \int_0^{V^*(t)} X^*(i,t)di$ *and* $\dot{K}^*(t) = I_K^*(t) - \delta K^*(t)$; ***xii)*** *the aggregate resource constraint* (9) *holds;* ***xiii)*** $V^*(t)$ *satisfies* (13); *and* ***xiv)*** $N^*(t)$ *satisfies* (14).

Along the balanced growth path (b.g.p.) equilibrium, each variable $x(t)$ *grows at a constant rate* g_x*, so that* $x(t) = e^{g_x t}\tilde{x}(t)$*, where* $\tilde{x}(t)$ *is the stationary version of* $x(t)$*.*

Appendix C.1 provides the details of the solution, and Appendix C.1.10 shows that the economy admits a balanced growth path and derives the growth rates.[78]

5.5 The mechanism at work

How can the model reproduce the empirical facts of section 3? I answer the question by studying its response to an *anticipated* patent term increase.[79]

First empirical fact At the news of a future extension of T, the *current* value of a new variety $\upsilon^*(t)$ in equation (11) does not increase markedly until the effective implementation of the policy, because it still discounts profits over the old, shorter, patent term. However, equation (15) shows that the value of an idea $P^*(t)$ increases on impact, because it discounts the higher value of *future* patents.[80]. Optimal investment decision rules (16) and (17) shape the reaction of research and development to news-induced changes in $\upsilon^*(t)$ and $P^*(t)$.

$$\iota_D^*(t) = \left[\frac{(\upsilon^*(t) - P^*(t))}{\theta \mu \upsilon^*(t)}\right]^{\frac{1}{\theta - 1}} \tag{16}$$

$$I_R^*(t) = \left(P^*(t)\left[d^{-1}\int_{t-d}^t \iota_D^*(s)ds\right]^\chi V^*(t)^{\phi_1}\right)^{\frac{1}{1-\phi_2}} \tag{17}$$

In equation (16), convex costs ($\theta > 1$) imply a desire to smooth development intensity in normal times. However, the combined movements of $\upsilon^*(t)$ and $P^*(t)$ induce a slowdown of the pace of development $\iota_D^*(t)$ before implementation. As $P^*(t)$ increases, (17) highlights that research investment rises at first. For most parameter values–and at the structural estimates of section 6–the drop of development investment is stronger than the rise in research. Therefore, total R&D falls. In addition, since firms obtain patents only at the end of development, innovation also drops.

[78] Appendix F.4 shows that the model can be equivalently formulated assuming that R&D uses labor rather than units of the final good.

[79] Appendix C.4 contains the model-implied responses to the simulated policy change for key variables.

[80] Let A be anticipation, the value of future patents is represented by $\upsilon^*(s), \quad s \geq t + A$ in the integral of equation (15)

Second empirical fact After implementation of the longer T, the *current* value of a novel variety $\upsilon^*(t)$ fully adjusts to its new higher level: Both $\upsilon^*(t)$ and $P^*(t)$ fully incorporate the impact of longer T. Therefore, equation (16) shows that the pace of development goes back to its pre-news levels after the initial drop upon news. However, the latter development slowdown negatively impacts subsequent research productivity through the spillover term $E(t)^\chi = \left[d^{-1}\int_{t-d}^{t} \iota_D^*(s)ds\right]^\chi$, with the effect showing its maximum magnitude $d = 4$ years after the policy implementation. The 4-years delay is chosen to match the observed behavior of within-field backward citations intensity of section 4. Lower productivity depresses research investment despite the higher value $P^*(t)$ of research output, i.e., ideas. Lower research investment and lower innovation persist until *i*) the effect of the technology disclosure externality from the pre-implementation development pace drop ends, and *ii*) the missing mass of ideas caused by lower research investment is gradually replenished.[81]

Long-run implications As the new steady state approaches, the long-run incentives implied by the longer patent term become dominant. Total R&D investment and innovation increase to a higher level. Higher value of ideas $P^*(t)$ reflects higher patent value and promotes higher research investment. More research investment generates a larger mass of projects $N(t)$, which result in a higher flow of innovations $\iota_D^*(t)N^*(t)$ and higher aggregate development investment $\mu\iota_D^*(t)^\theta\upsilon^*(t)N^*(t)$–because the pace of development returns to its pre-news levels after policy implementation. Therefore, the long-run behavior of the model is consistent with standard frameworks. However, the novel ingredients are crucial to understand the positive and normative consequences of patent term changes on innovation and output. As documented in section 3, anticipation and powerful technological spillovers may generate unexpected effects, which would be overlooked in standard frameworks. This motivates the structural estimation of the model to conduct policy counterfactuals and evaluate these forces far from the status quo.

6 Structural estimation

Subsection 6.1 describes the estimation of the key structural parameters of the model, and subsection 6.2 presents the results and the quantitative performance of the model.

6.1 Quantitative assessment of the model

I use a mix of calibration–for ρ, n, g_h, and η–and structural estimation via generalized method of moments–for the other 8 parameters $\phi_1, \phi_2, \theta, \mu, \psi, \chi, \alpha$, and δ–to match the empirical facts.

[81]The key role of the spillover to generate persistence is highlighted by the responses of innovation and R&D flow in a model where the externality channel is shut down. In the latter setting, innovation and R&D flow increase immediately after the implementation of the longer term. Figures F.3 and F.4 of Appendix F.2 illustrate the point.

6.1.1 Calibrated parameters

I calibrate $\rho = 0.03$ and population growth at $n = 1.1\%$, the US yearly average in the post-war period. The growth rate of exogenous productivity g_h is fixed so that the growth rate of output-per-capita $g_y = \frac{1-\phi_1}{1-\phi_1-\phi_2}(n + g_h) - n$ (derivations in Appendix C.1.10) is equal to 2% for every ϕ_1 and ϕ_2. Finally, in the extended model, η is set so that the steady-state block probability $1 - (1 - \zeta_{ss})^{\eta\phi_1}$ is equal to 1%, which is computed by combining information on patent litigation rates (1.5%) and plaintiff win or voluntary settlement rates (65%).[82]

6.1.2 Model-data mapping and estimation

I estimate key structural parameters by generalized method of moments to match *i*) the reduced-form causal DiD estimates for patents and R&D effort of subsection 3.1, and *ii*) 3 long-run moment restrictions. The DiD estimates are used to compute the predicted response of patenting and patent-read R&D effort to a 100-day increase in the patent term, known 2 years and 8 months before implementation–i.e., the lapse of time between November 1992 and June 1995. I re-express the DiD estimates in percentage deviations from the pre-announcement quarter (1992Q3) baseline level of the two variables and use the latter as targets for the simulated response of the model to a policy change of the same size and anticipation. I assume that the system is at its old steady state before news. The model counterpart of granted patents is $\iota_D^*(t)N^*(t)$–i.e. the probability that each project is successfully turned into a patent *times* the number of projects. The model-data mapping of R&D is less immediate. The first reason is that, in the model, research and development are distinct, but I cannot observe them separately in the data. The second reason is that, in field-level aggregate data, R&D effort must be inferred from inventors listed on patents, which reflect all past R&D done on the patent rather than the instantaneous flow of R&D expenditure. Since I cannot make the data more granular, I derive below the theoretical aggregate that best matches the patent-inferred R&D effort observed in the data. Let $n(\tau, \tau) = E(\tau)^\chi V(\tau)^{\phi_1} I_R(\tau)^{\phi_2}$ be the number of new projects generated at time τ by a total research investment $I_R(\tau)$, and let $n(t, \tau) = e^{-\int_\tau^t \iota_D^*(s)ds} n(\tau, \tau)$ be the number of such projects of vintage τ that are not completed yet at time t. The total number of active projects is $N(t) = \int_{-\infty}^t n(t, \tau)d\tau$. The total amount of R&D spent on any project of vintage τ and successfully developed into a patent at time $t \geq \tau$ is

$$r\&d(t, \tau) = I_R(\tau)/n(\tau, \tau) + \int_\tau^t \mu\iota_D^*(s)^\theta \upsilon^*(s)ds$$

[82]Patent litigation rates are taken from Figure S9 of WIPO report "Special theme - An overview of patent litigation systems across jurisdictions" (https://www.wipo.int/edocs/pubdocs/en/wipo_pub_941_2018-chapter1.pdf). Plaintiff win and settlement rates are taken from https://law.stanford.edu/wp-content/uploads/2016/07/Revised-Stanford-August-4-2016-Class-Presentation.pdf

I assume that each project of vintage τ absorbs an equal fraction $1/n(\tau, \tau)$ of the original research investment $I_R(\tau)$. In addition, I sum the total resources spent on the development of the specific project, from its inception at τ to its completion at t. Since, at any instant, all projects have equal completion probability irrespective of their vintage, aggregate R&D inferred from patents generated at t is

$$R\&D(t) = \int_{-\infty}^{t} r\&d(t, \tau) \Big[\iota_D^*(t) n(t, \tau) \Big] d\tau$$

i.e. the aggregation over all vintages of the total investment on projects of vintage τ, weighted by the mass thereof that are successful at time t–i.e., $\iota_D^*(t)n(t, \tau)$.

Finally, the long-run targeted moments are: *i*) a capital-output ratio of 3, *ii*) a consumption-output ratio of 0.65, and *iii*) an R&D expenditure-output ratio of 2.5%. The loss function is quadratic in deviations of the model-simulated moments from their data counterparts.

6.1.3 Identification of structural parameters from empirical moments

The rich dynamics of the reduced-form empirical responses of innovation and R&D to policy shocks are extremely informative about the key structural parameters of the innovation process, i.e. θ, χ, ϕ_1, and ϕ_2. The discussion of subsection 5.5 on how the model can qualitatively replicate the empirical facts is useful to infer which moments are informative about which parameters.

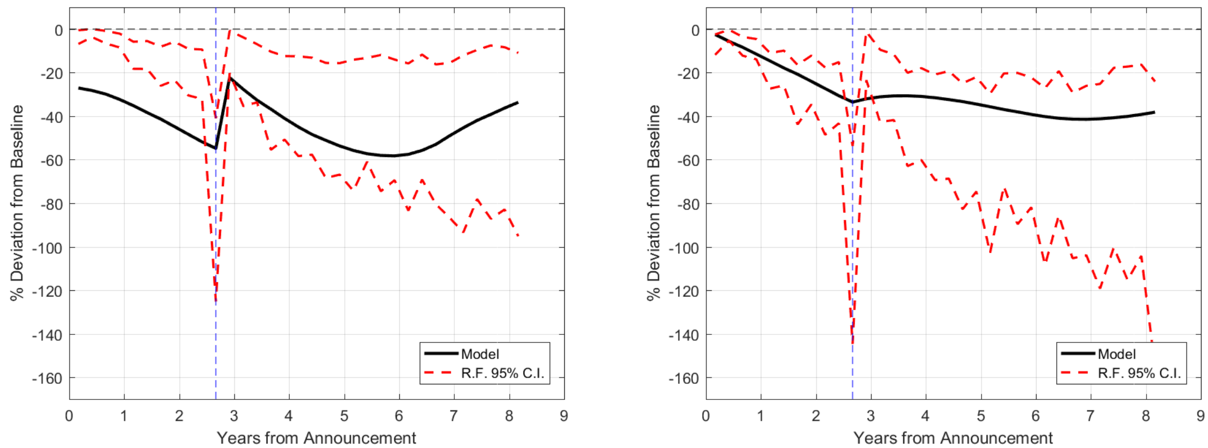
In the model, the adjustment of the pace of existing projects' development drives the reaction of innovation and R&D to news of a future patent term change. Therefore, the strength of the empirical response to the news shock is informative about θ, which governs the cost convexity of development pace. Subsequently, the disclosure spillover translates the initial fall in development into lower research productivity, driving post-implementation persistence with a delay of $d = 4$ years. Therefore, the magnitude of post-implementation negative effect observed in the DiD estimates informs χ, which governs the strength of the new externality. Instead, the speed of the recovery to the new steady state with higher innovation and R&D identifies ϕ_1, that shapes the "standing on the shoulders of giants effects" from the stock of varieties and influences research productivity more smoothly. Finally, for any given θ, χ, and ϕ_1, the long-run R&D-output ratio determines ϕ_2. A lower ϕ_2 implies more severe decreasing returns to research investment and lower aggregate R&D intensity.

The plots of Appendix C.5 illustrate these points by showing how the model-implied responses change when parameters deviate from the optimal estimates of subsection 6.2.

6.1.4 Solution algorithm

Estimation requires solving, for several parameter vectors, the system of equations describing the b.g.p. solution of the model, which includes the delayed differential equation (8). The

Figure 11: **Model-based simulation of the policy and targeted reduced-form estimates**



(a) New patents (b) R&D Investment

The black solid lines are the model-based responses of the model with parameter values reported in Table 2, and the red dashed lines are 95% confidence bands of the reduced form estimates of Section 3. The system is assumed to be at the pre-policy change steady state at $t = 0$, when the news of 100-days increase in protection time implemented after 2 years and 8 months (blue vertical line) happens.

solution algorithm is described in detail in Appendix C.2. In a nutshell, it starts with a guess for the full dynamic path of $V(t)$–and hence for $\lambda(t)$–and it solves the model treating the delay term $\dot{V}(t-T)e^{-\int_{t-T}^{t}\lambda(t')\mathrm{d}t'}$ of equation (8) as fixed from the guess. The guess is then updated, and the process is iterated until the series of $\lambda(t)$ converges.

6.2 Estimation results and quantitative performance

Table 2 reports the estimated parameters and their standard errors, and Figure 11 plots the model-based responses together with the confidence bands of the reduced-form estimates of subsection 3.1.2, divided by the 1992Q3 baseline.[83] The plots show that the model performs well in tracking the response of the two main aggregates. It captures the drop of R&D and innovation at news–due to a fall in the pace of development–and the persistence of the negative effect even after the policy implementation–due to the technology disclosure externality. Subsection 5.5 discusses the details of the mechanism through which this happens. In the long-run–not shown in the figure–innovation and R&D converge to their new higher steady-state levels implied by the longer T.

Some of the parameter estimates of Table 2 offer interesting insights. First, the cost convexity of development intensity is very mild, with θ very close to 1. It implies an optimal steady-state development intensity $\iota^*_{D,ss} =$ 0.3 that is consistent with an average project duration of approximately 3 years. This is slightly longer than the lags estimated by Pakes and

[83]The computation of standard errors is described in Appendix C.2.

Table 2: **Estimated and Calibrated Structural Parameters**

Symbol	Value	S.E.	Parameter	Target/Source
A: Calibration				
ρ	0.03		Discount rate	
g_h	0.011		Exog. Prod, Growth	2% p.c. Output Growth
n	0.011		Population Growth	World Bank
B: Estimation				
α	0.4463	0.6753	Capital Share	
δ	0.0701	0.0071	Capital Depreciation	
ϕ_1	0.6906	2.1484	Research V-Curvature	
ϕ_2	0.0840	0.6282	Research I_R-Curvature	
χ	8.3554	4.9520	Spillover Exponent	
θ	1.0129	0.1990	Dev.'t Curvature	
μ	0.6574	2.9311	Dev.'t M. Cost	
ψ	0.0001	8.5134	Endog. Creative Destruction	
C: Extension				
η	0.1002		Curv. Monopoly Distortion	1% Block Probability

The table reports the calibrated parameters (ρ, n, and g_h) and the structural estimates for the other parameters, obtained by simulated method of moments targeting *i*) the reduced form estimates of the response of granted patents and R&D effort presented in Section 3 and *ii*) three long-run moments: a capital-output ratio of 3, a consumption-output ratio of 0.65, and a R&D investment-output ratio of 0.025. Finally, the parameter η of the extended version of the model featuring blocking innovation is calibrated to match a 1% block probability of new projects. Details on how this target is derived are reported in subsection 6.1.1 of the paper.

Schankerman (1984) but it may reflect that projects have now become more complex and longer to complete on average. Second, ϕ_2 close to 0.1 suggests substantial decreasing returns to research investment, which may explain why models where R&D occurs in a single step usually employ quadratic costs of R&D intensity (e.g. Acemoglu et al. (2018)), capturing a combination of θ close to one and a small ϕ_2. Third, $\phi_1 = 0.69$ implies that returns from existing varieties are mildly decreasing, i.e. that marginal gains in research productivity slowly fall as the number of varieties expands.[84] Finally, the estimated model allows to infer the *long-run* elasticity of innovation and R&D to patent length, which could not be directly desumed from the reduced-form estimates of section 3. A 1% increase of T from 17 years is associated to +0.35% patents and +1.3% of total R&D spending in the new steady state. The implied elasticity of innovation to patent length is similar in size to the elasticity of innovation to market size (Jaravel (2021)), which also increases the incentives to innovate due to higher profits.

7 Policy simulations and normative analysis of patent length

This section conducts the normative analysis of the paper. It uses the structurally estimated model to quantify the key trade-offs and evaluates the welfare and output consequences

[84] This is qualitatively in line with the findings of Bloom et al. (2020), even though they estimate more severe decreasing returns, with ϕ_1 close to 0 for the aggregate economy

of patent term changes of various nature. I focus on two dimensions. First, subsection 7.1 presents the normative trade-off in the steady state of the model, and subsection 7.2 compare it to that arising from the transitional dynamics induced by the *unanticipated* implementation of a new patent length. I estimate the welfare-maximizing patent length in the absence of anticipation and discuss how key structural parameters affect it. Second, subsection 7.4 investigates anticipation and the normative implications of technology disclosure externalities, which subsection 5.5 showed to be crucial to understand the positive consequences of patent term changes. Indeed, due to the action of spillovers, an anticipation of 5 months would be enough for the implementation of the optimal patent term not to generate any output gains.

7.1 Steady state trade-off

The steady-state trade-off is similar in spirit to the classic trade-off highlighted by Nordhaus (1967). In the model, any stationary equilibrium featuring a longer patent term T implies a higher number of varieties V_{ss}, but also a larger share ζ_{ss} of them that are monopolistically produced. The expression of aggregate output in the steady-state equilibrium highlights the two forces:

$$Y_{ss} = \ V_{ss} \ \underbrace{(\alpha^{\alpha}\zeta_{ss} + (1-\zeta_{ss}))}_{\approx 0.9} \ L_0^{1-\alpha} X_{nm,ss}^{\alpha} \tag{18}$$

The first force pushes up output, while the second tends to depress it, because monopolistic varieties are produced in smaller quantity than competitive ones. Therefore, as ζ_{ss} increases, the distortion term in brackets gets closer to zero. However, in the structural model, $(\alpha^{\alpha}\zeta_{ss} + (1-\zeta_{ss}))$ is above 0.9, which implies small distortions and calls for a long patent length. The second column of Table 3 shows that the patent length that would maximize steady-state consumption at the baseline parameter estimates is 171 years (first row).

There is an additional channel through which patent length may affect welfare: A longer monopoly over specific products may block additional innovations if the latter infringe on existing patents. While this feature is not present in the benchmark model of section 5, for normative analysis I also consider a model extension where new projects may be blocked with a probability $1-(1-\zeta(t))^{\phi_1\eta}$, which is increasing in $\zeta(t)$ and whose severity is governed by the parameter η.[85] In the remainder of the paper, I will refer to this alternative setting as the model with "blocking innovation". The fourth column of Table 3 shows the optimal steady-state patent length in this setting. The additional distortion implies that a longer patent length is relatively less desirable, but the estimates remain longer than the status quo

[85]The production function of projects becomes $(1-\zeta(t))^{\phi_1\eta}E(t)^{\chi}V(t)^{\phi_1}I_R(t)^{\phi_2}$, where the new term $(1-\zeta(t))^{\phi_1\eta}$ is a distortion term that represents the probability of *not* having an idea blocked by an existing monopoly. Conversely, $1-(1-\zeta(t))^{\phi_1\eta}$ can be seen as the probability of being blocked and is increasing in $\zeta(t)$.

of 20 years.

7.2 Transitional dynamics trade-off

The second trade-off relates to the transitional dynamics of the model in response to the *unanticipated* implementation of a new patent length. Consider a longer patent length T. Its unanticipated implementation starts a transition to a new steady state that features more varieties V and, potentially, higher output and consumption. The increase of V occurs through research and development investment. However, due to the presence of development lags, the initial investment does not immediately translate into higher productivity and output, as it takes time for new ideas to be transformed into varieties. Therefore, by the resource constraint of the economy, a reduction of consumption must finance R&D at first. The social planner trades-off short-run losses and long-run gains using the time-zero utility of the representative agent, i.e.

$$\Theta = \int_0^\infty e^{-(\rho - g_c^*)t} \tilde{c}(t) dt \tag{19}$$

where $\tilde{c}(t)$ is per-capita consumption in the b.g.p. model, ρ is the discount rate, and g_c^* is per-capita consumption growth in the b.g.p..

To estimate the patent length that would maximize (19) in the absence of anticipation, I simulate a sudden policy change from $T_0 = 17$, the status quo before the TRIPs, to several possible values of T'. I compute the welfare index $\Theta^{(T')}$ along the transition to the new steady state for each T'.[86] and I compute the welfare gain or loss as the percentage deviation of $\Theta^{(T')}$ from $\Theta^{(T_0)} = \tilde{c}_{T_0,ss}/(\rho - g_c^*)$, which is the welfare index under the hypothesis that no policy change occurs. Similarly, I can compute the present discounted value of output and innovation relative to the status quo.

The first column of Table 3 shows in bold the welfare-maximizing patent length in the absence of policy anticipation, and reports in brackets the implied welfare change in percentage of the status quo. The first row shows that, for the baseline parameter values obtained in section 6, the optimal term would be 28 years. The other rows examine how this figure changes when key structural parameters vary, highlighting their importance for normative considerations. First, more severe cost convexity of the pace of development–i.e. higher θ– reduces optimal patent length. With higher convexity, development is slower and productivity gains take more time. Therefore, the initial consumption loss protracts for longer. In addition, a slower pace of development reduces research productivity through the spillover term, lowering the benefit of each unit of the final good spent on research investment relative to consumption. Second, more severe decreasing returns to research–i.e. lower ϕ_2– reduce optimal patent length. The effect of ϕ_2 is strong. Analogously to previous discussion, more

[86] In practice, the transition is simulated for 2,000 years and the welfare integral is cut there.

Table 3: **Optimal patent term - Transitional dynamics vs. steady state**

Specification	Benchmark		Blocking innovation	
	Dynamic	Steady State	Dynamic	Steady State
Baseline	**28** (+0.5%)	171 (+47.7%)	**23** (+0.2%)	61 (+21.5%)
$\theta = 1.001$	**29** (+0.6%)	168 (+47.3%)	**23** (+0.2%)	61 (+21.4%)
$\theta = 1.005$	**28** (+0.6%)	169 (+47.4%)	**23** (+0.2%)	61 (+21.4%)
$\theta = 1.02$	**27** (+0.5%)	174 (+48.0%)	**23** (+0.2%)	62 (+21.6%)
$\theta = 1.05$	**27** (+0.4%)	191 (+49.9%)	**22** (+0.2%)	63 (+22.1%)
$\phi_2 = 0.05$	**16** (+0.0%)	122 (+20.8%)	**15** (+0.0%)	60 (+10.6%)
$\phi_2 = 0.07$	**23** (+0.1%)	145 (+34.7%)	**19** (+0.0%)	61 (+16.5%)
$\phi_2 = 0.10$	**34** (+1.2%)	244 (+67.7%)	**27** (+0.6%)	62 (+28.3%)
$\phi_2 = 0.12$	**43** (+2.5%)	500 (+104.4%)	**32** (+1.6%)	62 (+38.9%)
$\phi_1 = 0.63$	**28** (+0.4%)	144 (+34.7%)	**23** (+0.2%)	69 (+19.3%)
$\phi_1 = 0.66$	**28** (+0.5%)	155 (+40.2%)	**23** (+0.2%)	65 (+20.3%)
$\phi_1 = 0.72$	**28** (+0.6%)	223 (+63.9%)	**23** (+0.2%)	57 (+20.7%)
$\phi_1 = 0.75$	**29** (+0.6%)	500 (+104.4%)	**24** (+0.3%)	53 (+24.0%)

The table reports: *i*) The estimated optimal patent length in the absence of policy anticipation (bold figures, columns 1 and 3); *ii*) the related welfare change along the transitional dynamics arising from its implementation, relative to welfare in the absence of any policy change from $T = 17$ years (in brackets, columns 1 and 3); *iii*) the patent length that would maximize steady state consumption (columns 2 and 4); *iv*) the related consumption change in percentage deviation from steady state consumption at $T = 17$ years (in brackets, columns 2 and 4). The dynamic welfare index is (19). Columns 1 and 2 refer to the benchmark model. Columns 3 and 4 refer to the model with blocking innovation. The rows report the estimates for different values of the structural parameters. In each row, all the other parameters are kept at the values reported in Table 2.

severe decreasing returns lower the benefit of each unit of the final good spent on research relative to consumption. Therefore, the optimal policy tries to induce a shift from the former to the latter. Finally, the third column of Table 3 shows how the optimal patent length would change in the presence of blocking innovation. Due to worse monopolistic distortions, the optimal length is uniformly lower than in the benchmark model.

7.3 Evaluating steady-state and dynamic trade-offs in the data

The specific structure of the model may potentially play a big role in the quantification of the normative trade-offs, putting into question how model-specific previous conclusions are. While addressing this issue empirically would require an entire new paper, I provide below some evidence that the data qualitatively support the quantification implied by the model. I use the NBER CES Manufacturing database to empirically investigate the impact of patent term changes on productivity and welfare.[87] As a measure of welfare, I consider the value of shipments deflator, consistently with the idea that a lower price level reflects lower expendi-

[87] All the details of the empirical strategy and of the results are reported in Appendix B.5.

tures to achieve any utility target. Productivity is instead measured by TFP, estimated as the residual of a 5 factors production function.

To investigate the steady-state trade-off, I use a yearly panel of 6-digit NAICS industries and perform an IV regression where innovation–measured by patents or citations-weighted patents–is instrumented by the policy-induced change in the patent term, interacted with yearly dummies. Results show that 100 more patents per year and industry increase TFP by 3.3% and welfare by 2.7%, with an implied pass-through of TFP gains into lower prices of around 0.83.[88] The effect is similar across innovation measures and confirms the model's predictions: At the status quo, productivity gains from a longer patent term outweigh monopolistic distortions, with a large pass-through of the former into higher consumer welfare.

As to the transitional dynamics trade-off, the model highlights a slow adjustment of productivity and welfare, which is confirmed in the data. I use a yearly DiD analysis that replicates specification (1) at the sectoral level. I study TFP and value of shipments deflator as outcomes. The results show that a longer patent length increases productivity and welfare, but the dynamic effect is very close to zero at first and becomes quantitatively relevant only gradually over time.[89]

7.4 The role of anticipation

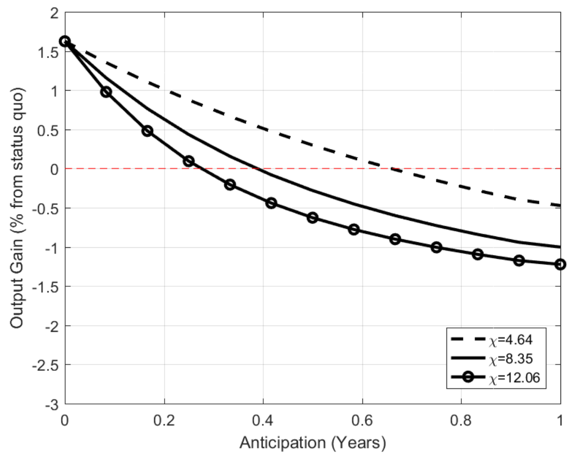
Subsection 7.2 examined the welfare trade-offs implied by the transitional dynamics of the model in response to an *unanticipated* patent term change. However, subsection 5.5 showed that policy anticipation is crucial to understand the impact of patent length on innovation and R&D, due to the presence of powerful technological spillovers. In this final subsection, I investigate the normative consequences of implementing a 28 years patent term in the presence of anticipation. The left (right) panel of Figure 12 shows the change in per-capita consumption (output)–in percentage deviations from the status quo–with policy anticipation between zero and 1 year. To highlight the crucial role of the new technological spillover, I plot consumption and output changes for three values of χ. The solid line refers to the baseline estimate of section 6. With no anticipation, consumption and output increase by 0.5% and 1.6% relative to the status quo, respectively. However, with a 1-month anticipation, almost all the consumption gains would be dissipated and a 1-year anticipation would induce a loss of 1.9%. Similarly, per-capita output change would be null with a policy anticipation of around 5 months, and a 1-year anticipation would generate a loss of 1%. Stronger spillovers (dotted line) would make anticipation even more costly, while weaker spillovers (dashed line) would attenuate the negative effects of news.

The mechanism is analogous to the one described in subsection 5.5. News of a future

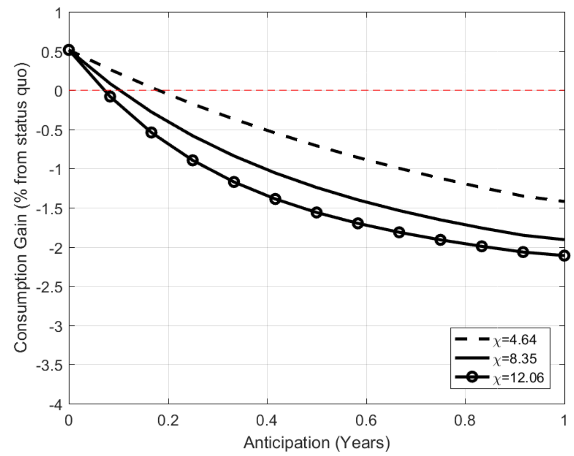
[88] The average number of yearly patents by 6-digit NAICS industry is 280.

[89] All the results are reported in Appendix B.5.2 and in Figures B.45 and B.46 more specifically.

Figure 12: **Change in output and welfare from $T = 28$ years by anticipation**



(a) Output per capita



(b) Consumption per capita

The left (right) panel shows the model-implied change in the present discounted value of per capita output (consumption) due to the *anticipated* implementation of a patent length of 28 years–starting from the 17 years status quo of 1995. The plots show how output and consumption vary by changing the anticipation of the policy. Output and consumption variations are expressed in percentage deviation from the present discounted value of per capita output (consumption) in the absence of any policy change.

patent term increase generates a drop in the pace of development, which immediately translates into less new varieties, lower productivity, and lower output. While the fall in aggregate R&D would initially free-up resources for consumption, the rapid decline of output overturns this effect. Therefore, this analysis reinforces the idea that the evil is the policy details. Output, innovation, and welfare consequences of patent term changes crucially depend on how the new policy is implemented and technology disclosure externalities play a pivotal role.[90]

8 Concluding remarks

While patent length is considered a key policy tool for innovation, economic research has struggled to provide causal estimates of its effects. This paper provides causally-identified evidence of the impact of an *anticipated* change in patent length on innovation and R&D. Policy anticipation, the distinct roles of ongoing projects' development and new projects' research, and powerful technological spillovers are crucial to understand the positive consequences of the policy, which depresses innovation and R&D at news of a future longer term and continues to negatively effect them even after implementation. Due to anticipation, firms decrease the pace of ongoing projects' development at news, as they want to profit of longer protection after policy implementation. Due to a technology disclosure externality, the fall

[90]Policy anticipation is crucial also in other settings. For example, anticipation of a tax increase on durable goods would substantially decrease tax revenues, because consumers would accelerate the replacement of the goods before implementation.

in the pace of development slows publication of patent documents, depressing diffusion of knowledge about recent technological advances. This temporarily hinders other innovators' research productivity in finding novel related ideas. Once controlling for the effect of the externality, the direct effect of a term extension on innovation is weakly positive.

The technology disclosure spillover allows to identify that the elasticity of future innovation to current innovation shocks is 0.997. Once controlling for the effect of the spillover, the direct policy effect implies that the elasticity of patenting to a 1% patent term increase is 2.7 in the short run. In addition, the empirical effect of the news shock allows to estimate that the elasticity of patenting to a 1% future increase of patent length available in 2 years is 3.1.

The paper develops a structural model of semi-endogenous growth that formalizes the distinct roles of development and research. The latter responds to long-run incentives and discovers new abstract ideas. Development subsequently converts ideas into intermediate products, obtaining patents at the end of the process. Intertemporal comparison of the values of patents obtained today relative to patents obtained in the future drives development incentives and the short-run impact of anticipated policies. Moreover, research productivity increases with a faster pace of development through an externality, because a more rapid diffusion of fresh technical knowledge through patent documents helps the creation of new ideas. Thanks to the new R&D structure, the model can replicate the empirical facts. A structural estimation of model's key parameters highlights a mild convexity of development intensity costs but severe decreasing returns to research investment. The long-run elasticities of innovation and R&D to permanent patent term changes are 0.35 and 1.3, respectively. Moreover, both the model and the data feature a high pass-through of productivity gains into consumer's welfare.

The main normative implication of the model is that the welfare-maximizing patent term would be 28 years in the absence of policy anticipation. This term is longer than the one currently in place (20 years). The presence of development lags makes the rich transitional dynamics of the model in response to implementation of the new term crucial for normative considerations. However, evil is in policy implementation details. Even short policy anticipation of the longer 28-years term would generate negative short-run effects on innovation and output, due to development's reaction to news and technology disclosure externalities. This would dissipate all the welfare gains from the optimal policy.

References

Abrams, David S. 2009. "Did TRIPS Spur Innovation? An Empirical Analysis of Patent Duration and Incentives to Innovate." *University of Pennsylvania Law Review*, 157(1): 1613.

Acemoglu, Daron, Ufuk Akcigit, Harun Alp, Nicholas Bloom, and William Kerr. 2018. "Innovation, Reallocation, and Growth." *American Economic Review*, 108(11): 3450–3491.

Aghion, Philippe, and Xavier Jaravel. 2015. "Knowledge Spillovers, Innovation and Growth." *The Economic Journal*, 125(March): 533–573.

Akcigit, Ufuk, Doug Hanley, and Nicolas Serrano-Velarde. 2020. "Back to Basics: Basic Research Spillovers, Innovation Policy and Growth." *Review of Economic Studies*, Forthcoming.

Akcigit, Ufuk, Santiago Caicedo, Ernest Miguelez, Stefanie Stantcheva, and Valerio Sterzi. 2020. "Dancing with the Stars: Innovation through Interactions." *NBER Working Paper No. 24466*.

Bartelsman, Eric J., and Wayne Gray. 1996. "The NBER manufacturing productivity database." *NBER Technical Working Paper 205*.

Bell, Alex, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen. 2019. "Who Becomes an Inventor in America? The Importance of Exposure to Innovation." *Quarterly Journal of Economics*, 134(2): 647–713.

Bloom, Nicholas, Charles I. Jones, John Van Reenen, and Michael Webb. 2020. "Are Ideas Getting Harder to Find?" *American Economic Review*, 110(4): 1104–44.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen. 2013. "Identifying Technology Spillovers and Product Market Rivalry." *Econometrica*, 81(4): 1347–1393.

Boldrin, Michele, and David K. Levine. 2013. "The Case against Patents." *Journal of Economic Perspectives*, 27(1): 3–22.

Budish, Eric, Benjamin N. Roin, and Heidi Williams. 2015. "Do Firms Underinvest in Long-Term Research? Evidence from Cancer Clinical Trials." *American Economic Review*, 105(7): 2044–85.

Budish, Eric, Benjamin N. Roin, and Heidi Williams. 2016. "Patents and Research Investments: Assessing the Empirical Evidence." *American Economic Review*, 106(5): 183–187.

Comin, Diego, and Mark Gertler. 2006. "Medium-Term Business Cycles." *American Economic Review*, 96(3): 523–551.

Doane, Michael. 1994. "TRIPS and International Intellectual Property Protection in an Age of Advancing Technology." *American University International Law Review*, 9(2): 465–497.

Futagamia, Koichi, and Tatsuro Iwaisako. 2007. "Dynamic analysis of patent policy in an endogenous growth model." *Journal of Economic Theory*, 133(1): 306–334.

Galasso, Alberto, and Mark Schankerman. 2015. "Patents and Cumulative Innovation: Causal Evidence from the Courts." *The Quarterly Journal of Economics*, 130(1): 317–369.

Gilbert, Richard, and Carl Shapiro. 1990. "Optimal Patent Length and Breadth." *RAND Journal of Economics*, 21(1): 106–12.

Goldschlag, Nathan, Travis Lybbert, and Nikolas Zolas. 2019. "An 'Algorithmic Links with Probabilities' Crosswalk for USPC and CPC Patent Classifications with an Application Towards Industrial Technology Composition." *Economics of Innovation and New Technology*, 1–21.

Gross, Daniel P., and Bhaven N. Sampat. 2021. "Crisis Innovation Policy from World War II to COVID-19." *NBER Working Paper 28915*.

Hall, Bronwyn H., Adam B. Jaffe, and Manuel Trajtenberg. 2001. "The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools." *NBER Working Paper No. 8498*.

Jaffe, Adam B. 1986. "Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value." *American Economic Review*, 76(5): 984–1001.

Jaravel, Xavier. 2021. "Inflation Inequality: Measurement, Causes, and Policy Implications." *Annual Review of Economics*, 13(1): :599–629.

Jones, Charles I. 1995. "R&D-Based Models of Economic Growth." *Journal of Political Economy*, 103(4): 759–784.

Klemperer, Paul. 1990. "How Broad Should the Scope of Patent Protection Be?" *RAND Journal of Economics*, 21(1): 113–30.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman. 2017. "Technological Innovation, Resource Allocation, and Growth." *The Quarterly Journal of Economics*, 132(2): 665–712.

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

Lemus, Jorge, and Guillermo Marshall. 2018. "When the Clock Starts Ticking: Measuring Strategic Responses to TRIPS's Patent Term Change." *Research Policy*, 47(4): 796–804.

Lerner, Josh. 2009. "The Empirical Impact of Intellectual Property Rights on Innovation: Puzzles and Clues." *American Economic Review*, 99(2): 343–48.

Lin, Hwan, and L. F. Shampine. 2018. "R&D-based Calibrated Growth Models with Finite-Length Patents: A Novel Relaxation Algorithm for Solving an Autonomous FDE System of Mixed Type." *Computational Economics*, 51(1): 123–158.

Lychagin, Sergey, Joris Pinkse, Margaret E. Slade, and John Van Reenen. 2016. "Spillovers in Space: does geography matter?" *Journal of Industrial Economics*, 64(2): 295–335.

Martin, Lisa B., and Susan L. Amster. 1994. "International Intellectual Property Protections in the New GATT Accord." *Journal of Property Rights*, 9.

Matthews, Duncan. 2002. *Globalising Intellectual Property Rights.* Routledge.

Montalvo, Patricia. 1996. "How Will the New Twenty-Year Patent Term Affect You? A Look at the Trips Agreement and the Adoption of a Twenty- Year Patent Term." *Santa Clara High Tech. Law Journal*, 139.

Moretti, Enrico. 2020. "The Effect of High-Tech Clusters on the Productivity of Top Inventors." *Working Paper*.

Moretti, Enrico, Claudia Steinwender, and John Van Reenen. 2019. "The Intellectual Spoils of War? Defense R&D, Productivity and Spillovers." *Working Paper*.

Morgese, Giuseppe. 2009. *L'accordo sull'aspetto dei diritti di proprietà intellettuale attinenti al commercio (TRIPs).* Cacucci Editore.

Moscona, Jacob. 2021. "Flowers of Invention: Patent Protection and Productivity Growth in US Agriculture." *Working Paper*.

Moser, Petra. 2005. "How Do Patent Laws Influence Innovation? Evidence from Nineteenth-Century World's Fairs." *American Economic Review*, 95(4): 1214–1236.

Moser, Petra, and Alessandra Voena. 2012. "Compulsory Licensing: Evidence from the Trading with the Enemy Act." *American Economic Review*, 102(1): 396–427.

Nordhaus, William D. 1967. "The Optimal Life of a Patent." Cowles Foundation for Research in Economics, Yale University Cowles Foundation Discussion Papers 241.

Nordhaus, William D. 1969. *Invention, Growth, and Welfare A Theoretical Treatment of Technological Change.* MIT Press.

Pakes, Ariel, and Mark Schankerman. 1984. "The Rate of Obsolescence of Patents, Research Gestation Lags, and the Private Rate of Return to Research Resources." *R&D, Patents and Productivity*, Chicago: University of Chicago Press: 98–112.

Pakes, Ariel, and Mark Schankerman. 1986. "Estimates of the Value of Patent Rights in European Countries During the Post-1950 Period." *Economic Journal*, 1052–1076.

Reichman, Jerome H. 1993. "The TRIPs Component of the GATT's Uruguay Round: Competitive Prospects for Intellectual Property Owners in an Integrated World Market." *Fordham Intellectual Property, Media & Entertainment Law Journal*, 9: 171–266.

Romer, Paul. 1990. "Endogenous Technological Change." *Journal of Political Economy*, 98(5): S71–102.

Sakakibara, Mariko, and Lee Branstetter. 2001. "Do Stronger Patents Induce More Innovation? Evidence from the 1988 Japanese Patent Law Reforms." *RAND Journal of Economics*, 32(1): 77–100.

Schankerman, Mark, and Florian Schuett. 2017. "Screening for Patent Quality." *CEPR Discussion Paper 11688*.

Appendices

Appendix A Data description

A.1 Data sources

For the empirical analysis, I rely on six data sources. The first is PATSTAT, which I use to build the technical field-level innovation and treatment variables. The second is the NBER Patent Database, which contains patent information and provides a match of applicants to firm identifiers in COMPUSTAT. The NBER Patent database reports rich patent information for patents that are filed at the USPTO between 1976 and 2006 and it provides harmonized applicants' identifiers which have been dynamically matched to COMPUSTAT identifiers. This allows to merge patent information with firm-level balance sheet information. The third source of data is COMPUSTAT, that reports balance sheet and financials for listed US firms. These are a selected subsample of innovating firms that are nonetheless responsible for a relevant share of the aggregate US GDP, R&D and innovation. The fourth source of data is the economic value of patents as taken from Kogan et al. (2017). The Authors use a long historical sample of patent records and match applicants to listed firms for which stock market data are available in the CRSP database. They estimate the private economic value of patents by exploiting the stock market reaction to patent grants. The fifth source of data for sectoral analyses is the NBER CES manufacturing database, which is jointly built by the National Bureau of Economic Research (NBER) and U.S. Census Bureau's Center for Economic Studies (CES). It contains annual industry-level data for 1958-2011 and by 6-digit NAICS industries. The sixth data source the is 'Algorithmic Links with Probabilities' crosswalk by Goldschlag, Lybbert and Zolas (2019). It maps technical field (4-digit IPC classes) into industries (6-digit NAICS) and viceversa. The probabilistic links are based on text-analysis of patents abstracts and descriptions the activity of different sectors. I refer to the Online supplementary materials D for details on variables construction.

A.2 Summary Statistics

Summary statistics by technical field and at the firm-level are reported below. Online supplementary material D reports summary statistics by 6-digit NAICS industry.

Table A.1: **Summary statistics by technical field and quarter**

Variable	Mean	S.D.	10th p.	90th p.
Granted Patents	36.09	136.35	0	78
5 years Citations	195.15	1070.1	0	360
Patent value (Million Dollars)	351.01	3341.87	0	406.01
Pending Days (days)	1022.1	496.2	558.93	1660.94
Change in patent length (days)	472.66	117.42	343.55	590.79
Standard dev. of Δ patent term (days)	37.13	38.93	11.53	72.55
Patents share w. p.p.>3y.	.06	.08	0	.12
Share of patents renewed to max. term	.29	.26	0	.63
Share of entrant applicants	.54	.21	.28	.82
Share of patents granted to entrants	.49	.23	.2	.8
Patents-based HHI	1191.51	1982.33	116.35	2800
N. Patent with w.-field bckwd. cit.s	4.89	27.07	0	10
Sh. Patent with w.-field bckwd. cit.s	.19	.28	0	.59
Patents at EPO	1.47	4.65	0	4
Avg. Pending Period at EPO (days)	1702.43	272.63	1386.5	2012.84
Share of second filing applications	.56	.15	.37	.73

The Table reports the summary statistics of different variables used in the paper by technical field (4-digit IPC class) and quarterly date

Table A.2: **Summary statistics by firm and year**

Variable	Mean	S.D.	10th p.	90th p.
Granted Patents	13.57	94.75	0	12
Citations-weighted Patents	199.02	1589.75	0	173.72
Patents value (Million Dollar)	287.16	3434.04	0	74.8
Expected change in protection time	445.02	118.57	273.29	571.4
Sales (Million Dollar)	2337.66	9997.57	2.59	4456.27
Age	14.73	13.88	1	37
Employment (Thousands)	10.85	42.01	.05	22.92
R&D Expenditure (Million Dollar)	60.87	359.06	0	58.78

The Table reports the summary statistics of different variables used in the paper by firm (COMPUSTAT firms) and year

Table A.3: **Summary statistics by industry and year**

Variable	Mean	S.D.	10th p.	90th p.
Granted Patents	199.23	591.3	.44	491.7
Citations-weighted Patents	1433.73	6930.68	1.4	2656.01
Patents value (Million Dollar)	3015.86	16684.91	.38	4595.21
Expected change in protection time	474.27	87.17	377.46	564.72
Avg. TFP Growth (p.p.)	.39	6.43	-6.09	6.86
Avg. Inflation (p.p.)	1.89	4.87	-1.53	6.26

The Table reports the summary statistics of different variables used in the paper by industry (6-digit NAICS) and year

Table A.4: **Correlation between Avg. Pending Period and Field-specific Characteristics**

Variable	Correlation	Weighted Corr.
Number of Applications	.13	
Number of Second Filings	.18	
Perc. Growth of Patents	-.02	
Number of First Grants	.08	
Patents at EPO	.14	
Avg. Pending Period EPO	.27	.44
Share of Second Filings	.3	.3

The first column reports the simple correlation between the average ex-ante pending period by field and several average characteristics of the field. The second column reports the same correlations, weighted by the field-specific number of patents.

Appendix B Additional empirical evidence

B.1 Differences with Abrams (2009) explained

The aim of this subsection is to show how the different results obtained in Abrams (2009) and in this paper can be reconciled in light of the different assumptions made about the timing of the policy–unanticipated in Abrams (2009) and anticipated in my setting–and in terms of the specification and sample restrictions adopted. In particular, Abrams (2009) *i*) assumes no policy anticipation, *ii*) cuts the sample to a narrow window of data (6, 12, or 24 months) around the policy implementation, and *iii*) employs a static DiD specification that includes a field-specific liner trend. The analyses below show that the combination of these three factors can lead to a violation of the parallel trends assumption underlying the DiD exercise and bias the estimated coefficients.

As a preliminary step, I replicate Abrams (2009)'s results in my data. Table B.1 shows the estimates of the static DiD specification

$$Y_{j,t} = \alpha_j + Post_t + \delta T_j + \beta Post_t T_j + \chi_j t + \mathbf{X}_{j,t}\gamma + \varepsilon_{j,t} \quad (20)$$

which is equivalent to specification (2) in Abrams (2009). As in the main analysis, j indexes technical fields–a 4-digit IPC class in my setting–while t identifies a specific month, consistently with Abrams (2009). So $Y_{j,t}$ is either the number of patents filed in month-t and field-j, or number of citations-weighted patents. α_j are field fixed effects, $Post_t$ is a dummy variable that takes value 1 if month t comes after the policy implementation of June 1995 and 0 otherwise, T_j is the policy-induced change in protection time described in subsection 2.3.1 of the paper, $\chi_j t$ identifies a field-specific monthly linear trend, $\mathbf{X}_{j,t}$ are field-specific controls including the average number of inventors per patent and the average number of claims per patent, and $\varepsilon_{j,t}$ is the error term. The difference-in-difference coefficient of interest is β, i.e. the coefficient of the interaction term between the treatment and the post-implementation dummy variable. As in Abrams (2009), specification (20) is estimated on the period April 1994 to July 1996, i.e. a 12 months time window around the policy implementation of June 1995, excluding a 2-months inner window around the policy event because of bunching.

Columns (1) and (2) of Table B.1 refer to the number of granted patents as outcome and show that the results are consistent with the sign of Abrams (2009)'s estimates. The magnitude is smaller in my replication because *i*) Abrams (2009) restricts the sample to technical fields with at least thirty patents in every year, and *ii*) defines a technical field as a USPC class, which is slightly broader definition than the 4-digit IPC I use.[91] These two differences imply that the baseline average number of patents in each field is smaller in my sample than in Abrams (2009) and, therefore, marginal effects are smaller too. Indeed, when I also restrict the sample to fields with at least thirty patents in every year, I get coefficients estimates that are approximately four times bigger than the ones reported in Table B.1, and that are in line with the original ones by Abrams (2009). Columns (3) and (4) refer instead to citations-weighted patents as outcome.

As a second step of the analysis, I will try to show how the interaction of the timing assumptions of Abrams (2009) with the employed specification tend to cause problems with the parallel trends assumption of the DiD exercise, and to deliver estimates of β which are biased upward. As an illustrative example, I take two technical fields, C12R and A01H, which have a negative expected change in protection time (-75 days) and a slightly positive change in patent term (+50 days), respectively. Figure B.1 plots the number of granted patents in the two technical fields over the period 1990-2000. The first vertical line refers to November 1992–the date of the Blair House Accord, which I use as reference date for the policy "news"; the second vertical line refers to December 1994, when the TRIPs were formally adopted in

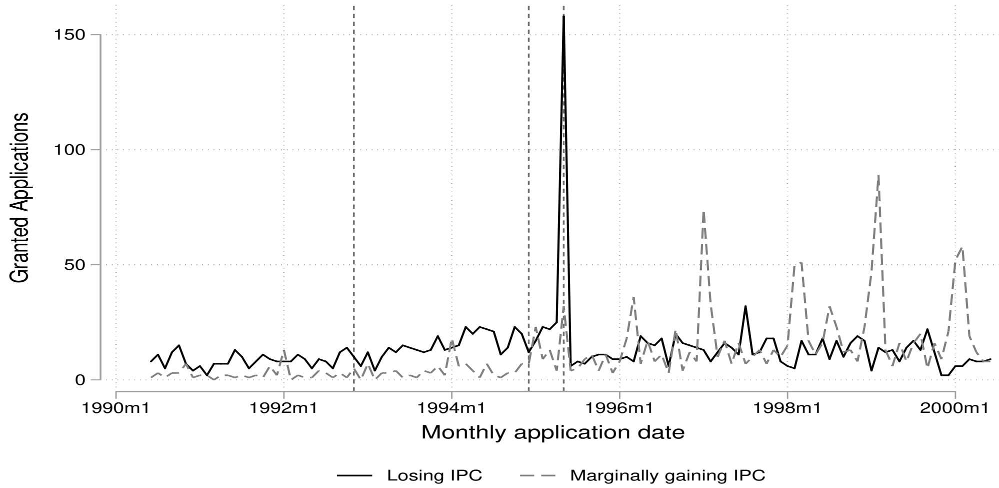
[91] Overall, there are more than 600 4-digit IPC classes, but only 400 USPC classes.

Table B.1: **Replication of Abrams (2009)'s results**

	(1) Patents	(2) Patents	(3) Citations	(4) Citations
$Post_t$	-12.238***	-17.261***	-55.151**	-76.602**
	(3.106)	(4.051)	(22.098)	(31.091)
$Post_t \times Tj$	0.020***	0.029***	0.072*	0.109*
	(0.006)	(0.008)	(0.041)	(0.059)
Avg. Num. of Inventors		0.216***		4.121***
		(0.048)		(0.539)
Avg. Num. of Claims		-0.005		0.537***
		(0.004)		(0.080)
Constant	14.640***	16.980***	108.129***	111.448***
	(0.200)	(0.235)	(1.763)	(2.372)
Field F.E.	Y	Y	Y	Y
Field-specific Trend	Y	Y	Y	Y
Observations	14904	12603	14904	12603

Columns (1) and (2) report the OLS estimates of specification (20) using granted patents filed in month-t and classified in field-j as dependent variable. Columns (3) and (4) report the OLS estimates of specification (20) using citations-weighted granted patents filed in month-t and classified in field-j as dependent variable. Standard errors are clustered by technical field. Statistical significance levels: $^{*}(p < 0.05),^{**} (p < 0.01),^{***} (p < 0.001)$

Figure B.1: **Number of monthly patents in a losing and a gaining field**



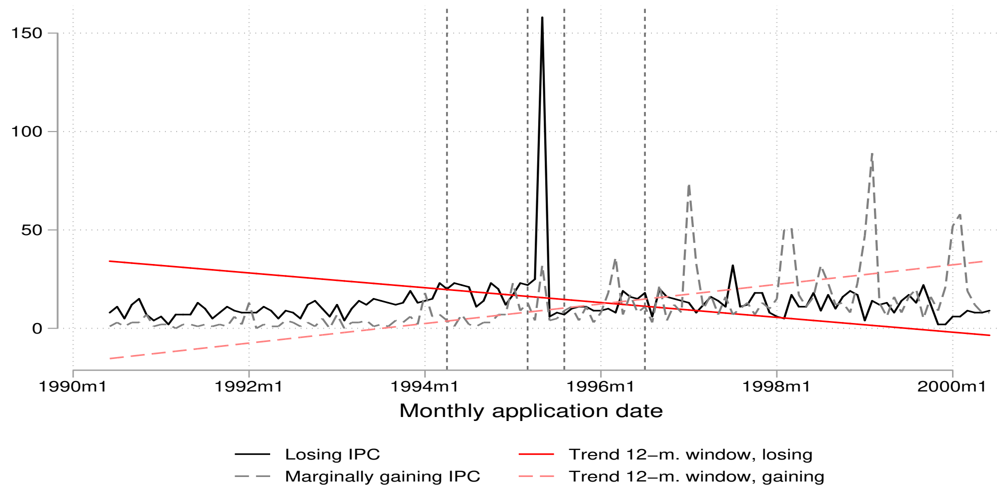
The plot shows the time series of granted patents applied for in month-t and classified in field C12R, the "losing field", and field A01H, the "marginally gaining field".

the US; and the third line refers to June 1995, the date of the policy implementation. From the Figure, it is quite evident that patenting starts accelerating in the field losing protection well before December 1994. Figure B.2, instead, shows the implications of this anticipation for the estimated field-specific time trend in Abrams (2009)'s specification. Here the vertical lines refer to the bounds of the outer 12-months estimation window used in Abrams (2009), with an inner gap of 2 months before and after June 1995. The red solid and dashed lines are the estimated trends for the field losing and gaining protection, respectively. It is immediately clear from the Figure that these trends do not capture the long-run behavior of patenting in the two fields.

This has relevant consequences for the interpretation of the β DiD estimate of specification (20). Indeed, by the Frisch-Waugh-Lowell theorem, we could get β from the "residuals" regression of the outcome variable and the regressors on a field-specific linear trend. In practice, we would like to check that the pre-trends assumption underlying the DiD exercise holds in the residuals of the patenting outcomes from the estimated linear trend. The time series of these residuals is plotted, for the two fields of interest, in Figure B.3. It is clear from the plot that, while the parallel trend assumption seems to hold in the raw data–as confirmed by section 3 of the paper –, the same is not true when focusing on the trend-deviations. This clearly undermines any causal interpretation of the static difference-in-difference estimates in Abrams (2009).

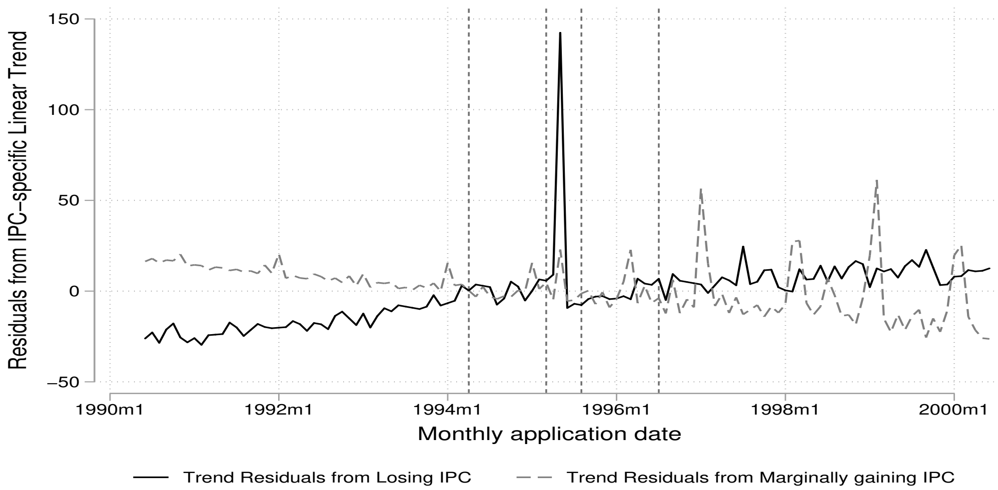
The main issues with Abrams (2009)'s assumptions seem to be two. The first is that the assumption of no policy anticipation is very restrictive. As argued in the paper, if the latter assumption is not true in the data–as it seems to be the case–then the static specification

Figure B.2: **Number of monthly patents in a losing and a gaining field - Fitted trends from Abrams (2009)**



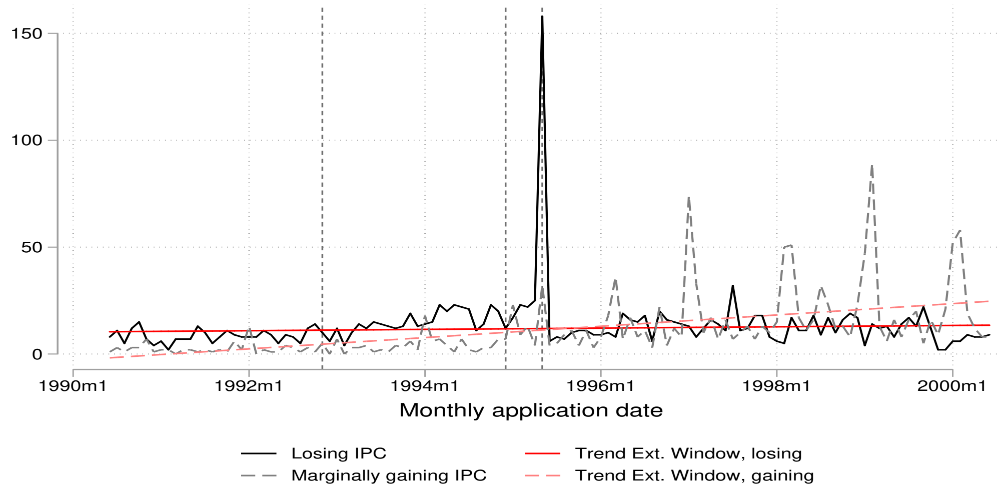
The plot shows the time series of granted patents applied for in month-t and classified in field C12R, the "losing field", and field A01H, the "marginally gaining field". In red, it also plots the fitted field-specific time trends implied by Abrams (2009) specification and sample restriction.

Figure B.3: **Number of monthly patents in a losing and a gaining field - Trend deviations**



The plot shows the time series of the deviation of granted patents from a field-specific linear trend on a time-window of data corresponding to April 1994 to April 1995 and from August 1995 to July 1996. Field C12R is the "losing field" and field A01H is the "marginally gaining field". The fitted field-specific time trends are those implied by Abrams (2009) specification and sample restriction.

Figure B.4: **Number of monthly patents in a losing and a gaining field - Fitted trends on extended sample**



The plot shows the time series of granted patents applied for in month-t and classified in field C12R, the "losing field", and field A01H, the "marginally gaining field". In red, it also plots the fitted field-specific time trends obtained on an extended sample compared to Abrams (2009). The new sample overs the periods June 1990 to November 1994 and December 1995 to June 2000.

(20) delivers biased estimates, because the chosen reference level for the first "diff" is also affected by the policy. Assuming anticipation is a more conservative assumption is this respect. Particularly when using a dynamic DiD specification such as equation (1), the estimated dynamic DiD coefficients for the pre-implementation period can flexibly capture any reactions to the news or the absence thereof. The second issue is related to the choice of a sample that uses data just in a narrow time-window around the implementation date. This implies that the fitted field-specific trend does not capture well the actual evolution of field-specific patenting farther away from the policy. As a result, the parallel trends assumption does not hold for patenting data in trend-deviations, making the cross-fields comparison misleading, because fields bad counterfactuals of one-another. The simplest possible fix to the problems above is to extend the estimation window, so that the fitted field-specific trend gives a better representation of the general behavior of the series, reducing concerns related to the violation of the parallel trends assumption. Therefore, Figure B.4 replicates the plot of Figure B.2, but it extends the sample from June 1990 to November 1994, and from December 1995 to June 2000. So, both the outer and the inner window are expanded. The rationale for expanding also the inner window is to further reduce anticipation concerns, excluding the 6 months between the formal signing of the URAA (December 1994) and the implementation of the policy. As Figure B.4 shows, the fitted trends represent much better the behavior of the series in the sample.

Encouraged by this experiment, I replicate the estimation of specification (20) used by

Table B.2: **Replication of Abrams (2009)'s results - Extended Sample**

	(1) Patents	(2) Patents	(3) Citations	(4) Citations
$Post_t$	1.357***	1.975***	122.192***	170.237***
	(0.440)	(0.568)	(8.691)	(11.239)
$Post_t \times T_j$	-0.003***	-0.005***	-0.226***	-0.320***
	(0.001)	(0.001)	(0.018)	(0.023)
Avg. Num. of Inventors		0.152***		3.489***
		(0.045)		(0.900)
Avg. Num. of Claims		0.002		0.649***
		(0.005)		(0.093)
Constant	13.444***	15.681***	110.019***	114.221***
	(0.059)	(0.130)	(1.168)	(2.574)
Field F.E.	Y	Y	Y	Y
Field-specific Trend	Y	Y	Y	Y
Observations	72657	60969	72657	60969

Columns (1) and (2) report the OLS estimates of specification (20) using granted patents filed in month-t and classified in field-j as dependent variable. Columns (3) and (4) report the OLS estimates of specification (20) using citations-weighted granted patents filed in month-t and classified in field-j as dependent variable. The sample is extended to the period June 1990 - November 1994 and December 1995 - June 2000. Standard errors are clustered by technical field. Statistical significance levels: $^*(p < 0.05),^{**}(p < 0.01),^{***}(p < 0.001)$

Abrams (2009) on the extended sample June 1990 - November 1994 and December 1995 - June 2000. Table B.2 reports the results and shows that, once correcting for the problems outlined above, the DiD coefficient changes sign compared to Abrams (2009) analysis, and becomes coherent with the reduced-form estimates of section 3 of the paper.

B.2 Technical field-level analyses

B.2.1 Private economic value of patents

Results of specification (1) with patent value as dependent variable are in Figure B.5.

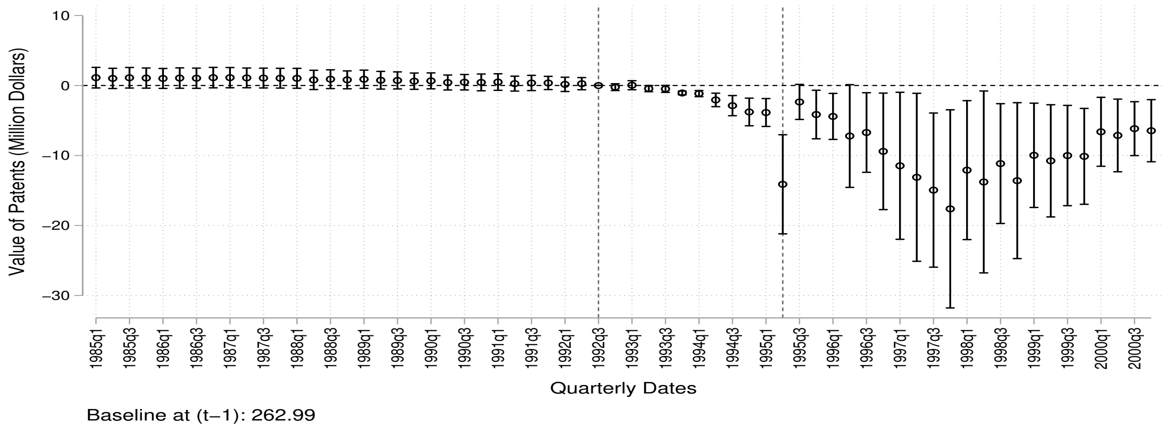
B.2.2 Claims-weighted patents

Results of regression (1) with claims-weighted patents as outcome are in Figure B.6.

B.2.3 R&D effort by technical field

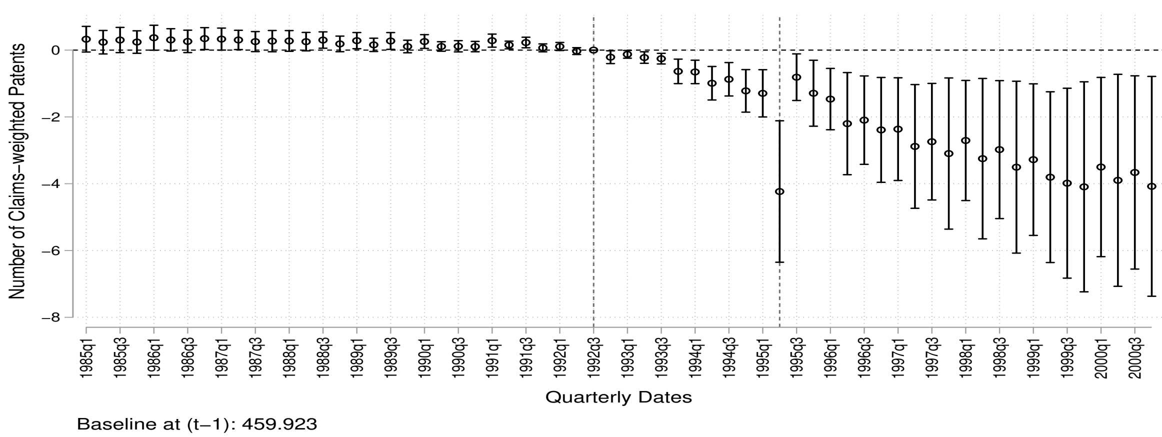
Figure B.7 plots the β_k coefficients of the difference-in-difference specification (1) having as dependent variable the number of inventors listed on patents filed in a given quarter t and classified in field j. The variable corrects for double-counting of inventors listed on more than one patent in the same field and quarter.

Figure B.5: **Effect of 1 more day of protection on patents value**



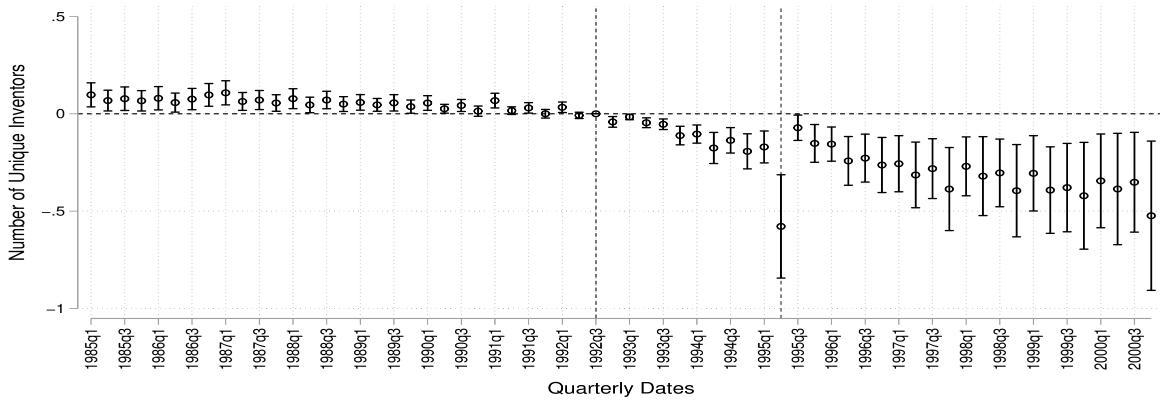
The plot shows the β_k coefficients of specification (1). Dependent variable is quarter-t and field-j dollar value of granted patents built from Kogan et al. (2017). Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure B.6: **Effect of 1 more day of protection on claims-weighted patents**



The plot shows the β_k coefficients of specification (1). Dependent variable is quarter-t and field-j claims-weighted granted patents. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure B.7: **Marginal effect of 1 more day of protection on number of inventors**



The plot shows the β_k coefficients of specification (1) having as dependent variable quarter-t and field-j number of unique inventors. Standard errors are clustered by technical field and 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

B.2.4 Reaction of the average pending period to the policy

Figure B.8 plots the β_k coefficients of the difference-in-difference specification (1) having as dependent variable the average pending period of patents filed in a given quarter t and classified in field j. The results show that the policy event has no impact on the average pending period by field in my data.

B.2.5 Extension of the analysis to 2010Q4

Results of specification (1), estimated on a sample extended to 2010Q4 and having citations-wieghted patents as dependent variable, are in Figure B.9.

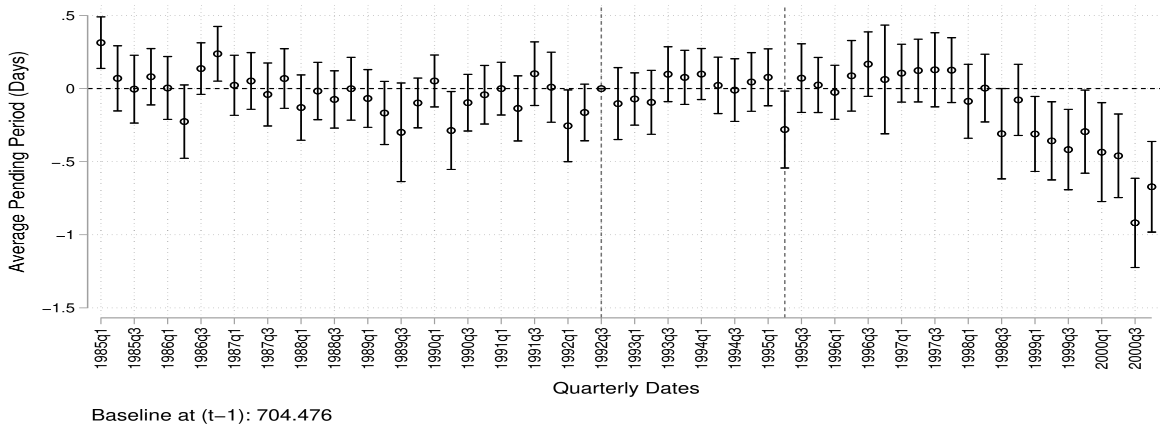
B.2.6 Triple difference with EPO patents

In order to assess endogeneity concerns motivated by the fact that the TRIPs were not limited to the change of the patent term, I run a triple difference specification that tries to clean potential field-specific trends in innovation by controlling for the evolution of patents filed at the European Patent Office in a given quarter-field. The specification is:

$$\begin{aligned} P_{r,j,t} = \psi_r + \alpha_j + \kappa \mathbf{1}_{(r=US)} T_j + \sum_{k=1985Q1}^{2000Q4} \gamma_k \mathbf{1}_{(t=k)} + \sum_{k=1985Q1}^{2000Q4} \eta_k \mathbf{1}_{(t=k)} \mathbf{1}_{(r=US)} \\ + \sum_{k=1985Q1}^{2000Q4} \theta_k \mathbf{1}_{(t=k)} T_j + \sum_{k=1985Q1}^{2000Q4} \beta_k \mathbf{1}_{(t=k)} T_j \mathbf{1}_{(r=US)} + \varepsilon_{j,t} \end{aligned} \tag{21}$$

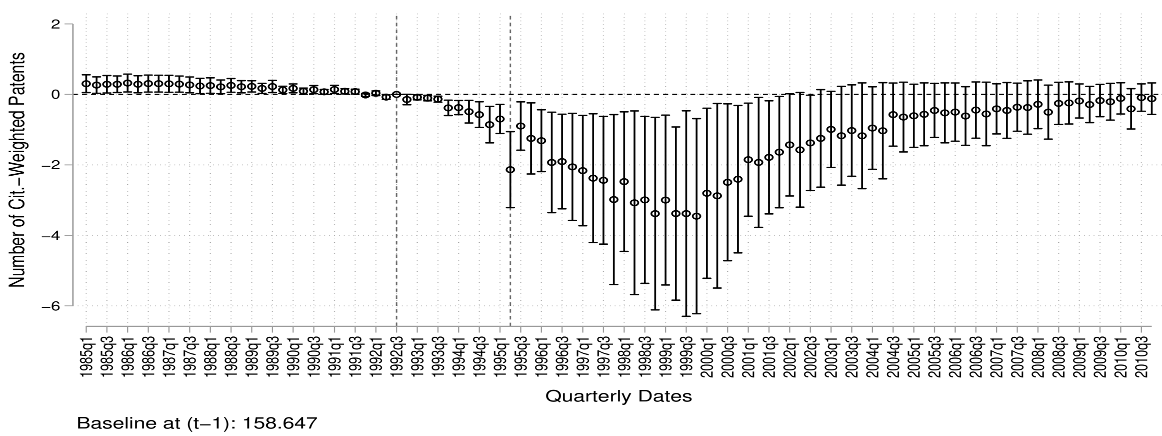
where $P_{r,j,t}$ is region-r, quarter-t, and field-j number of granted patents, T_j is the field-specific treatment described in Section 2.3.1, and $\mathbf{1}_{(r=US)}$ is a dummy variable taking value 1 if the region considered is the US. The motivation for this triple-difference specification is that European innovators were exposed to the same TRIPs-related changes as the US, particularly

Figure B.8: **Average pending time around the treatment date**



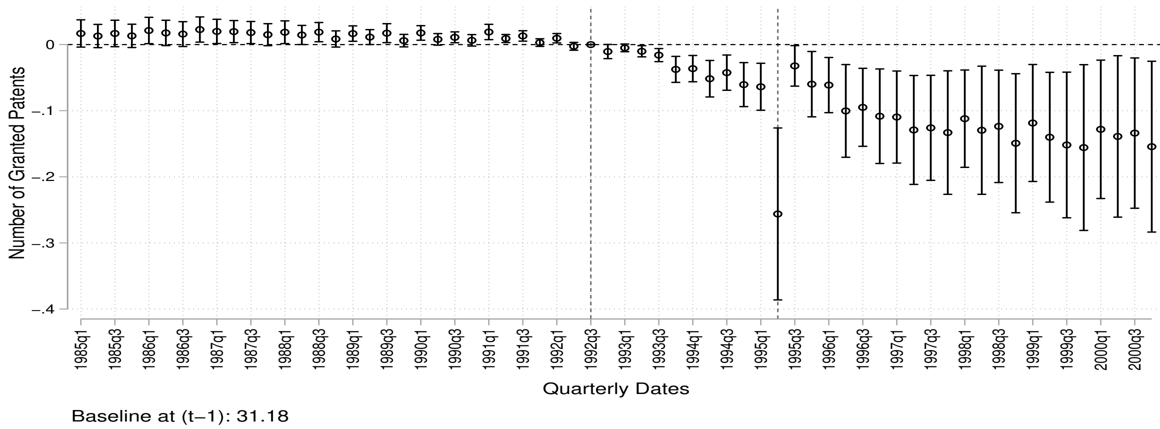
The plot shows the β_k coefficients of specification (1) having as dependent variable quarter-t and field-j average pending period. Standard errors are clustered by technical field and 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure B.9: **Effect of 1 more day of protection on cit.s-weighted patents**



The plot shows the β_k coefficients of specification (1) having as dependent variable quarter-t and field-j 5-years citations-weighted patents. The sample is extended to 2010Q4. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure B.10: **Marginal effect of 1 more day of protection on granted patents - Triple difference specification**



The plot shows the β_k coefficients of specification (21) having as dependent variable region-r's, quarter-t, and field-j number of granted patents. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

regarding access to developing countries. However, the only difference between the two regions was that the patent term in Europe was unaffected by the TRIPs. I omit the dummy for 1992Q3, which is the pre-treatment quarter. Standard errors are clustered by technical field, and 95% confidence bands are plotted. Figure 5 shows the results, which are fully consistent with the evidence of Section 3.

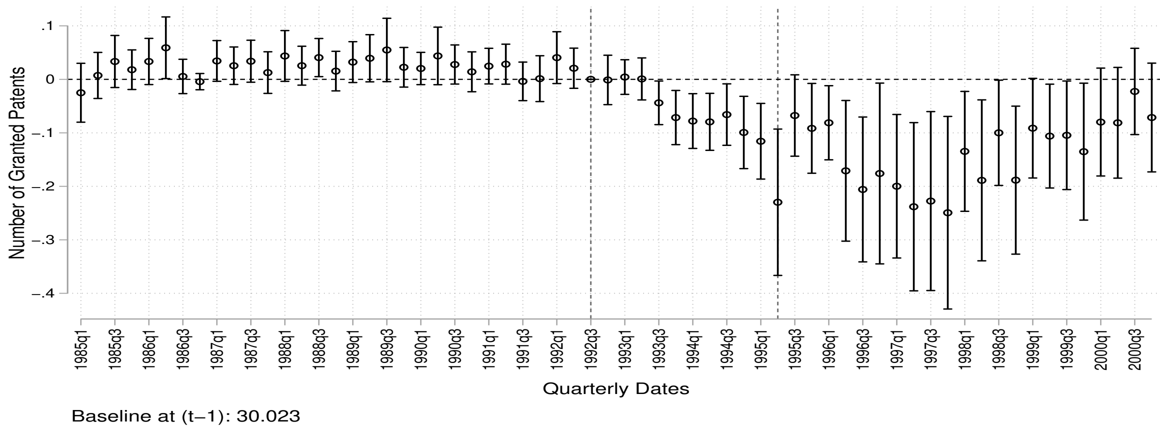
B.2.7 ex-post effective treatment instrumented by ex-ante treatment

This subsection presents the results of an IV regression where the treatment is computed based on the ex-post realized average pending period for patents filed in quarter-t and field-j. The latter is instrumented by the field-specific treatment–based on the ex-ante average pending period, as described in subsection 2.3.1–interacted with quarterly dummy variables. This analysis should address the concern that the policy-induced treatment computed using the ex-ante pending period is not representative of the ex-post effective change in patent length. Therefore, the analysis uses the latter in the second stage regression, and it employs the former to induce plausibly exogenous variation in the ex-post patent term change. As a result, variation still comes from the original treatment of subsection 2.3.1 but the β_k coefficients measure now the response of innovation outcomes to a 1-day change in *ex-post effective* protection time.

The specification of the second stage regression is

$$Y_{j,t} = \alpha_j + \sum_{k=1985Q1}^{2000Q4} \gamma_k \mathbf{1}_{(t=k)} + \sum_{k=1985Q1}^{2000Q4} \beta_k \mathbf{1}_{(t=k)} \tilde{T}_{j,t} + \varepsilon_{j,t} \tag{22}$$

Figure B.11: **Marginal effect of 1 more day of effective ex-post protection change on granted patents**



The plot shows the β_k coefficients of specification (22) having as dependent variable quarter-t and field-j number of granted patents. I omit the dummy for 1992Q3, which is the pre-treatment quarter. Standard errors are clustered by technical field and 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

where all the variables have the same meaning as in specification (1), and $\tilde{T}_{j,t}$ is the treatment based on the ex-post, realized effective average pending period computed for patents filed in quarter-t and classified in field-j. In turn, the first stage regressions are

$$\mathbf{1}_{(t=k)}\tilde{T}_{j,t} = \eta_j + \sum_{k=1985Q1}^{2000Q4} \psi_k \mathbf{1}_{(t=k)} + \sum_{k=1985Q1}^{2000Q4} \delta_k \mathbf{1}_{(t=k)} T_j + u_{j,t} \quad \forall k \tag{23}$$

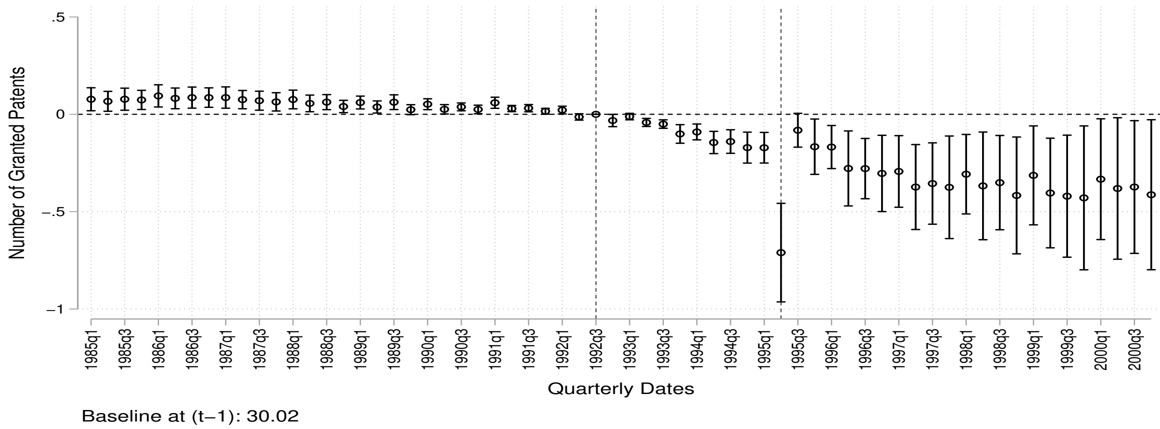
where T_j is the usual treatment variable based on the ex-ante average pending period. Figure B.11 plots the β_k coefficients of (22) having as dependent variable the number of granted patents. Results are fully consistent with the main evidence of section 3.[92]

B.2.8 Triple difference analysis with the standard deviation of the pending period

In this regression, I use a triple difference specification interacting the main treatment variable T_j–i.e. the expected change in protection time (in days) for field j–with a dummy variable taking value 1 if the standard deviation of the average pending period by technical field–as computed using patents granted before the policy news–is above the median value across technical fields. The aim of this specification is to corroborate the idea that the effects estimated in the benchmark specification are genuinely related to the TRIPs-induced patent term change, rather than to other factors. This would be the case if we observe that the magnitude of the estimated coefficients increases when the "signal" from the policy is more precise–i.e. the standard deviation of the average protection change is smaller. The specification of the

[92] This is the case for citations-weighted patents too. Results are not reported for space constraints.

Figure B.12: **Marginal effect of 1 more day of protection on granted patents - Triple difference specification**



The plot shows the β_k coefficients of regression (24) having as dependent variable $P_{j,t}$, i.e. quarter-t and field-j number of granted patents. I omit the dummy for 1992Q3, which is the pre-treatment quarter. Standard errors are clustered by technical field and 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

regression is

$$\begin{aligned} Y_{j,t} = \alpha_j + d_{\sigma_j \leq \sigma^m} + \sum_{k=1985Q1}^{2000Q4} \gamma_{1,k} \mathbf{1}_{(t=k)} + \sum_{k=1985Q1}^{2000Q4} \gamma_{2,k} \mathbf{1}_{(t=k)} d_{\sigma_j \leq \sigma^m} + \\ \chi T_j d_{\sigma_j \leq \sigma^m} + \sum_{k=1985Q1}^{2000Q4} \psi_k \mathbf{1}_{(t=k)} T_j + \sum_{k=1985Q1}^{2000Q4} \beta_k \mathbf{1}_{(t=k)} T_j d_{\sigma_j \leq \sigma^m} + \varepsilon_{j,t} \end{aligned} \tag{24}$$

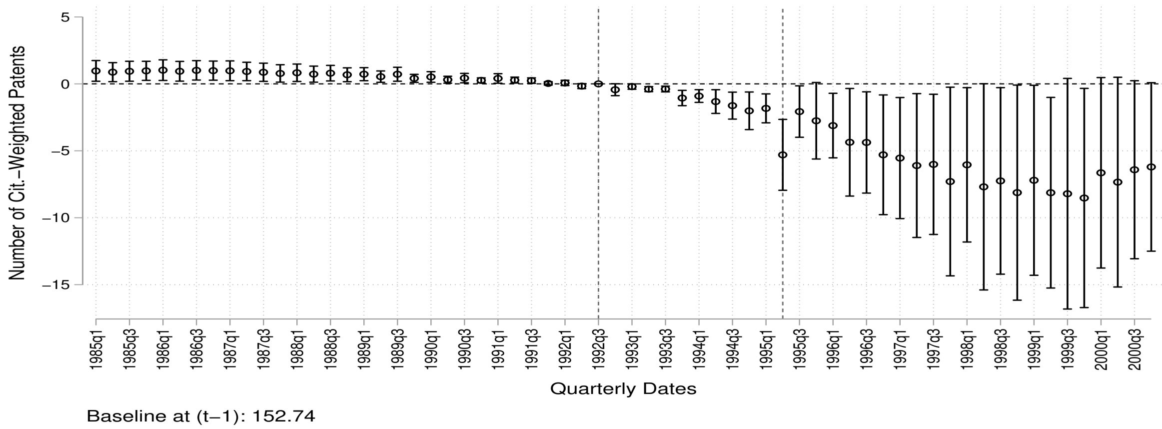
where all the variables follow the usual notation, σ_j is the field-specific standard deviation of the pre-policy-news pending period, and σ^m is the median value of such standard deviation across technical fields.

Figure B.12 and B.13 plot the $\hat{\beta}_k$ coefficients from previous regression, having as dependent variables the number of granted applications and citations-weighted patents, respectively. The negative coefficients mean that the negative magnitude of the baseline DiD specification is stronger when $d_{\sigma_j \leq \sigma^m} = 1$, i.e. when the standard deviation of the average pending period is below the median and the policy-induced treatment is more precise.

B.2.9 IV strategy

The IV strategy aims to address the potential endogeneity of the average ex-ante pending period across technical fields. I link heterogeneous pending periods by technical field to two factors. The first is differential congestion of the technical offices that examine patents from different fields. The second is the different technical difficulty of the patent examination process in different technical areas. To capture presumably exogenous variation in the average

Figure B.13: **Marginal effect of 1 more day of protection on citations-weighted granted patents - Triple difference specification**



The plot shows the β_k coefficients of regression (24) having as dependent variable $C_{j,t}$, i.e. quarter-t and field-j number of citations-weighted granted patents. I omit the dummy for 1992Q3, which is the pre-treatment quarter. Standard errors are clustered by technical field and 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

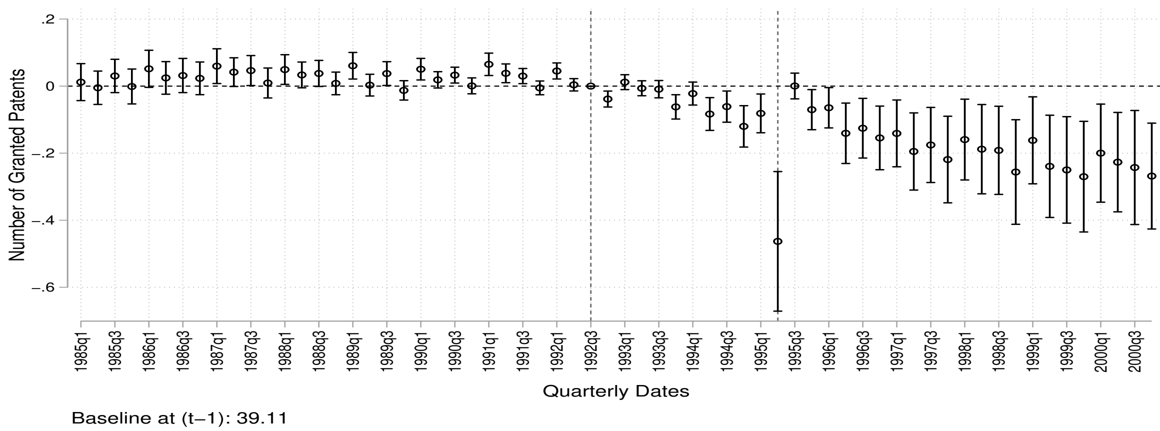
pending period, I build two instruments: i) the technical field-specific share of patents classified as second filings before the policy news - which should capture pre-existing congestion of the examination offices - and ii) the technical field-specific pending period at the European Patent Office - to capture the technical difficulty of examination.[93] The first-stage F-statistic of the excluded instruments is 33.09. Figure B.14 shows results of the second stage regression for granted patents as outcome variable. Evidence is analogous for citations-weighted patents (Figure B.15).

B.2.10 Inclusion of a flexible trend by 3-digit IPC class

Ideally, we would like to control for any field-specific and quarter-specific factor that may affect innovation independently from the policy-induced change in effective patent length. The inclusion of field $\times$ quarter fixed effects is of course not feasible, and Appendix B.1 has shown how the inclusion of a field-specific liner trend, as in Abrams (2009), can severely bias the DiD estimates if the data are not linear-in-levels before the policy. Given these restrictions, the best available alternative is the inclusion of a flexible trend by 3-digit IPC class, which is a broader definition of technical field than the one employed in the main text, in specification

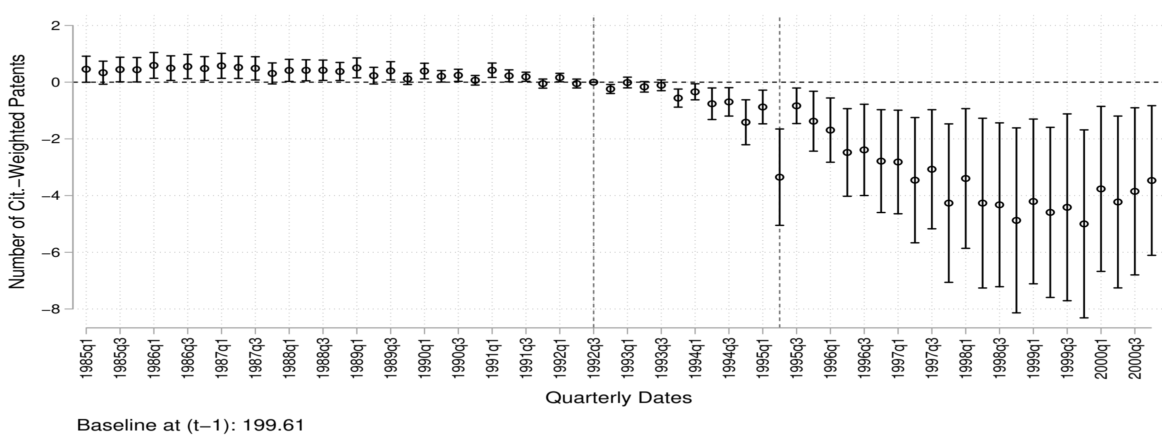
[93] A patent application is defined as a second filing if its filing date at the USPTO is subsequent to its priority date, i.e. its earliest filing date at any foreign patent office. Both instruments are computed using patents granted between the 1^{st} of January 1990 and the 31^{st} of May 1992, i.e. before the policy news in 1992Q4, as done for the treatment variable, in order to minimize potential endogeneity concerns.

Figure B.14: **Effect of 1 more day of protection on granted patents - IV**



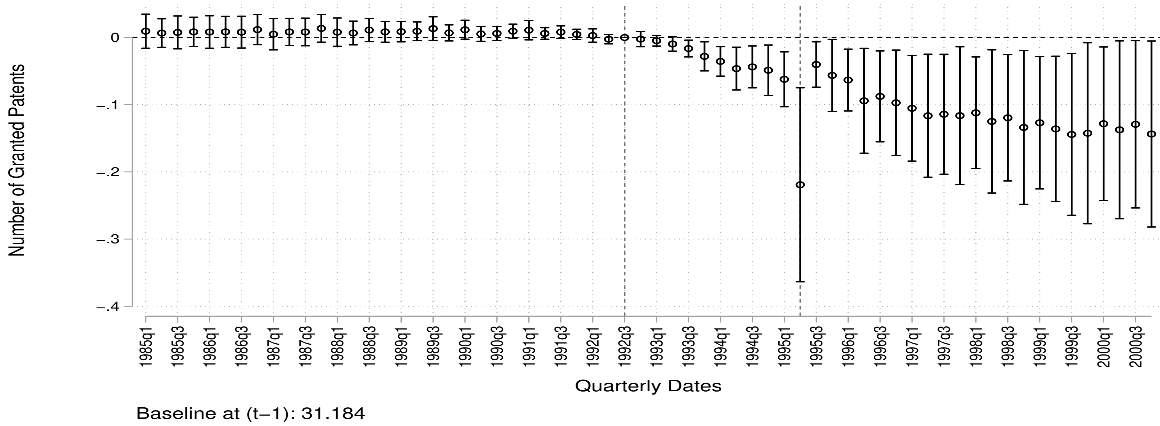
The plot shows the β_k coefficients of specification (1) where T_j is instrumented by i) congestion by foreign patents, and ii) technical difficulty of examination. See Appendix B.2.9 for details. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure B.15: **Effect of 1 more day of protection on cit.-weighted patents - IV**



The plot shows the β_k coefficients of specification (1) where T_j is instrumented by i) congestion by foreign patents, and ii) technical difficulty of examination. See Appendix B.2.9 for details. Dependent variable: 5-years citations-weighted patents filed in field j and quarter t. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure B.16: **Effect of 1 more day of protection on granted patents**



The plot shows the β_k coefficients of specification (25). Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

(1).[94] In practice, the enriched specification is

$$Y_{j,t} = \alpha_j + \sum_{k=1985Q1}^{2000Q4} \gamma_k \mathbf{1}_{(t=k)} + \sum_{k=1985Q1}^{2000Q4} \beta_k \mathbf{1}_{(t=k)} T_j + \sum_f \sum_{k=1985Q1}^{2000Q4} d_{j\in f} \mathbf{1}_{(t=k)} + \varepsilon_{j,t} \tag{25}$$

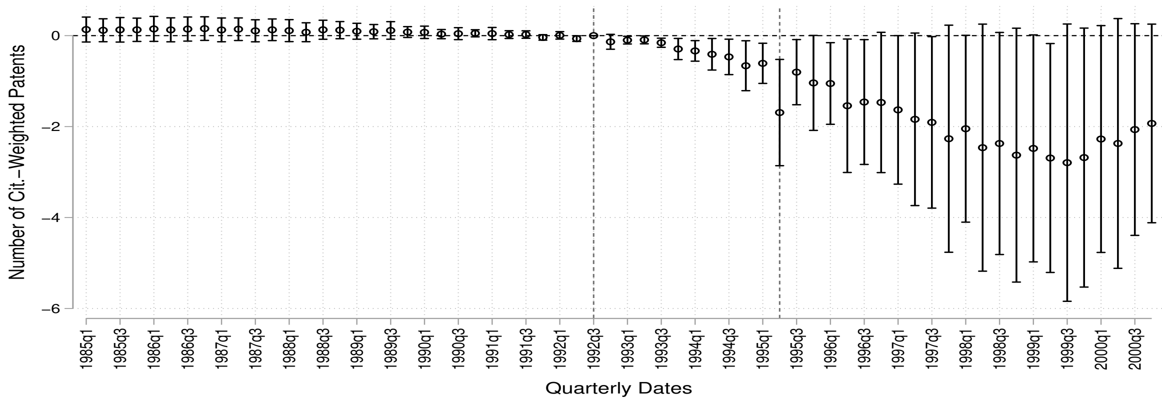
where all the terms have the same meaning as in (1). The new term $\sum_f \sum_{k=1985Q1}^{2000Q4} d_{j\in f} \mathbf{1}_{(t=k)}$ collects all the interactions between 3-digit IPC dummy variables $d_{j\in f}$ and quarterly dummies. A 3-digit IPC is indexed by f and $d_{j\in f}$ takes value one if the 4-digit field j belongs to the 3-digit field f. The coefficients of interest remain the β_k's. Figures B.16 and B.17 plot them for granted patents and citations-weighted patents as outcome variables, respectively. Results are fully consistent with Section 3.

B.2.11 Triple difference with maintenance fees

In this triple difference specification, I interact the main treatment variable T_j with a dummy variable discretizing the field-specific share of patents for which maintenance fees at 11.5 years from grant are paid. These patents are those for which the maximum patent term–which is the one affected by the policy–is binding. In such classes, the treatment effect is expected to be stronger in magnitude, as the relevance of the policy is higher. The specification of the regression is

[94] For example, the 4-digit IPC "A23D" is "Edible Oils or Fats, e.g. Margarines Shortenings, Cooking Oils". It is included in the 3-digit IPC "A23", "Food or Foodstuffs; Their Treatment, not covered by other classes" and in the 1-digit IPC "A", "Human Necessities". It further includes two 8-digit IPCs: "A23D 7/00", "Edible oil or fat compositions containing an aqueous phase, e.g. margarines", and "A23D 9/00", "Other edible oils or fats, e.g. shortenings, cooking oils".

Figure B.17: **Effect of 1 more day of protection on citations-weighted patents**



The plot shows the β_k coefficients of specification (25). Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

$$P_{j,t} = \alpha_j + d_{R_j>25\%} + \sum_{k=1985Q1}^{2000Q4} \gamma_{1,k}\mathbf{1}_{(t=k)} + \sum_{k=1985Q1}^{2000Q4} \gamma_{2,k}\mathbf{1}_{(t=k)} d_{R_j>25\%} + \chi T_j d_{R_j>25\%} + \sum_{k=1985Q1}^{2000Q4} \psi_k \mathbf{1}_{(t=k)} T_j + \sum_{k=1985Q1}^{2000Q4} \beta_k \mathbf{1}_{(t=k)} T_j d_{R_j>25\%} + \varepsilon_{j,t} \tag{26}$$

where $P_{j,t}$ is quarter-t and field-j number of granted patents, T_j is the field-specific treatment described in Section 2.3.1, and $d_{R_j>25\%}$ is a dummy taking value 1 if R_j, the field-specific percentage of patents for which the last maintenance fee at 11.5 years since the grant is paid, is above 25%. I omit the dummy for 1992Q3, which is the pre-treatment quarter. Standard errors are clustered by technical field and 95% confidence bands are plotted. Figure B.18 shows that the results go in the expected direction.

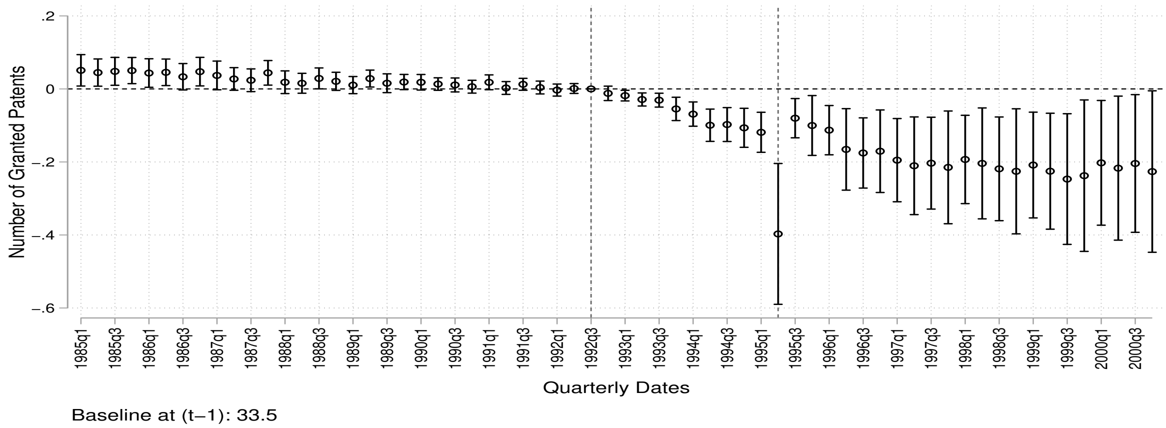
B.2.12 Placebo date

In this subsection, I present the results of a further robustness analysis at the technical field-level. I run the same specification (1) of Subsection 3.1.1, but I shift everything 10 years before the actual treatment time, where no effect is supposed to be observed. Figure B.19 shows that this is the case.

B.2.13 Alternative specifications of the model

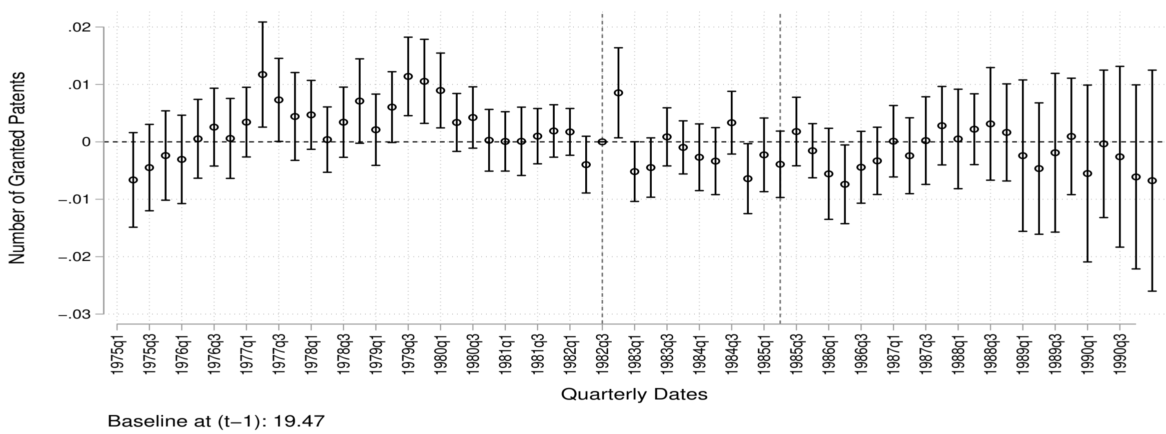
This subsection presents the results obtained from the estimation of model (1) with different transformations of the outcome variable. or the results of the estimation of a negative binomial model for count data.

Figure B.18: **Effect of 1 more day of protection on granted patents - Triple difference specification**



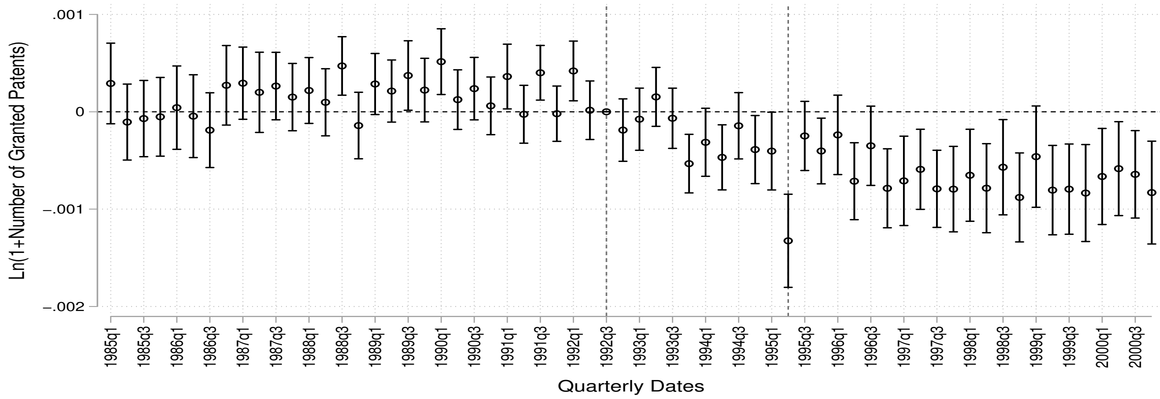
The plot shows the β_k coefficients of specification (26) having as dependent variable quarter-t and field-j number of granted patents. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure B.19: **Marginal effect of 1 more day of protection on granted patents**



The plot shows the β_k coefficients of the specification (1) having as dependent variable quarter-t and field-j number of granted patents. The sample covers 1975Q1-1990Q4. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1982Q3) and the second vertical line refers to the quarter before the policy implementation (1985Q2).

Figure B.20: **Marginal effect of 1 more day of protection on granted patents**



The plot shows the β_k coefficients of the specification (27) having as dependent variable the log. of one plus quarter-t and field-j number of granted patents. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Natural logarithms The first alternative specification of regression (1) is to take the outcome variable in natural logarithms rather than in levels. Given the presence of zeroes in the data, I run the specification

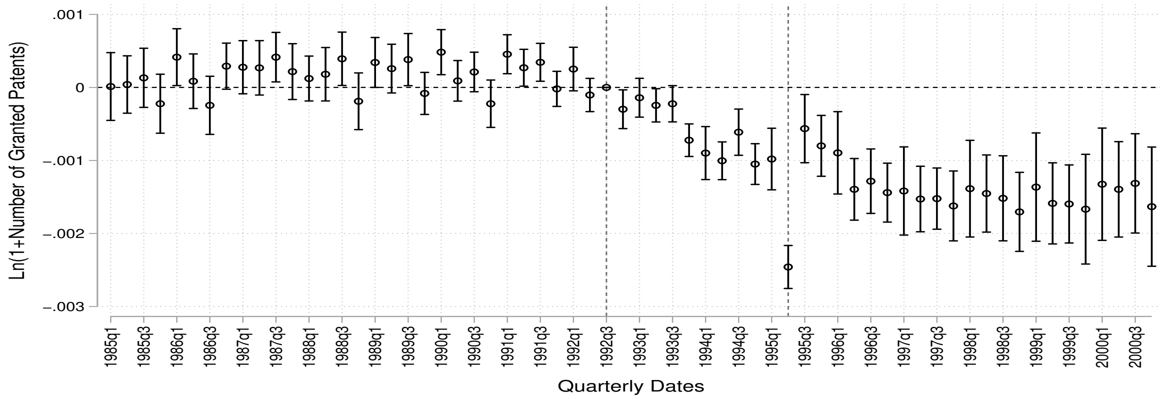
$$ln(1+P_{j,t}) = \alpha_j + \sum_{k=1985Q1}^{2000Q4} \gamma_k \mathbf{1}_{(t=k)} + \sum_{k=1985Q1}^{2000Q4} \beta_k \mathbf{1}_{(t=k)} T_j + \varepsilon_{j,t} \tag{27}$$

where all the variables have the same meaning as in subsection 3.1.The $ln(1+P_{j,t})$ transformation is chosen, among the possible alternatives, e.g. $ln(10+P_{j,t})$ or $ln(0.1+P_{j,t})$, because it is the one that implies the skewedness of the dependent variable closest to zero. Results are reported here for granted patents as outcome.[95] Figure B.20 plots the β_k coefficients of specification (27) when the regression is run considering all technical fields. However, given the presence of many zeros in the data, I also run the same regression excluding from the sample the fields with a total number of patents below the sample median. The remaining half of the fields generates more than 90% of total patents in the sample. Figure B.21 plots the β_k coefficients of specification (27) on such restricted sample, showing that results of section 3 of the paper are largely confirmed.

Negative binomial model The second alternative model to specification (1) is to use a negative binomial model for count data. The explanatory variables are the same as in the linear specification, i.e. the assumed model is

[95] In the Appendix E.1.5, there is equivalent evidence for citations-weighted patents, and unique number of inventors.

Figure B.21: **Marginal effect of 1 more day of protection on granted patents**



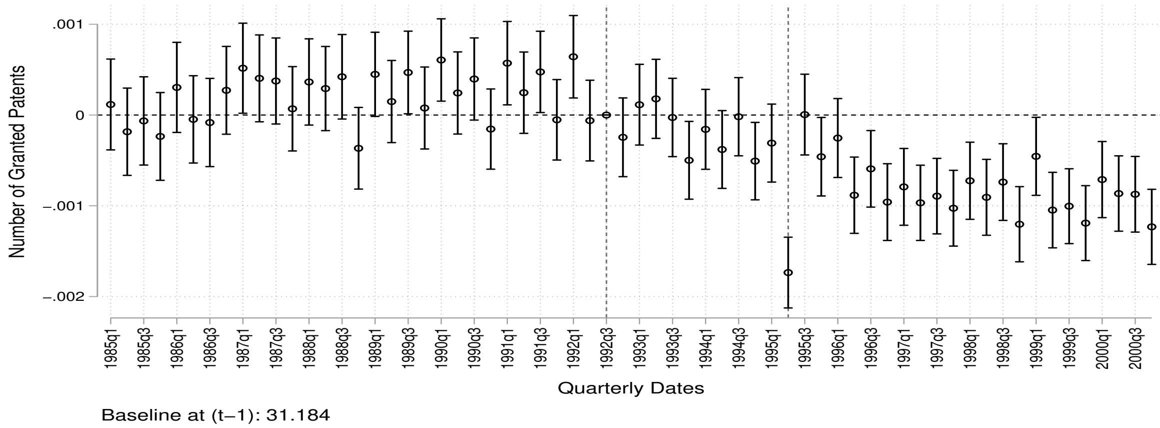
The plot shows the β_k coefficients of the specification (27) having as dependent variable the log. of one plus quarter-t and field-j number of granted patents. The sample covers only technical fields with a number of total patents above the sample median. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

$$\mathbf{E}[P_{j,t}|\mathbf{X}_{j,t}] = e^{\alpha_j + \sum_{k=1985Q1}^{2000Q4} \gamma_k \mathbf{1}_{(t=k)} + \sum_{k=1985Q1}^{2000Q4} \beta_k \mathbf{1}_{(t=k)} T_j} \tag{28}$$

Figure B.22 plots the β_k coefficients of specification (28). These are fully consistent with the evidence shown above for the linear models with logs-transformed outcome variable. When the sample is restricted to the technical fields with total patenting above the median– in order to take care of the many zeros in the full sample–the evidence (Figure B.23) is even closer to the one presented in section 3 of the paper. Appendix E.1.5 reports the results for the negative binomial model using citations-weighted patents and the unique number of inventors as outcome variables. Evidence is also in these cases fully consistent.

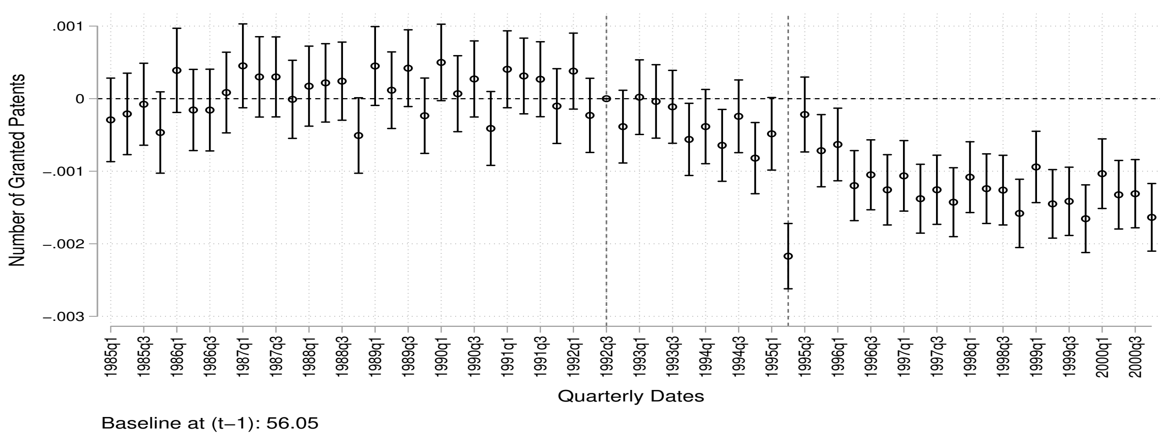
Additional transformations Finally, Appendix E.1.5 reports the results obtained *i*) using the inverse sine transformation of the outcome variable in place of the natural logarithms, and *ii*) expressing the outcome variable in percentage deviations from the number of patents filed in the specific technical field in 1985Q1. In both cases, results are analogous to the ones shown here. Appendix E.1.6 reports the results of restricting the estimation sample to technical fields with a number of quarterly patents above 25 and below 500 in every quarter. The logic of this sample restriction is to reduce the skewedness of the outcome variable, and keep the same linear-in-levels specification of section 3 of the paper. Results are robust.

Figure B.22: **Estimated treatment coefficients of a negative binomial model**



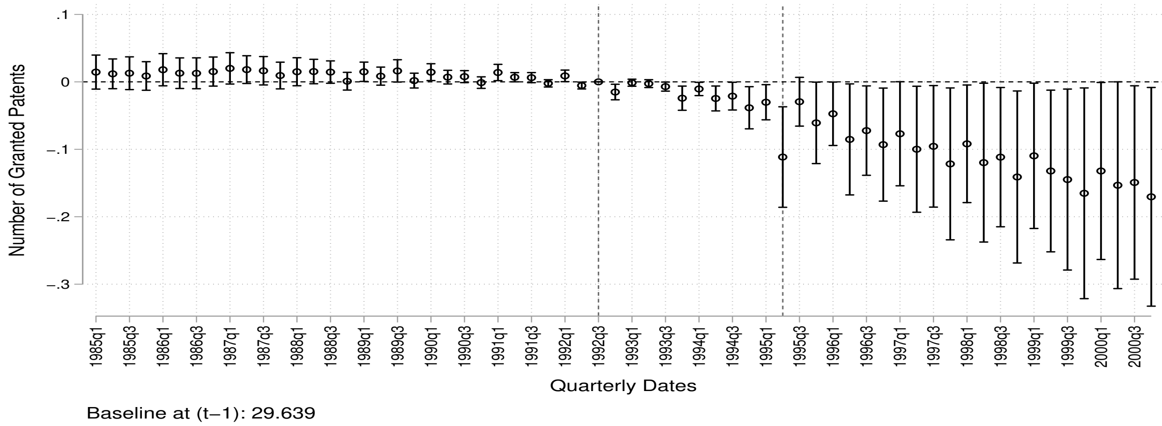
The plot shows the β_k coefficients of the specification (28) having as dependent variable quarter-t and field-j number of granted patents. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure B.23: **Estimated treatment coefficients of a negative binomial model**



The plot shows the β_k coefficients of the specification (28) having as dependent variable quarter-t and field-j number of granted patents. The sample covers only technical fields with a number of total patents above the sample median. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure B.24: **Marginal effect of 1 more day of protection on granted patents - No pharma-related fields**



The plot shows the β_k coefficients of the specification (28) having as dependent variable quarter-t and field-j number of granted patents. Pharma-related technical fields A01H, A61K, A61P, C07D, C02F, C07G, C07H, C07J, C07K, C12M, C12N, C12P, C12Q, C12S, and G01N are dropped from the sample. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

B.2.14 Dropping technical fields related to the pharmaceutical sector

Kyle and McGahan (2012) point out that some US pharmaceutical firms increased R&D investment after the TRIPs because the new regulations imposed on developing countries the patentability of several pharmaceutical products that were previously not patentable. In order to reduce the concerns that the results presented in Section 3 of the paper are biased or driven by such fields, I report here the results of specification 1 on a restricted sample that excludes all technical fields whose technologies can be related to the pharmaceutical and biotech. industries.[96] Figure B.24 shows the results for granted patents as dependent variable, and Figure B.25 the results for citations-weighted patents. The evidence of section 3 is virtually unaffected.

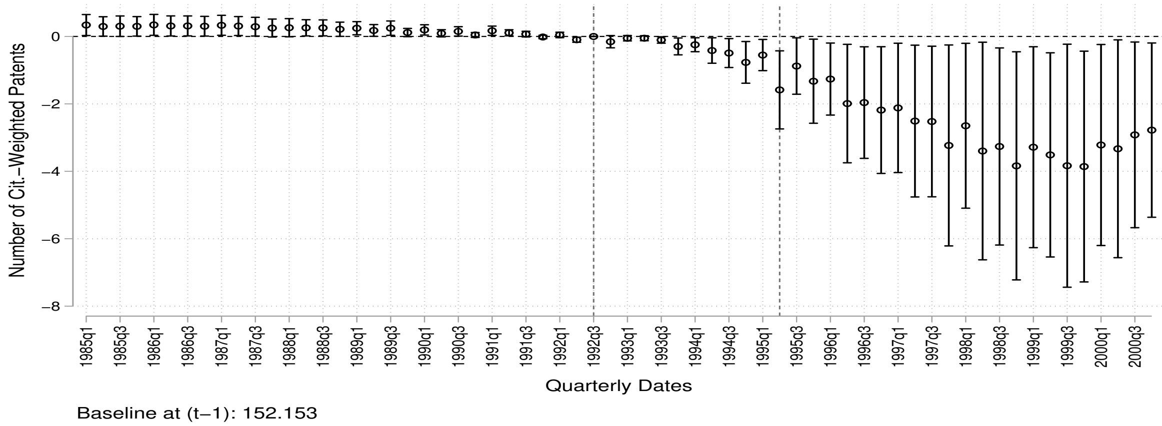
B.3 Firm-level analyses

B.3.1 Number of patents

Figure B.26 shows β_k coefficients of the firm-level difference-in-difference specification (2). The β_k capture the effect of a one-day increase of patent length on yearly firm-level patenting, in percentage deviation from the 1991 baseline average. Evidence is consistent with the behavior of patenting observed at the technical field level. On average, a 30 days future increase of patent length decreases yearly patenting by 2.6% at the firm level before implementation.

[96] These are the 4-digit IPCs: A01H, A61K, A61P, C07D, C02F, C07G, C07H, C07J, C07K, C12M, C12N, C12P, C12Q, C12S, and G01N.

Figure B.25: **Marginal effect of 1 more day of protection on cit.-weighted patents - No pharma-related fields**



The plot shows the β_k coefficients of the specification (28) having as dependent variable quarter-t and field-j number of citations-weighted granted patents. Pharma-related technical fields A01H, A61K, A61P, C07D, C02F, C07G, C07H, C07J, C07K, C12M, C12N, C12P, C12Q, C12S, and G01N are dropped from the sample. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

This estimate is close to the field-level effect. After the implementation, the impact of the same policy change implies a decrease of yearly firm-level patenting of 2.1%.

B.3.2 Citations-weighted patents

Figure B.27 shows that firm-level results are robust to the use of citation-weighted patents as a measure of innovation. The citations-weighted version of the patent count is the one provided in the NBER Patent database, and computed according to Hall, Jaffe and Trajtenberg (2001).

B.3.3 Private economic value of patents

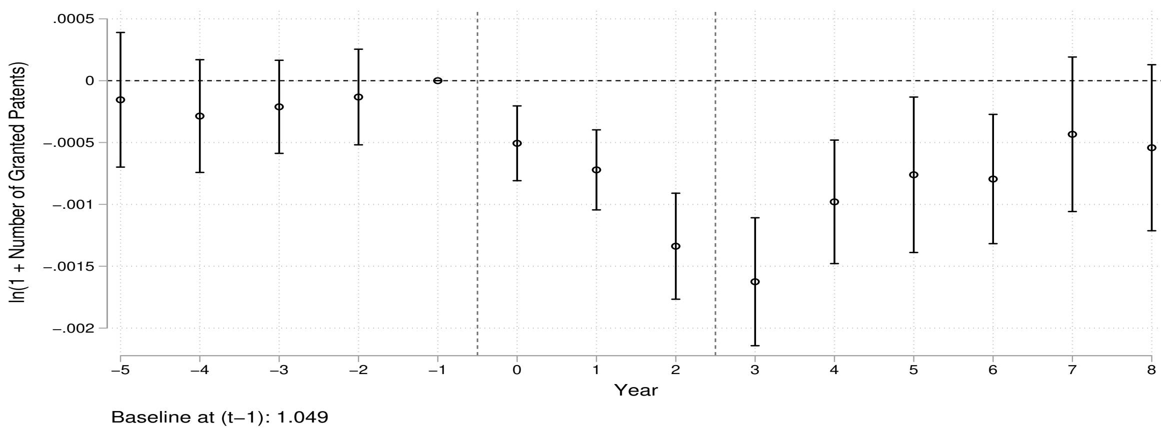
I match the dataset by Kogan et al. (2017) with the NBER patent database using USPTO patent numbers, and I aggregate patent values at the firm-level and by year. Figure B.28 plots the β_k coefficients of specification (2), run having as dependent variable one plus the natural logarithm of such patent value. The estimated effects are consistent with the evidence obtained for patents.

B.3.4 Negative binomial model for patent counts

An alternative model specification to the linear regression (2) is to use a negative binomial model for count data. The dependent variable is the number of subsequently-granted patents filed by firm i in year t. The specification of the conditional mean of the model is:

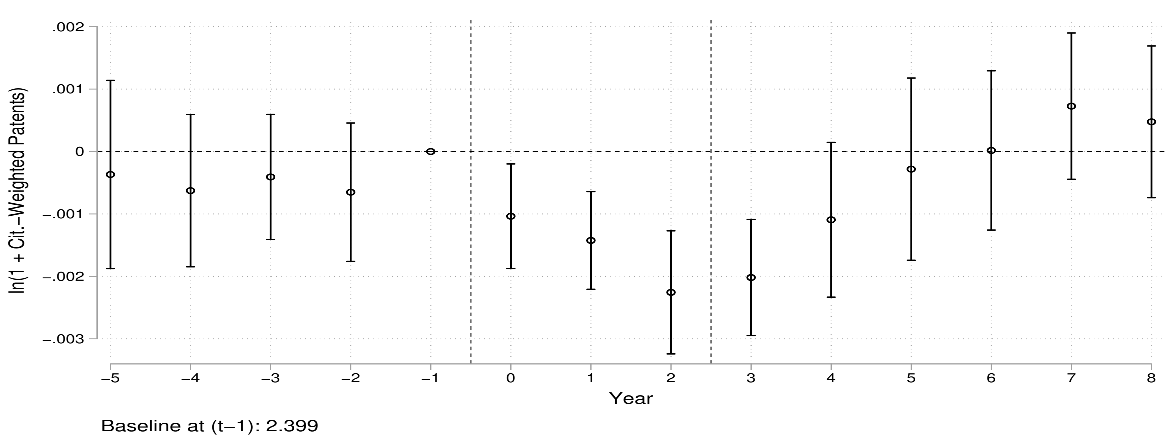
$$\mathbf{E}[P_{i,t}|\mathbf{X}_{i,t}] = e^{\alpha_i + \sum_j \eta_{1,j} sic_j + \sum_j \sum_{k=1987}^{2000} \eta_{2,j,k} sic_j \mathbf{1}_{(t=k)} + \sum_{age \in A} \delta_{age}} \times e^{\theta ln(1+S_{i,t}) + \sum_{k=1987}^{2000} \gamma_k \mathbf{1}_{(t=k)} + \sum_{k=1987}^{2000} \beta_k \mathbf{1}_{(t=k)} T_i} \tag{29}$$

Figure B.26: **Marginal effect of 1 more day of protection on firm-level patenting**



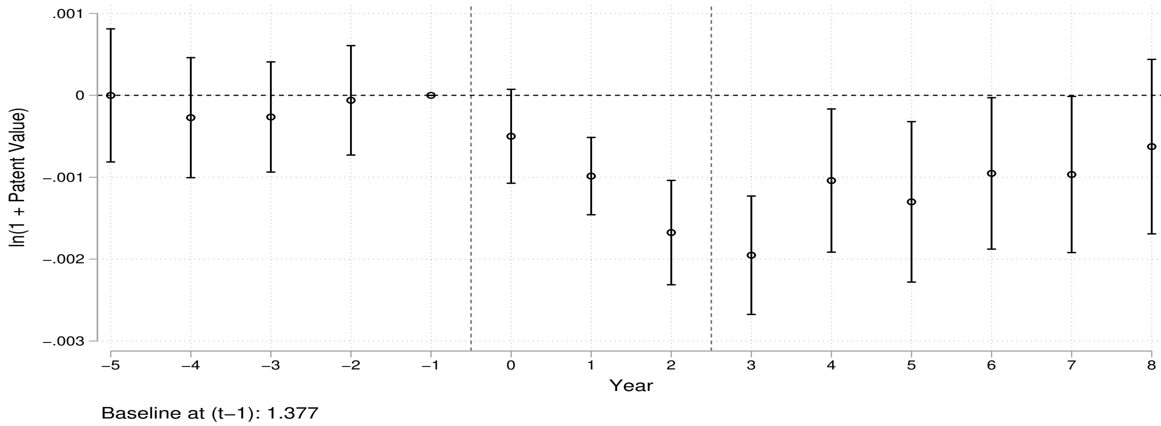
The plot shows the β_k coefficients of regression (2) having as dependent variable $ln(1+P_{i,t})$, where $P_{i,t}$ is year-t and firm-i number of granted patents. Standard errors are clustered by 2-digit SIC industry. 95% confidence bands are plotted. The first vertical line lies just before the news year (1992) and the second vertical line lies just before the implementation year (1995).

Figure B.27: **Marginal effect of 1 more day of protection on firm-level citations-weighted granted patents**



The plot shows the β_k coefficients of regression (2) having as dependent variable year-t and firm-i citations-weighted granted patents. Standard errors are clustered by 2-digit SIC industry. 95% confidence bands are plotted. The first vertical line lies just before the news year (1992) and the second vertical line lies just before the implementation year (1995).

Figure B.28: **Marginal effect of 1 more day of protection on firm-level patent value**



The plot shows the β_k coefficients of regression (2) having as dependent variable year-t and firm-i total patent value. Standard errors are clustered by 2-digit SIC industry. 95% confidence bands are plotted. The first vertical line lies just before the news year (1992) and the second vertical line lies just before the implementation year (1995).

Figure B.29 plots the β_k coefficients of specification (29). These are fully consistent with the evidence shown in subsection 3.2 of the paper.

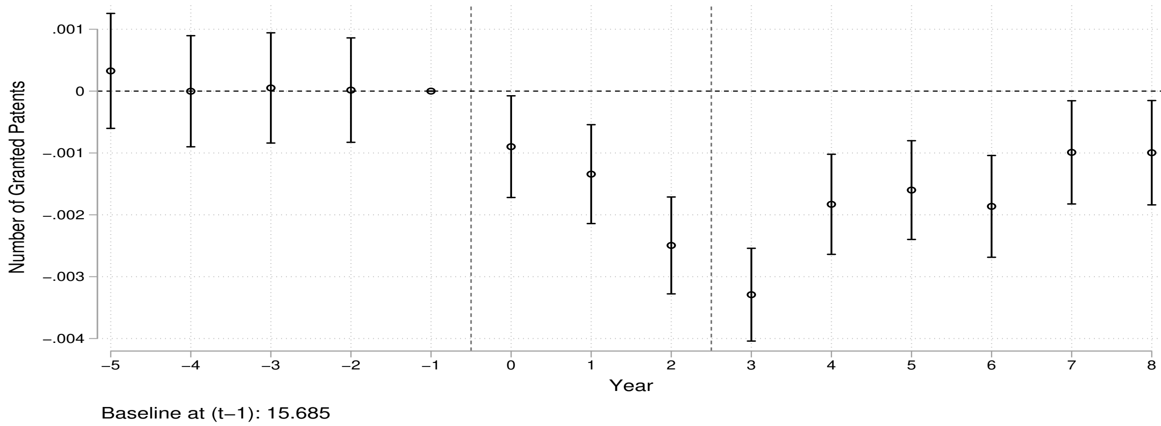
B.3.5 Within-firm analysis of innovation outcomes

The aim of this analysis is to study how a patent term change, heterogeneous by technical field, affects innovation activity *within* a specific firm, *across* technical fields. The first step is to build a panel dataset where the cross-sectional unit is a firm in a given technical field (firm $\times$ technical field). The time-dimension features 5 periods: (0) 1983-1985, (1) 1986-1988, (2) 1989-1991, (3) 1992-1995, (4) 1996-1999. Periods 0 and 1 are used to check pre-trends in the regression, period 2 is the pre-treatment period, period 3 is the period between the policy news and the policy implementation of 1995, and period 4 is the post-implementation period. The starting point for this panel dataset is the NBER Patent Database, matching patents to COMPUSTAT identifiers of applicant firms. I add up granted patents, citations-weighted patents, and patent value by firm, technical field, and time period, using the 4-digit IPC class reported in the NBER Patent Database for each patent. The specification of the regression is

$$\begin{aligned} ln(1+Y_{i,j,p}) = \alpha_i + \chi_j + \sum_{age \in A} \delta_{age} + \sum_{p=1}^{4} \gamma_k \mathbf{1}_{(p=k)} + \sum_{p=1}^{4} \eta_k \mathbf{1}_{(p=k)} s_{i,j} + \\ + \sum_{p=1}^{4} \beta_k \mathbf{1}_{(p=k)} (T_j/100) + \sum_{p=1}^{4} \xi_k \mathbf{1}_{(p=k)} s_{i,j} (T_j/100) + \varepsilon_{i,j,p} \end{aligned} \tag{30}$$

where i indexes firms, j technical fields, and p the time period. α_i are firm fixed effects, χ_j are technical field fixed effects, δ_{age} are fixed effects by median age of the firm during the period, $\mathbf{1}_{(p=k)}$ is an indicator taking value 1 when period $p = k$, T_j is technical field j's

Figure B.29: **Estimated treatment coefficients of a negative binomial model**



The plot shows the β_k coefficients of the specification (29) having as dependent variable quarter-t and firm-i number of granted patents. Standard errors are clustered by 2-digit SIC industry code. 95% confidence bands are plotted. The first vertical line lies just before the news year (1992) and the second vertical line lies just before the implementation year (1995).

treatment variable described in subsection 2.3 of the paper, $s_{i,j}$ is the share of patents by firm i produced in technical field j in the years before the policy news, i.e. up to 1991, and $\varepsilon_{i,j,p}$ is an idiosyncratic error term. In the specification, I allow for the ex-ante technological position of the firm across technical fields to influence the trends in innovation outcomes along the same dimension. The outcome variables $Y_{i,j,p}$ used in the regression are: $P_{i,j,p}$, i.e. the number of patents granted to firm i in field j and filed in period p; $C_{i,j,p}$, i.e. the citations-weighted patents granted to firm i in field j and filed in period p; $V_{i,j,p}$, i.e. the economic value of patents granted to firm i in field j and filed in period p; $f_{i,j,p}$, i.e. the share of firm i's period p's patents classified in technical field j out of the total number of firm i's, period p's patents. $f_{i,j,p}$ is not regressed in natural logs but in levels.

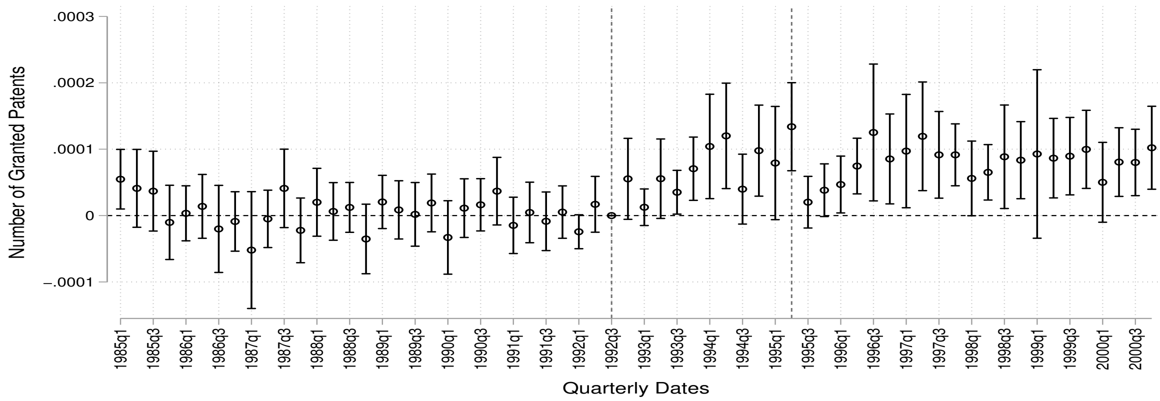
The coefficients in *Panel B* of Table B.3 confirm that the treatment is unrelated to differential trends in innovation across technical fields. The coefficients in *Panel C* capture the baseline effect of the policy and provide evidence that firms tend to reallocate innovation effort towards technical fields with a more favorable patent term. This is clearly consistent with the patterns documented in subsections 4.1 and 3.2.

Table B.3: **Within-firm cross-technical fields effect of a patent term change**

	(1) Patents	(2) Citations	(3) Value	(4) Patent Share
Panel A: Interaction coefficients				
$\mathbf{1}_{(p=0)} \times s_{i,j} \times (T_j/100)$	-0.09541*	-0.58772***	-0.24771**	0.00006
	(0.04925)	(0.17764)	(0.11098)	(0.04802)
$\mathbf{1}_{(p=1)} \times s_{i,j} \times (T_j/100)$	-0.01424	-0.20658**	0.01926	-0.03135
	(0.02765)	(0.10376)	(0.05532)	(0.03838)
$\mathbf{1}_{(p=3)} \times s_{i,j} \times (T_j/100)$	-0.07684***	-0.12578*	-0.23457***	0.00077
	(0.02088)	(0.07564)	(0.04986)	(0.03088)
$\mathbf{1}_{(p=4)} \times s_{i,j} \times (T_j/100)$	-0.01909	0.17715**	-0.11911*	-0.02483
	(0.02329)	(0.08227)	(0.06210)	(0.03253)
Panel B: Pre-trends				
$\mathbf{1}_{(p=0)} \times (T_j/100)$	-0.00006	-0.00078	0.00002	-0.00001
	(0.00015)	(0.00049)	(0.00040)	(0.00004)
$\mathbf{1}_{(p=1)} \times (T_j/100)$	-0.00002	-0.00052*	0.00007	0.00004
	(0.00009)	(0.00028)	(0.00023)	(0.00003)
Panel C: Policy-impact				
$\mathbf{1}_{(p=3)} \times (T_j/100)$	-0.00082***	-0.00197***	-0.00209***	-0.00037***
	(0.00010)	(0.00031)	(0.00030)	(0.00004)
$\mathbf{1}_{(p=4)} \times (T_j/100)$	-0.00129***	-0.00165***	-0.00325***	-0.00070***
	(0.00014)	(0.00044)	(0.00043)	(0.00004)
Observations	7,687,980	7,687,980	7,687,980	4,874,231

The Table reports the OLS estimates of specification (30). See subsection B.3.5 for details. Statistical significance levels: $^{*}(p < 0.05),^{**}(p < 0.01),^{***}(p < 0.001)$

Figure B.30: **Effect of 1 more day of protection on patents - HHI as interactor**



The plot shows the β_k coefficients of regression (31) having as dependent variable quarter-t and field-j number of granted patents. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

B.4 Evidence on the mechanism

B.4.1 Evidence on concentration as an interactor

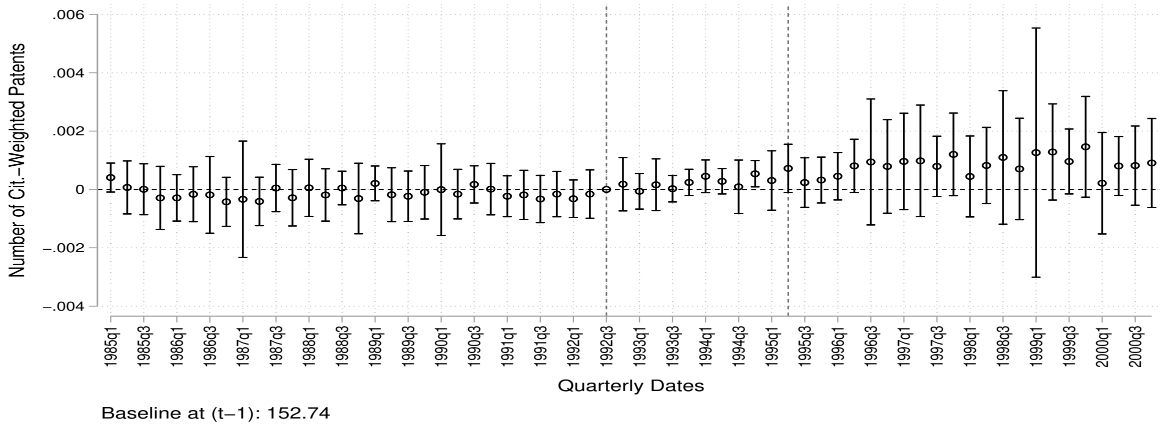
To test whether the patent term change has stronger effects on innovation in more competitve fields, I run the following triple difference specification

$$\begin{aligned} Y_{j,t} = \alpha_j + \kappa HHI_j + \sum_{k=1985Q1}^{2000Q4} \gamma_k \mathbf{1}_{(t=k)} + \sum_{k=1985Q1}^{2000Q4} \eta_k \mathbf{1}_{(t=k)} HHI_j \\ + \sum_{k=1985Q1}^{2000Q4} \theta_k \mathbf{1}_{(t=k)} T_j + \sum_{k=1985Q1}^{2000Q4} \beta_k \mathbf{1}_{(t=k)} T_j HHI_j + \varepsilon_{j,t} \end{aligned} \tag{31}$$

where $Y_{j,t}$ will be either the number of patents or citations-weighted patents, α_j are technical field fixed effects, $\mathbf{1}_{(t=k)}$ are quarterly dummy variables, HHI_j is the Herfindahl-Hirschman Index of concentration based on the share of patents granted to different applicants in a given field before the policy news, and T_j is the policy-induced, field-specific change in effective protection time. Since the HHI is smaller in less concentrated technical fields, for the treatment to affect more strongly innovation in more competitive fields it should be the case that the β_k's of previous regression are positive.

Figure B.30 plots the estimated β_k coefficients of the previous specification for $Y_{j,t}$ being the number of granted patents in field j and quarter t. This confirms that in technical fields where the degree of concentration is lower, innovators respond more strongly to patent protection time. Evidence for citations-weighted patents is identical (Figure B.31).

Figure B.31: **Marginal effect of 1 more day of protection on citations-weighted patents - HHI as interactor**



The plot shows the β_k coefficients of regression (31) having as dependent variable quarter-t and field-j 5-years citations-weighted patents. See Appendix B.4.1 for all the details about the empirical strategy and the specification. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

B.4.2 Evidence on entry as an interactor

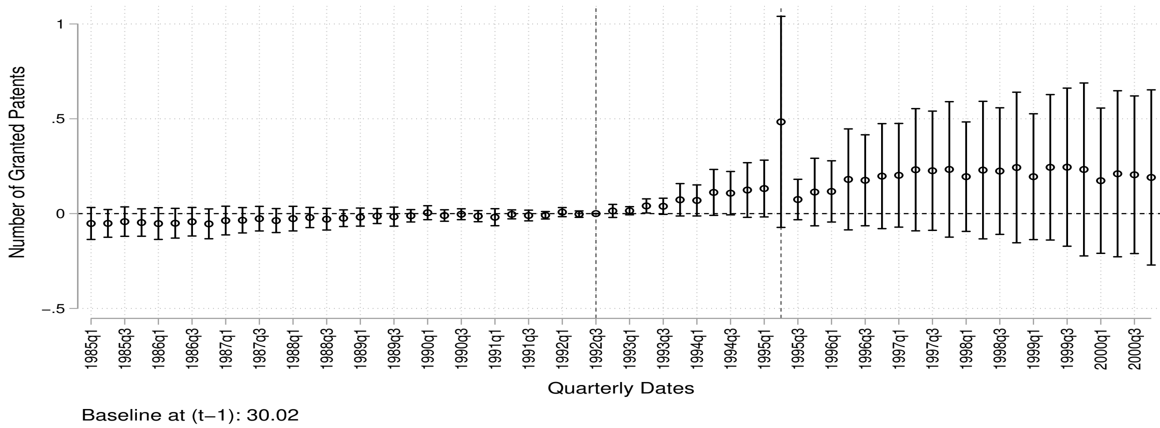
As a further test for the fact that the change in patent length affects innovation more strongly in more competitive fields, I run the triple difference specification (B.31) with s_j^E–the percentage of patents that are granted to applicants that have never patented in a given field, computed before the policy news–in place of HHI_j as interactor. Since the s_j^E is higher in technical fields where entry intensity is higher, for the working hypothesis to be verified it should be the case that the β_k's of previous regression are negative.

Figure B.32 plots the estimated β_k coefficients. Using entry rates as a measure of competition, it is not possible to conclude that the treatment affects innovation more strongly in more competitive fields. Results are the same using citations-weighted patents as outcome (see Figure B.33).

B.4.3 Evidence on concentration as outcome

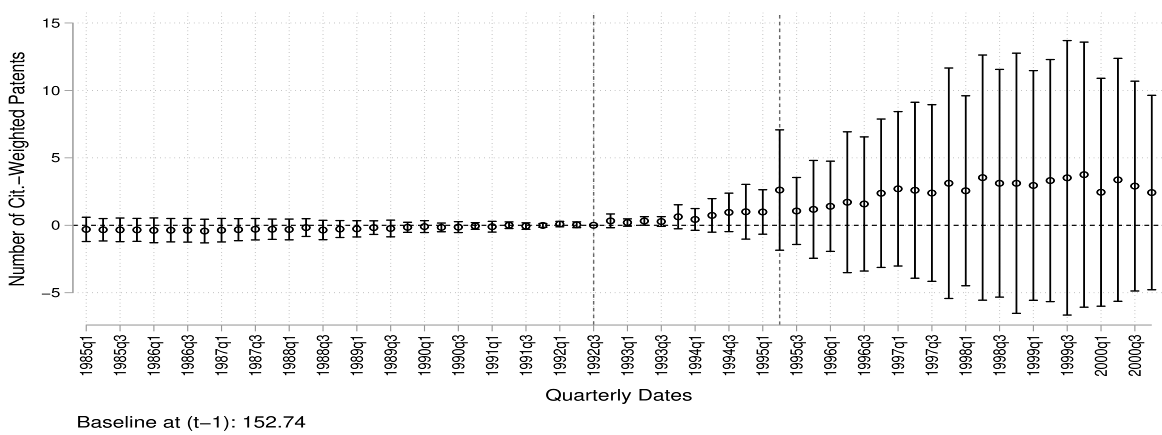
To test the effect of a change in patent length on competition, I run specification (1) using as dependent variable $HHI_{j,t}$, i.e. the Herfindahl–Hirschman Index based on the flow of patents filed by different applicants in quarter t and field j. If competition falls as a consequence of a longer patent length, the post-implementation estimated β_k's of (1) should be positive. A lower T_j should increase competition and lower the HHI. Figure B.34 plots the estimated β_k's, and shows that concentration is unaffected by the policy.

Figure B.32: **Effect of 1 more day of protection on patents - Entry rate as interactor**



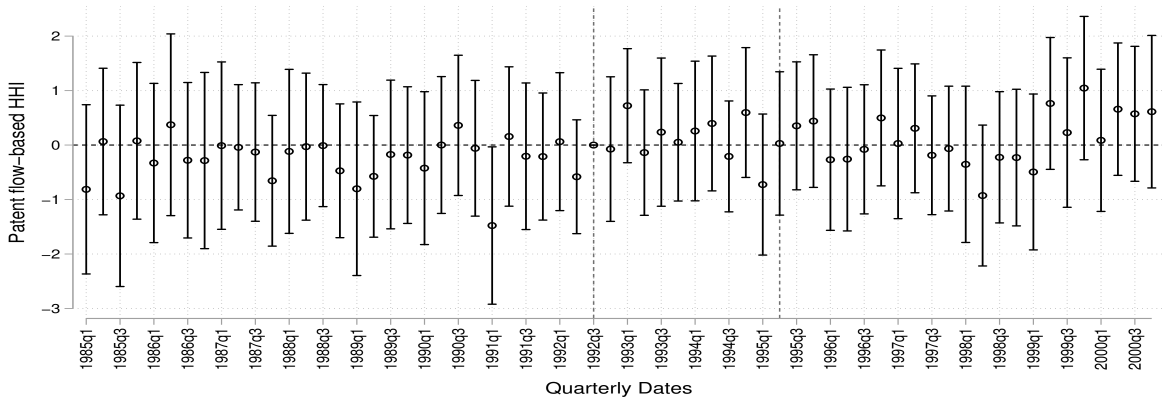
The plot shows the β_k coefficients of regression (31) having as dependent variable quarter-t and field-j number of granted patents and s_j^E as interactor of T_j. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure B.33: **Marginal effect of 1 more day of protection on citations-weighted patents - Entry intensity as interactor**



The plot shows the β_k coefficients of regression (31) having as dependent variable quarter-t and field-j 5-years citations-weighted patents and s_j^E as interactor of T_j. See Appendix B.4.2 for all the details about the empirical strategy and the specification. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure B.34: **Marginal effect of 1 more day of protection on the HH Index of concentration**



The plot shows the β_k coefficients of regression (1) having as dependent variable quarter-t and field-j Herfindahl–Hirschman index. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

B.4.4 Evidence on entry as outcome

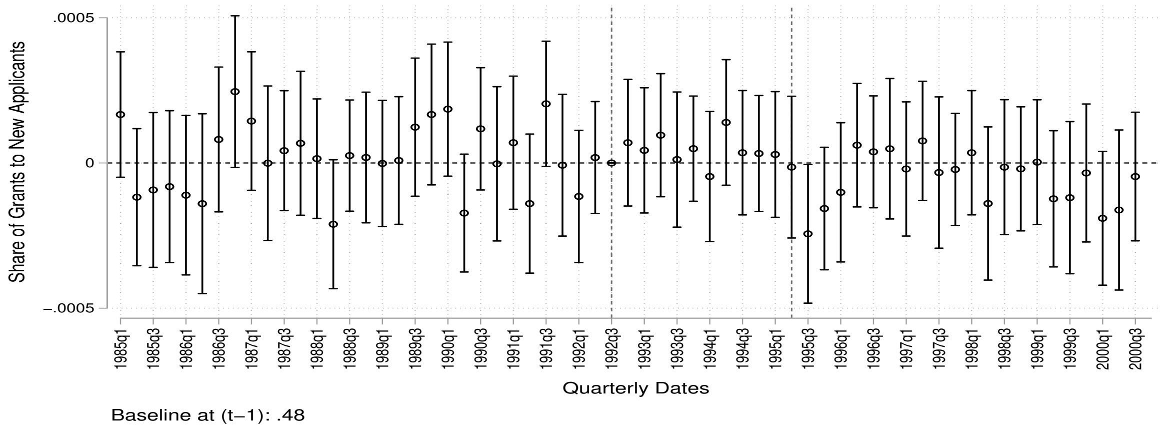
As an alternative test of the impact of patent length on competition, I run specification (1) using entry rate as outcome. To build the entry intensity measure, I determine which applicants in quarter-t and field-j are *new* to the field (entrants).[97] Entry intensity is defined as the *share* of granted patents filed by new applicants. I run specification (1) using entry as outcome variable. Figure B.35 shows that the latter does not respond to the policy. Evidence on the absolute number of new applicants and on the absolute number of patents granted to new applicants can be found in Figures B.36 and B.37.

B.4.5 Evidence on the average quality of incumbents' patents

An extension of patent length may increase the attitude of incumbents to use patent rights to foreclose the entry of innovators. However, this should be reflected in a fall of the average quality of patents granted to incumbents, both in absolute terms and relative to the quality of patents granted to new entrants. Figure B.38 shows the results of specification (1) having as dependent variable the average quality of patents granted to incumbent firms. In absolute terms, the average quality of incumbents' patents slightly falls in response to a patent term increase. However, the effect is quantitatively weak and barely different from zero statistically. Figure B.39 shows the results of the same regression using as dependent variable the average quality of patents granted to incumbent firms **divided** by the average quality of patents granted to new applicants. In relative terms, the average quality of incumbents'

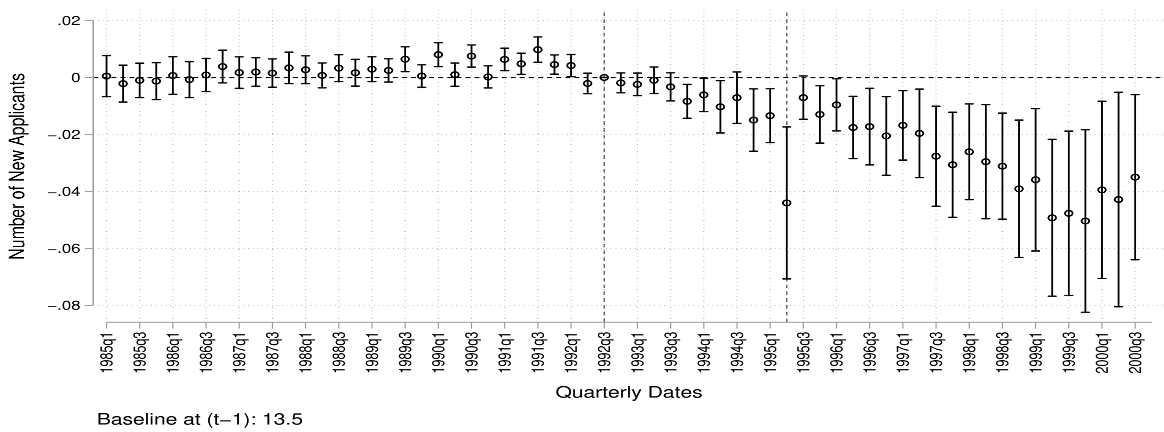
[97] New applicants at the quarterly level are determined using STAN harmonized applicant's identifiers from the EPO Worldwide Bibliographic Database available in PATSTAT and selecting, among the applicants observed in a given field-quarter, those that are never observed before.

Figure B.35: **Marginal effect of 1 more day of protection on the share of total grants to new applicants**



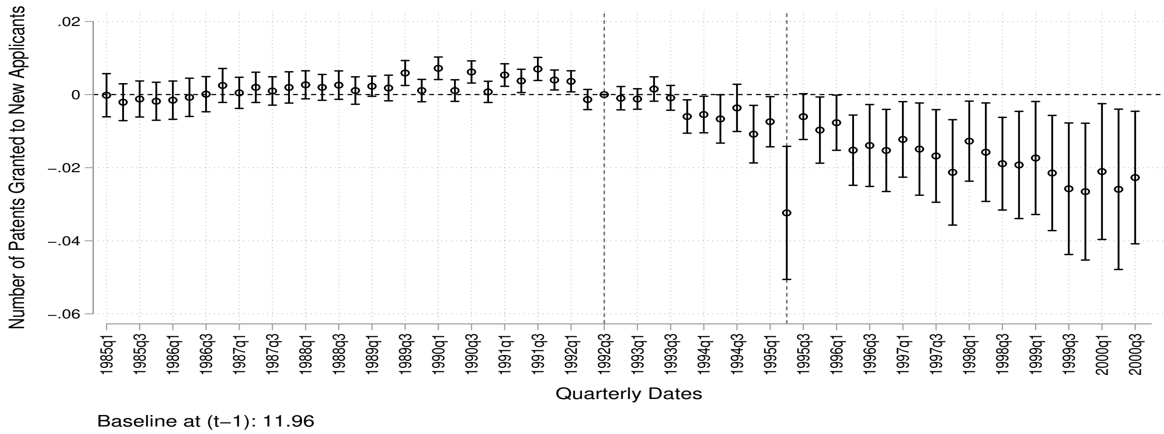
The plot shows the β_k coefficients of regression (1) having as dependent variable quarter-t and field-j *share* of granted patents filed by new applicants (entrants). Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure B.36: **Marginal effect of 1 more day of protection on the number of new applicants (entrants)**



The plot shows the β_k coefficients of regression (1) having as dependent variable quarter-t and field-j number of new applicants (entrants). Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure B.37: **Marginal effect of 1 more day of protection on the number of grants by new applicants**



The plot shows the β_k coefficients of regression (1) having as dependent variable quarter-t and field-j number of granted patents filed by new applicants (entrants). Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

patents shows no reaction to the policy change. Overall, I interpret this evidence as speaking *against* the anti-competitive use of longer patents by incumbent innovators.

B.4.6 Within-field backward citation intensity as interactor

To test **H1**, I run the triple-difference specification (31) having as interactor of the treatment T_j the within-field j share of patents S_j that have at least one applicant-made backward citation to at least one previous patent classified in the same technical field j. S_j replaces HHI_j in (31) and is computed before the policy news. Figure B.40 reports the estimated β_k coefficients, confirming the validity of **H1** and the evidence of subsection 4.2 of the paper.

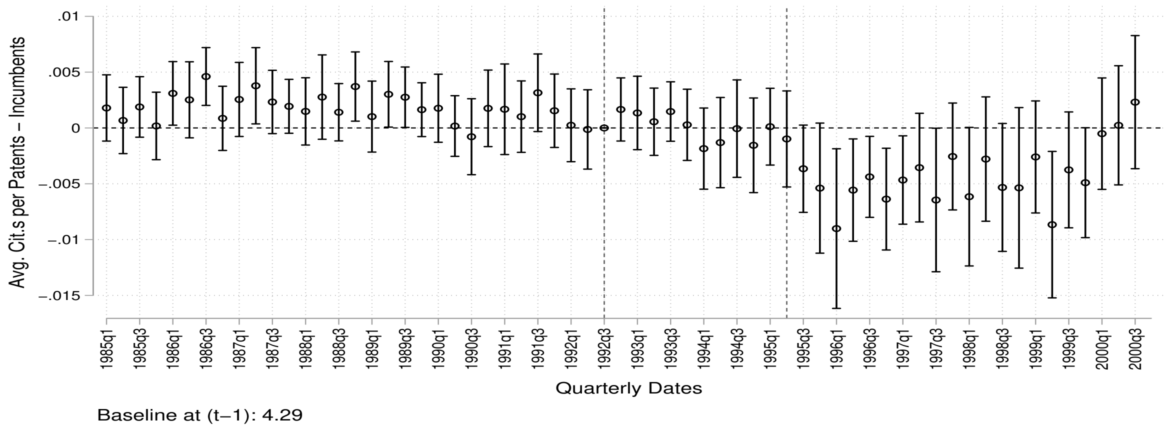
B.4.7 Within-field backward citations as an outcome

This subsection shows evidence in support of **H2** by running specification (1) on two outcomes. First, the number of patents, by technical field and quarter of application, that have at least one applicant-made backward citation to prior patents classified in the same technical field (Figure B.41). Second, the number of applicant-made backward citations from patents classified in field j and filed in quarter t to other patents classified in the same field and filed beforehand (Figure B.42). Figures B.41 and B.42 confirm the patterns documented in subsection 4.2 of the paper.

B.4.8 Direct within-field links among pre- and post-implementation innovations

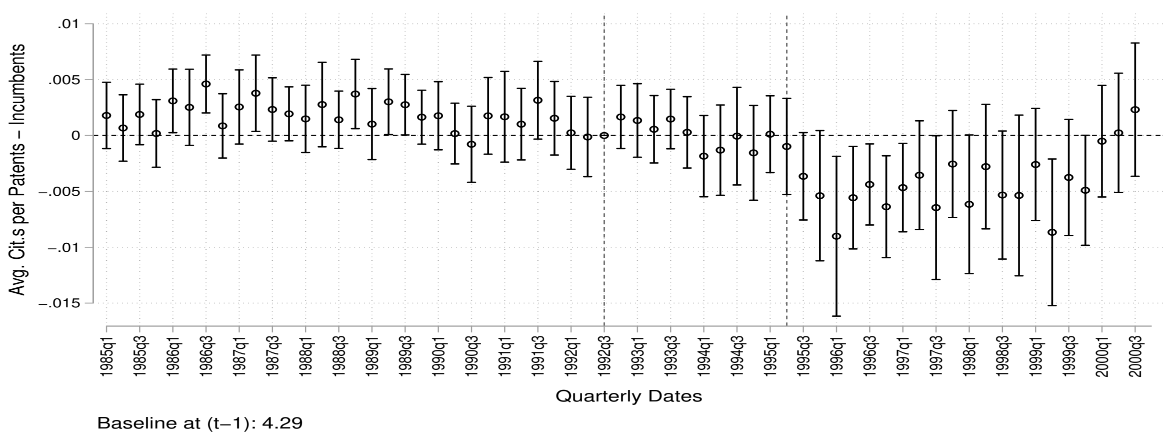
In this subsection, I try to provide evidence of the presence of a direct link between innovations produced during the pre- and post-implementation periods in the same technology field. My strategy is to count the number of applicant's backward citations from patents filed

Figure B.38: **Marginal effect of 1 more day of protection on the average number of citations per patent by incumbents**



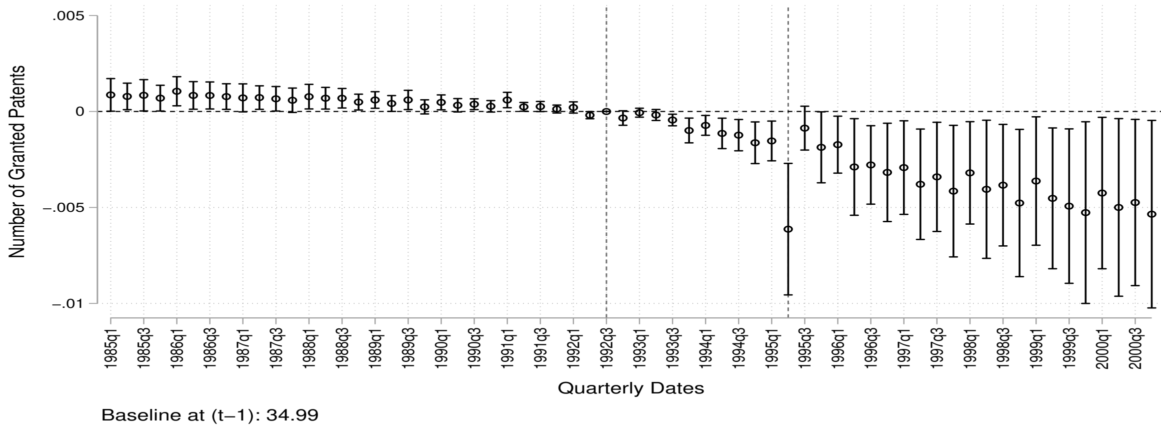
The plot shows the β_k coefficients of regression (1) having as dependent variable the average number of forward citations received by quarter-t and field-j patents granted to incumbent innovators. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure B.39: **Marginal effect of 1 more day of protection on the average number of citations per patent by incumbents relative to those of entrants**



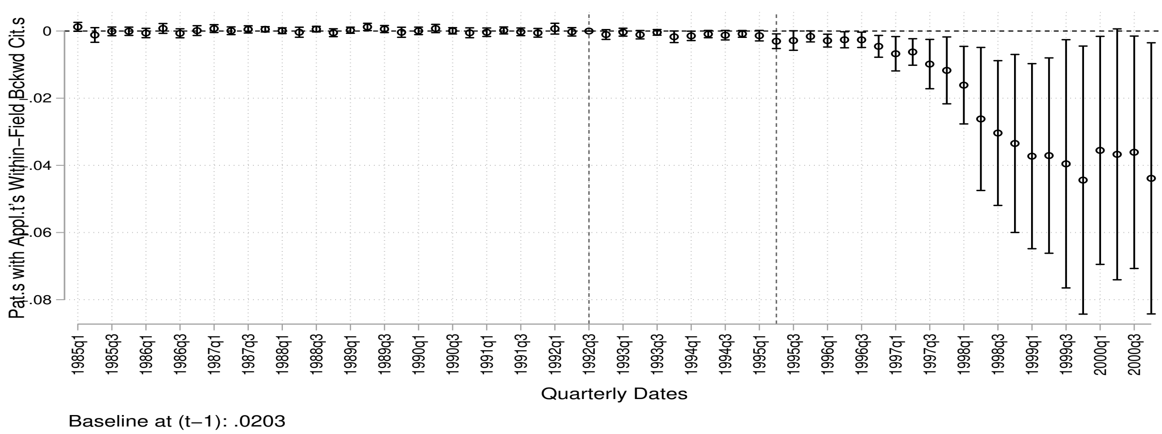
The plot shows the β_k coefficients of regression (1) having as dependent variable the average number of forward citations received by quarter-t and field-j patents granted to incumbent innovators divided by the average number of forward citations received by quarter-t and field-j patents granted to new entrants. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure B.40: **Heterogeneity analysis for the same-field citation intensity - Granted patents**



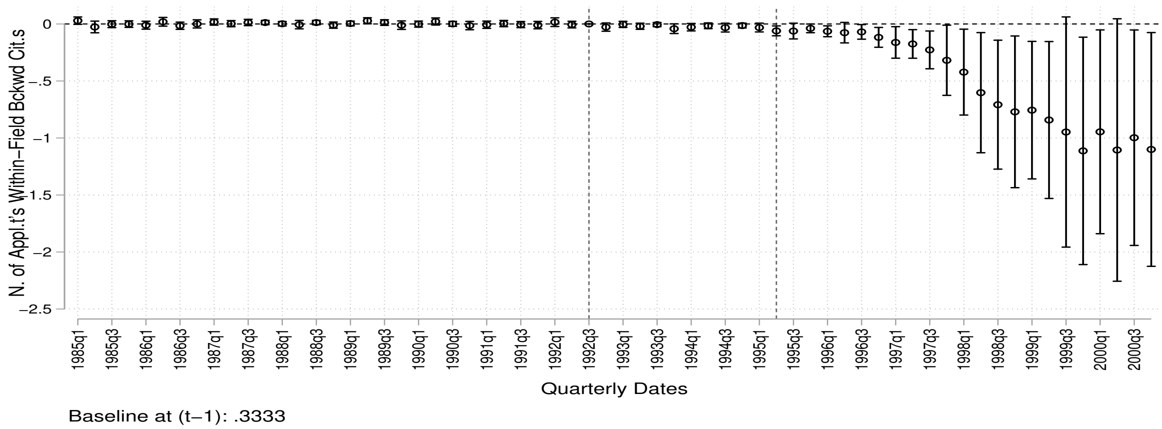
The plot shows the β_k coefficients of regression $P_{j,t} = \psi_r + \alpha_j + \kappa S_j + \sum_{k=1985Q1}^{2000Q4} \gamma_k \mathbf{1}_{(t=k)} + \sum_{k=1985Q1}^{2000Q4} \eta_k \mathbf{1}_{(t=k)} S_j + \sum_{k=1985Q1}^{2000Q4} \theta_k \mathbf{1}_{(t=k)} T_j + \sum_{k=1985Q1}^{2000Q4} \beta_k \mathbf{1}_{(t=k)} T_j S_j + \varepsilon_{j,t}$. $P_{j,t}$ is quarter-t, and field-j number of granted patents, T_j is the field-specific treatment described in Section 2.3.1, and S_j is the pre-announcement share of patents of field j that have at least one applicant-made backward citation to patents in field j. I omit the dummy for 1992Q3, which is the pre-treatment quarter. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure B.41: **Marginal effect of 1 more day of protection on the number of patents with at least one applicant-made backward citation to the same field**



The plot shows the β_k coefficients of specification (1) having as dependent variable the number of patents in field j and quarter t that have at least one applicant-made backward citations to prior patents classified in the same technical field. Clustered 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure B.42: **Marginal effect of 1 more day of protection on applicant-made backward citations to the same field**



The plot shows the β_k coefficients of specification (1) having as dependent variable the number applicant-made backward citations made by patents in field j and quarter t to prior patents classified in the same technical field. Clustered 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

in 1995Q3-1999Q4 to patents filed in 1992Q4-1995Q2 and classified in the same technological field as the citing patent. Then, I relate the change in protection by technical field to three outcomes: *i*) the share of patents filed in the post-implementation phase that have at least one applicant-made, within-field backward citation to a patent filed in the pre-implementation period; *ii*) the total number of such citations; and *iii*) the total number of such backward citing patents. As a control group, I use the same technical fields 10 years before, i.e. 1985Q3-1989Q4 and 1982Q4-1985Q2. The difference-in-difference regression equation is

$$Y_{j,p} = d_p + \delta T_j + \beta d_p T_j + \varepsilon_{j,p} \tag{32}$$

where d_p is a dummy variable taking value 1 if the outcome refers to the 1992Q4-1999Q4 period, and value 0 for 1982Q4-1989Q4. T_j is the usual policy-driven change in patent length for technical field j, and β is the difference-in-difference coefficient of interest. Columns (1), (3), and (5) of Table B.4 report the estimated coefficients of the previous regression for the share of patents, the number of backward citations, and the number of patents satisfying the criteria outlined in the previous paragraph. Standard errors are clustered by technological field. All outcome variables confirm **H2**. Columns (2), (4), and (6) augment the previous specification with fixed effects by technical field. Indeed, the negative DiD estimates imply that fields with a positive treatment, i.e. with a fall in innovation during the pre-implementation phase, show a lower backward citations intensity to patents produced exactly in the pre-implementation period and in the same technological field.

Table B.4: **Direct evidence on within-field intertemporal technology link**

	(1) Pat. Share	(2) Pat. Share	(3) Cit.s	(4) Cit.s	(5) Pat.s	(6) Pat.s
d_{post}	0.22399***	0.21880***	1080.02876***	1108.23022*	138.34799***	141.92759*
	(0.01945)	(0.02256)	(396.26864)	(580.73722)	(51.78394)	(75.87705)
T_j	-0.00004		-0.02388		-0.00722	
	(0.00003)		(0.01588)		(0.00486)	
$d_{post} \times T_j$	-0.02296***	-0.02194***	-170.02418**	-174.95083	-20.99621**	-21.60771
	(0.00375)	(0.00444)	(74.18362)	(108.83809)	(9.74579)	(14.29485)
Tech. field F.E.		Y		Y		Y
Obs.	1211	1211	1211	1211	1211	1211

The Table reports the OLS estimates of specification (32). See B.4.8 for all the details. Standard errors are clustered by technical field. Statistical significance levels: $^{*}(p < 0.10)$,** $(p < 0.05)$,*** $(p < 0.01)$

B.4.9 Does the technology link act within-firm or between-firms?

In this subsection, I detail the steps of the decomposition of subsection 4.2.2 of the paper. First, I define the theoretical objects of the decomposition.

$$\Delta \hat{P}^{A}_{j,p} \equiv \mathbf{E}[P_{j,p}|T = T_j] - \mathbf{E}[P_{j,p}|T = 0] \tag{33}$$

is the difference between expected patents–filed in period p and field j–conditional on the policy-induced change in protection time T_j, and expected patents absent any treatment. This represents the aggregate impact of the policy change on innovation in period p. $\Delta \hat{P}^{A}_{j,p}$ can be decomposed in two parts. The first is policy-driven innovation by incumbent firms ($\Delta \hat{P}^{A,I}_{j,p}$), and the second is the contribution of entrant firms ($\Delta \hat{P}^{A,E}_{j,p}$), i.e.

$$\Delta \hat{P}^{A}_{j,p} = \Delta \hat{P}^{A,I}_{j,p} + \Delta \hat{P}^{A,E}_{j,p} \tag{34}$$

Further, I assume that $\Delta \hat{P}^{A,I}_{j,p}$ can be broken down into: 1) $\Delta \hat{P}^{A,I,T}_{j,p}$, i.e. the direct impact of the patent term change on innovation, 2) $\Delta \hat{P}^{A,I,B}_{j,p}$, i.e. the between-firms component generated by the aggregate policy-driven innovation from the previous period ($\Delta \hat{P}^{A}_{j,p-1}$), and 3) $\Delta \hat{P}^{A,I,W}_{j,p}$, the within-firm component driven by within-firm technological linkages between past and present innovation. So,

$$\Delta \hat{P}^{A,I}_{j,p} = \Delta \hat{P}^{A,I,T}_{j,p} + \Delta \hat{P}^{A,I,B}_{j,p} + \Delta \hat{P}^{A,I,W}_{j,p} \tag{35}$$

I assume that for entrant firms the within-firm component of their contribution to the aggregate effect is 0. This leads to the following decomposition

$$\Delta \hat{P}^{A,E}_{j,p} = \Delta \hat{P}^{A,E,T}_{j,p} + \Delta \hat{P}^{A,E,B}_{j,p} \tag{36}$$

Finally, define the relative contribution of incumbents to the total effect of the policy in period p as

$$s^I_p = \frac{\Delta \hat{P}^{A,I}_{j,p}}{\Delta \hat{P}^{A}_{j,p}} \tag{37}$$

The second step concerns the estimation of these objects. The main dataset for this is the firm $\times$ technical field panel dataset described in Appendix B.3.5. The cross-sectional unit is a firm in a given technical field (firm $\times$ technical field) and there are 5 time-periods: (0) 1983-1985, (1) 1986-1988, (2) 1989-1991, (3) 1992-1995, (4) 1996-1999. To estimate $\Delta \hat{P}^{A}_{j,3}$, i.e. the aggregate policy impact in period (3), I consolidate the data by technical field and period, and I run the specification

$$P_{j,p} = \alpha_j + \sum_{k=0}^{3} \gamma_k \mathbf{1}_{(p=k)} + \sum_{k=0}^{3} \beta_k \mathbf{1}_{(p=k)} T_j + \varepsilon_{j,p} \tag{38}$$

All the variables have the same meaning as in specification (1) of Section 3, and $P_{j,p}$ is the number of granted patents filed in period p and classified in technical field j. Notice that the post-implementation period is excluded, and the regression is run including *all* firms in the sample regardless of when they enter, i.e. start innovating. $\Delta \hat{P}^{A}_{j,3}$ is estimated according to definition (33) using the linear specification (38).

Then, I estimate $\hat{P}^{A,I}_{j,3}$, i.e. the contribution of incumbent firms to the total effect in period (3), by aggregating the data by technical field as above, but *excluding* patents by firms that start innovating in period (3) itself, i.e. the entrants in period (3). I run specification (38) on such sample, and I estimate $\Delta \hat{P}^{A,I}_{j,3}$ using expression (33), given the new parameter estimates. Finally, the contribution of entrants to the aggregate policy effect in period (3) can be determined residually using (34).

Under the assumption that the between-firms policy-driven spillover is at work in the post-implementation period only, for period (3) expressions (35) and (36) can be rewritten as $\Delta \hat{P}^{A,I}_{j,3} = \Delta \hat{P}^{A,I,T}_{j,3}$ and $\Delta \hat{P}^{A,E}_{j,3} = \Delta \hat{P}^{A,E,T}_{j,3}$. Once we have the estimated $\Delta \hat{P}^{A,I}_{j,3}$ and $\Delta \hat{P}^{A,E}_{j,3}$ from the previous steps, these coincide with the direct policy effect.

To estimate $\Delta \hat{P}^{A}_{j,4}$, i.e. the aggregate policy impact in period (4), I consolidate the data by technical field and period and I run the specification

$$P_{j,p} = \alpha_j + \sum_{k=0}^{4} \gamma_k \mathbf{1}_{(p=k)} + \sum_{k=0}^{4} \beta_k \mathbf{1}_{(p=k)} T_j + \varepsilon_{j,p} \tag{39}$$

All the variables have the same meaning as in specification (1) of Section 3, but now specification (39) includes also period (4), differently from specification (38). Again, *all* firms are included, regardless of when they start innovating. $\Delta\hat{P}^{A}_{j,4}$ is estimated according to definition (33) using the linear specification (39).

Then, I estimate $\hat{P}^{A,I}_{j,4}$, i.e. the contribution of incumbent firms to the total effect in period (4), by aggregating the data by technical field as above, but *excluding* patents by firms that start innovating in period (4) itself, i.e. the entrants in period (4). I run specification (39) on such sample, and I estimate $\Delta\hat{P}^{A,I}_{j,4}$ using expression (33), given the new parameter estimates. Finally, the contribution of entrants to the aggregate policy effect in period (4) can be determined residually using (34). The relative contribution of incumbents to the total post-implementation policy effect is $\hat{s}^I_4 = \frac{\Delta\hat{P}^{A,I}_{j,4}}{\Delta\hat{P}^{A}_{j,4}}$.

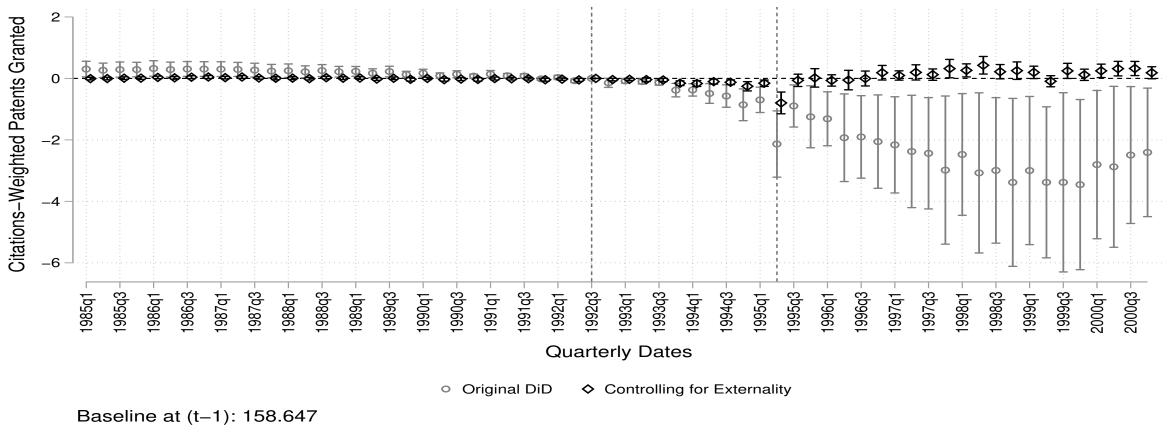
To isolate the direct effect of the policy on innovation in the post-implementation period for the case of incumbents, I follow the strategy described in subsection 4.4 and I augment the baseline difference-in-difference specification by lagged patenting in the field in the previous period, interacted with period-specific dummy variables. These terms capture the impact of the lagged spillover on innovation. So, the specification is

$$P_{j,p} = \alpha_j + \sum_{k=0}^{4} \gamma_k \mathbf{1}_{(p=k)} + \sum_{k=0}^{4} \beta_k \mathbf{1}_{(p=k)} T_j + \sum_{k=0}^{4} \eta_k \mathbf{1}_{(p=k)} P_{j,p-1} + \chi P_{j,p-1} + \varepsilon_{j,p} \tag{40}$$

The regression is run excluding from the sample patents from new entrants. Using definition (33) once more, I isolate the direct effect of the policy on incumbents, net of the intertemporal effects generated by the within and between components, i.e. $\Delta\hat{P}^{A,I,T}_{j,4}$. I assume that the direct effect of the policy affects entrants proportionally to their contribution to the total post-implementation effect. Therefore, I compute $\Delta\hat{P}^{A,E,T}_{j,4} = \frac{1-\hat{s}^I_4}{\hat{s}^I_4}\Delta\hat{P}^{A,I,T}_{j,4}$.

Equation (36) can be used to infer $\Delta\hat{P}^{A,E,B}_{j,4} = \Delta\hat{P}^{A,E}_{j,4} - \Delta\hat{P}^{A,E,T}_{j,4}$, i.e. the contribution of the between-firms technological spillover to aggregate entrants' post-implementation innovation. Assuming that such effect is proportional to the aggregate policy-induced innovation from the previous period, I can retrieve such coefficient of proportionality as $\hat{\kappa}^B = \frac{\Delta\hat{P}^{A,E,B}_{j,4}}{\Delta\hat{P}^{A}_{j,3}}$, and use it to infer $\Delta\hat{P}^{A,I,B}_{j,4}$, i.e. the between-firms component of the aggregate policy impact for incumbent firms in period (4), as $\Delta\hat{P}^{A,I,B}_{j,4} = \hat{\kappa}^B \frac{\hat{s}^I_4}{1-\hat{s}^I_4}\Delta\hat{P}^{A}_{j,3}$. The final step is to residu-

Figure B.43: **Direct marginal effect of 1 more day of protection on citations-weighted patents**



The plot shows in black the β_k coefficients of the augmented DiD specification (5) and in gray the β_k coefficients of specification (1), having as dependent variable quarter-t and field-j number of citations-weighted granted patents. Standard errors are clustered by technical field and 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

ally infer from expression (35) the contribution of the within-firm internality in period 4 as $\Delta\hat{P}^{A,I,W}_{j,4} = \Delta\hat{P}^{A,I}_{j,4} - \Delta\hat{P}^{A,I,T}_{j,4} - \Delta\hat{P}^{A,I,B}_{j,4}$, where all the terms on the right hand side are known from previous calculations.

Therefore, the aggregate policy impact in period (4) can be decomposed as

$$\Delta\hat{P}^{A}_{j,4} = \underbrace{\Delta\hat{P}^{A,I,T}_{j,4} + \Delta\hat{P}^{A,E,T}_{j,4}}_{\text{Direct policy effect}} + \underbrace{\Delta\hat{P}^{A,I,B}_{j,4} + \Delta\hat{P}^{A,E,B}_{j,4}}_{\text{Between-firms spillover}} + \underbrace{\Delta\hat{P}^{A,I,W}_{j,4}}_{\text{Within-firm internality}}$$

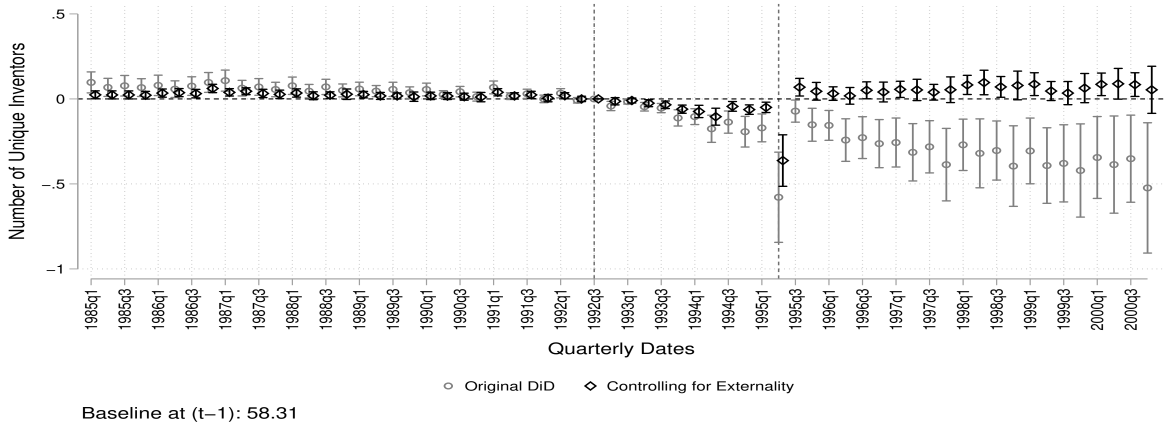
B.4.10 Additional results from the augmented DiD specification

In this subsection I estimate specification (5), reported below, having as dependent variable either citations-weighted patents or the quarter- and field-specific proxy of R&D effort based on the number of inventors listed on patents.

$$Y_{j,t} = \alpha_j + \sum_k \gamma_k \mathbf{1}_{(t=k)} + \sum_k \beta_k \mathbf{1}_{(t=k)} T_j + \sum_k \psi_k \mathbf{1}_{(t=k)} \underbrace{\bar{P}_{j,k-16:k-1}}_{\equiv (1/16)\sum_{q=k-16}^{k-1} P_{j,q}} + \varepsilon_{j,t} \tag{41}$$

Figures B.43 and B.44 report the results and they confirm the findings of subsection 4.4. Specifically, the direct effect of adopting a longer patent length is positive on quality-adjusted innovation and R&D effort after policy implementation. While news of a future patent length extension generates a decline in innovation and R&D until implementation.

Figure B.44: **Direct marginal effect of 1 more day of protection on R&D effort**



The plot shows in black the β_k coefficients of the augmented DiD specification (5) and in gray the β_k coefficients of specification (1), having as dependent variable quarter-t and field-j unique number of inventors listed on granted patents. Standard errors are clustered by technical field and 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

B.4.11 Direct effect of patent length *vs.* Technology disclosure externality in driving the post-implementation effect

The aim of this subsection is to provide further empirical support to the findings of subsection 4.4 of the paper. Specifically, I assess potential endogeneity concerns raised by the introduction of a lagged innovation term in specification (5). Also, I exploit the results of this additional analysis to estimate the elasticity of post-implementation innovation to a 1% change in innovation in the "news" period.

For this analysis, I aggregate the sample in 5 periods: (1) 1981Q1-1984Q4 and (2) 1985Q1-1988Q4 (control periods); (3) 1989Q1-1992Q3 (pre-news period); (4) 1992Q4-1995Q2 (post-news, pre-implementation period); and (5) 1995Q3-2000Q4 (post-implementation period). The primary rationale of this aggregation is the reduction of the computational burden of the instrumental variable analysis, which is very heavy on the quarterly sample due to the high number of external instruments. As a first step, I run the triple-difference specification (3) having as dependent variable the number of patents $P_{j,p}$ classified in field j and applied for in period $p = 1, 2, 3, 4, 5$. I store the field- and period-specific fitted values $\hat{P}_{j,p}$ of this regression, which captures the potentially heterogeneous effect of T_j on patenting, depending on the degree of technological dependence of the field. In the second step, I run the augmented DiD specification

$$P_{j,p} = \alpha_j + \gamma_p + \sum_{k=2}^{5} \beta_{p=k} T_j + \chi T_j + \sum_{k=2}^{5} \theta_{p=k} P_{j,p-1} + \psi P_{j,p-1} + \mathbf{X_{j,t}}\delta + \varepsilon_{j,t} \tag{42}$$

The specification is similar in spirit to (1) but it controls for the period-specific effect of the technological spillover, proxied by the flow of innovation generated in the same field during the previous period. α_j are field fixed effects, γ_p are period fixed effects, $\beta_{p=k}$ capture the period-specific impact of the change in patent length T_j, and $\theta_{p=k}$ captures the period-specific effect of the externality term $P_{j,p-1}$, in excess of the baseline impact ψ. $\mathbf{X_{j,t}}$ includes the within-field backward citations intensity variable interacted with period-specific dummies. Since the flow of innovation in the previous period might be endogenous, I instrument $P_{j,p-1}$ by $\hat{P}_{j,p-1}$, the fitted values of specification (3) run on the aggregated sample.

Table B.5 reports the estimation results. The first column shows, for sake of comparison, the estimates of the baseline DiD specification (1) on the aggregated sample. The second column reports the result of the first step of the exercise, i.e. the OLS estimates of (3). The third and the fourth columns report the OLS and IV estimates of (42), respectively. The fifth column reports the results of one of the first stage regressions. There are three takeaways from the IV estimates, and the red coefficients highlight them.

First, the direct effect of patent length on innovation during the pre-implementation period is unaffected by the inclusion of the externality term. It remains negative and close to the baseline DiD estimate of column (1). Second, the externality term is positive, statistically significant, and economically sizable in the post-implementation period only. A policy-induced drop of one patent during the pre-implementation period 1992Q4-1995Q2 causes a drop of around two patents in the post-implementation period 1995Q3-2000Q4. Subsection B.4.12 describes how to infer the elasticity of future innovation to current innovation shocks from this estimate. Third, the *direct* effect of patent length on innovation is *positive* in sign in the post-implementation period, once controlling for the impact of the spillover. The coefficient estimate of 0.275 is not statistically significant but is economically sizable. It implies that a 1-month (30-days) increase of patent length generates additional 8.25 patents per field in the period, which is equivalent to around 1.3% of the pre-news baseline level. The corresponding elasticity estimate is 2.7, which is very close to the one obtained using the coefficients of Figure 9 in the paper. This suggests that a longer patent length promotes innovation. Moreover, it provides support to the interpretation that innovators reduce the pace of innovation at news of a future patent term increase because they want to profit of the relatively longer protection available after implementation.

B.4.12 The elasticity of future innovation to current innovation shocks

The results of Table B.5 allow to estimate the elasticity of future innovation to current innovation shocks. Overall, the table captures that the policy-driven drop in innovation after implementation can be directly related to the initial fall in innovation caused by the news of a future patent term extension. Therefore, this is instructive about the impact of current patents on

Table B.5: **Decomposition of the post-implementation effect**

	(1) (D.i.D.) $Patents$	(2) (T.D.) $Patents$	(3) (OLS) $Patents$	(4) (IV) $Patents$	(5) (F.S.) $Patents_{t-1}$
$d_{85-88} \times T_j$	0.07628	-0.06653	-0.12583	-0.06501	0.00554
	(0.10747)	(0.56583)	(0.10168)	(0.05139)	(0.02893)
$d_{93-95} \times T_j$	**-0.63075*****	-0.23375	-0.62522***	**-0.59694*****	0.03527
	(0.16061)	(0.56787)	(0.15547)	(0.11668)	(0.03235)
$d_{96-00} \times T_j$	-3.43259**	-0.47862	0.18472	**0.27505**	-0.03125
	(1.34928)	(0.55725)	(0.44764)	(0.34143)	(0.06666)
$d_{85-88} \times T_j \times S_j$		0.00587			
		(0.01050)			
$d_{93-95} \times T_j \times S_j$		-0.01330			
		(0.01051)			
$d_{96-00} \times T_j \times S_j$		-0.09918***			
		(0.01045)			
$d_{85-88} \times Patents_{t-1}$			0.07163	-0.05456	
			(0.09368)	(0.34333)	
$d_{93-95} \times Patents_{t-1}$			-0.56134***	**-0.28330**	
			(0.21563)	(0.46021)	
$d_{96-00} \times Patents_{t-1}$			1.24929***	**1.99416*****	
			(0.17705)	(0.75624)	
$Patents_{t-1}$			2.84560***	1.75987	
			(0.75790)	(1.70780)	
$d_{85-88} \times \widehat{Patents}_{t-1}$					-0.08350***
					(0.01111)
$d_{93-95} \times \widehat{Patents}_{t-1}$					0.16842***
					(0.01248)
$d_{96-00} \times \widehat{Patents}_{t-1}$					0.27283***
					(0.03264)
$\widehat{Patents}_{t-1}$					0.52965***
					(0.11175)
Period F.E.	Y	Y	Y	Y	Y
Field F.E.	Y	Y	Y	Y	Y
Observations	3067	3067	2434	2434	2434

The table reports the estimates of several specifications run on a sample aggregated into five periods: (1) 1981Q1-1984Q4; (2) 1985Q1-1988Q4; (3) 1989Q1-1992Q3; (4) 1992Q4-1995Q2; and (5) 1995Q3-2000Q4. Column (1) reports the OLS estimates of specification (1). Column (2) reports the OLS estimates of specification (3). Column (3) reports the OLS estimates of specification (42). Column (4) reports the IV estimates of specification (42), where the $P_{j,p-1}$ terms are instrumented using the fitted values of the specification of column (2). Column (5) reports the results of one of the first-stage regressions, having as dependent variable the uninteracted term $P_{j,p-1}$. In all columns, standard errors are clustered by technical field. Statistical significance levels: $^{*}(p < 0.10),^{**} (p < 0.05),^{***} (p < 0.01)$

future patents' production. Quantitatively, 1 more patent during pre-implementation period 1992Q4-1995Q2 causes a drop of around 2 patents in the same technical field during post-implementation period 1995Q3-2000Q4. Taking as a baseline the number of patents by field during pre-news period 1989Q1-1992Q3 and re-scaling the effects by the different lengths of pre- and post-implementation phases, the estimate implies an elasticity of 0.997 (95% confidence bands: [0.256; 1.738]).[98].

B.5 Evidence by industry

B.5.1 The effect of innovation on welfare and TFP

In this subsection, I provide evidence on the static effect of innovation (measured by patents, citations-weighted patents, and patent value) on productivity and welfare. I perform a sectoral-level analysis using data from the NBER CES manufacturing database and taking the finest available definition of sector, i.e. the 6-digit NAICS.[99]

The (inverse) measure of welfare is the value of shipments deflator, and productivity is measured by the 5-factors TFP. For the analysis, the time dimension of the panel is restricted to the policy-related window 1985-2000. The aim of the empirical strategy is to isolate the static impact of innovation on welfare and productivity by exploiting the variation in innovation induced by the news and the subsequent implementation of the TRIPs-related patent term change. The second stage regression is

$$y_{s,t} = \alpha_s + \sum_{k=1985}^{2000} \gamma_k \mathbf{1}_{(t=k)} + \beta I_{s,t} + \Xi X_{s,t} + \varepsilon_{s,t} \tag{43}$$

where $y_{s,t}$ denotes the natural logarithm of either the value of shipments deflator or TFP for industry s in year t, α_s are industry fixed effects, $\mathbf{1}_{(t=k)}$ are yearly dummy variables, $X_{s,t}$ is a matrix of controls that include: 4-digit NAICS industry $\times$ yearly effects, and the log of the energy price deflator and energy consumption. $\varepsilon_{s,t}$ is an idiosyncratic error term. (43) is estimated by weighted least squares. Weights are the mass of patents produced in the sector in 1985, to take into account heterogeneous innovation-related industry sizes. $I_{s,t}$ is the innovation measure for industry s and year t. To aggregate measures of innovation by 6-digit

[98]The average number of patents by field during pre-news 1989Q1-1992Q3 period is 425, which implies an average of 28.3 patents per quarter. Taking such quarterly baseline and considering that there are 11 quarters in pre-implementation phase and 22 quarters in post-implementation phase, I estimate the elasticity as $\frac{1.99/(28.3\times22)}{1/(28.3\times11}$. The confidence bands for the elasticity are computed by replacing the extremes of 95% confidence intervals to the point estimate in the previous formula.

[99]An example of the depth of the sectoral classification I use in the analysis is the following. *31-33* is the aggregate 2-digit classification for *Manufacturing*; *324* is the 3-digit *Petroleum and Coal Products Manufacturing*, *3241* is the 4-digit *Petroleum and Coal Products Manufacturing*; which includes the 5-digit *32412 Asphalt Paving, Roofing, and Saturated Materials Manufacturing*, which includes the 6-digit sectors 324121 *Asphalt Paving Mixture and Block Manufacturing* and 324122 *Asphalt Shingle and Coating Materials Manufacturing*.

NAICS and year, I start from measures of innovation by technical field at the yearly level, and I map them into 6-digit NAICS through the formula $I_{s,t} = \sum_j I_{j,t}\pi_{s|j}$. $I_{j,t}$ is innovation in 4-digit IPC field j and year t, and $\pi_{s|j}$ is the probability that a patent classified in technical field j is linked to sector s. $\pi_{s|j}$ is taken from the 'Algorithmic Links with Probabilities' crosswalk by Goldschlag, Lybbert and Zolas (2019).

The first stage regression is

$$I_{s,t} = \kappa_s + \sum_{k=1985}^{2000} \iota_k \mathbf{1}_{(t=k)} + \sum_{k=1985}^{2000} \psi_k \mathbf{1}_{(t=k)} T_s + \Lambda X_{s,t} + u_{s,t} \tag{44}$$

where the LHS innovation variables used are: granted patents, citations-weighted patents, or patent value. T_s is the policy-induced change in protection time in sector s. The technical field level treatment T_j is converted into a sectoral treatment T_s by the formula $T_s = \sum_j T_j \pi_{j|s}$, where $\pi_{j|s}$ is the probability that, given that a patent is assigned to NAICS s, it comes from technical field j. These probabilities are again taken from Goldschlag, Lybbert and Zolas (2019).

Table B.6 reports the estimated impact of innovation on the natural logarithm of the value of shipments price deflator. An increase of 100 patents by industry$\times$ year implies a 2.7% reduction of the value of shipment deflator.[100] Similarly, an increase of 1,000 citations-weighted patents by industry$\times$ year implies 1.5% lower value of shipment deflator.[101] The F-statistics for the first stage regressions are always above 10.

Table B.7 reports the estimated impact of innovation on the natural logarithm of the 5-factors TFP estimated by the NBER. An increase of 100 patents by industry$\times$ year implies a 3.3% increase of TFP.[102] Similarly, an increase of 1,000 citations-weighted patents by industry$\times$ yea implies 1.8% higher productivity.r[103] The F-statistics for the first stage regressions are always above 10.

Pass-through of productivity gains The ratio of the estimated impact of innovation on TFP to the estimated impact of innovation on the value of shipments deflator (flipped in sign) is informative about the pass-through of productivity gains into higher consumer's welfare. Estimates of Tables B.6 and B.7 imply a pass-trough of TFP gains from more patents into lower prices of around 83%; using citations-weighted patents, the figure is 84%; using the

[100] The sectoral average of industry$\times$ year patents in the pre-treatment year 1991 is 280.

[101] The sectoral average of industry$\times$year citations-weighted patents in the pre-treatment year 1991 is approximately 1,450.

[102] The sectoral average of industry$\times$ year patents in the pre-treatment year 1991 is 280.

[103] The sectoral average of industry$\times$ year citations-weighted patents in the pre-treatment year 1991 is approximately 1,450.

Table B.6: **Sectoral evidence on prices**

	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
Patents/100	-0.024***	-0.027***				
	(0.003)	(0.008)				
Citations/1000			-0.012***	-0.015***		
			(0.001)	(0.003)		
Patent value (M)/1000					-0.002***	-0.007**
					(0.001)	(0.003)
6-d NAICS f.e.	Y	Y	Y	Y	Y	Y
Year f.e.	Y	Y	Y	Y	Y	Y
4-d NAICS × Year f.e.	Y	Y	Y	Y	Y	Y
Observations	6684	6684	6684	6684	6684	6684

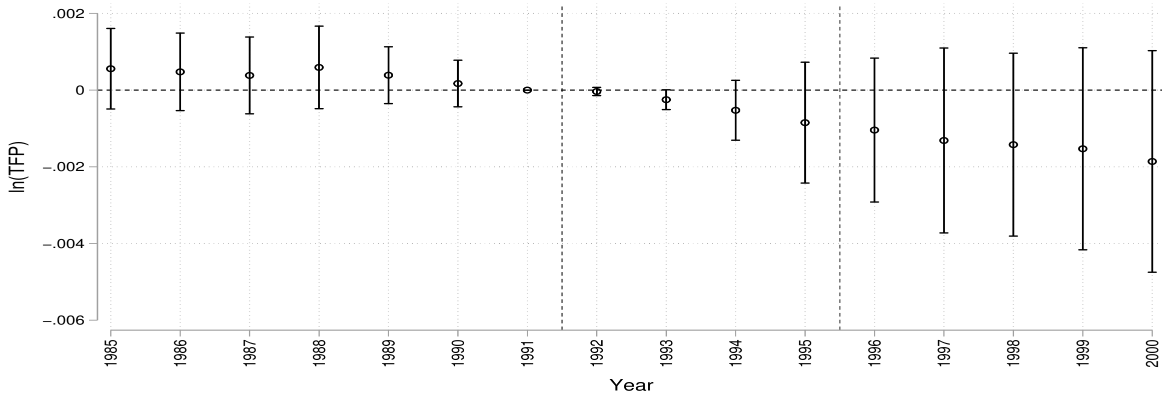
Columns (1), (3), and (5) report the OLS estimates of the β coefficient of specification (43) having as dependent variable the natural logarithm of the price of shipment deflator, normalized to 100 in 1997. Columns (2), (4), and (6) report the 2-stage estimates of the IV regression. Standard errors are clustered by 3-digit NAICS × year. Statistical significance levels: $^{*}(p < 0.05)$,$^{**}(p < 0.01)$,$^{***}(p < 0.001)$

Table B.7: **Sectoral evidence on TFP**

	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
Patents/100	0.027***	0.033***				
	(0.003)	(0.008)				
Citations/1000			0.013***	0.018***		
			(0.001)	(0.003)		
Patent value (M)/1000					0.003***	0.009**
					(0.001)	(0.004)
6-d NAICS f.e.	Y	Y	Y	Y	Y	Y
Year f.e.	Y	Y	Y	Y	Y	Y
4-d NAICS × Year f.e.	Y	Y	Y	Y	Y	Y
Observations	6684	6684	6684	6684	6684	6684

Columns (1), (3), and (5) report the OLS estimates of the β coefficient of specification (43) having as dependent variable the natural logarithm of the 5-factors TFP, normalized to 100 in 1997. Columns (2), (4), and (6) report the 2-stage estimates of the IV regression. Standard errors are clustered by 3-digit NAICS × year. Statistical significance levels: $^{*}(p < 0.05)$,$^{**}(p < 0.01)$,$^{***}(p < 0.001)$

Figure B.45: **Marginal effect of 1 more day of protection on sectoral TFP**



The plot shows the β_k coefficients of regression (45) having as dependent variable the natural logarithm of TFP in sector s and year t. Clustered 95% confidence bands by 3-digit NAICS industry and year are plotted. The first vertical line lies just before the news year (1992) and the second vertical line lies just before the implementation year (1995).

private economic value of patents, it is 76%. So, the pass-through implied by the IV estimates is high and quite stable across innovation measures.

B.5.2 The dynamic effect of the policy on welfare and TFP

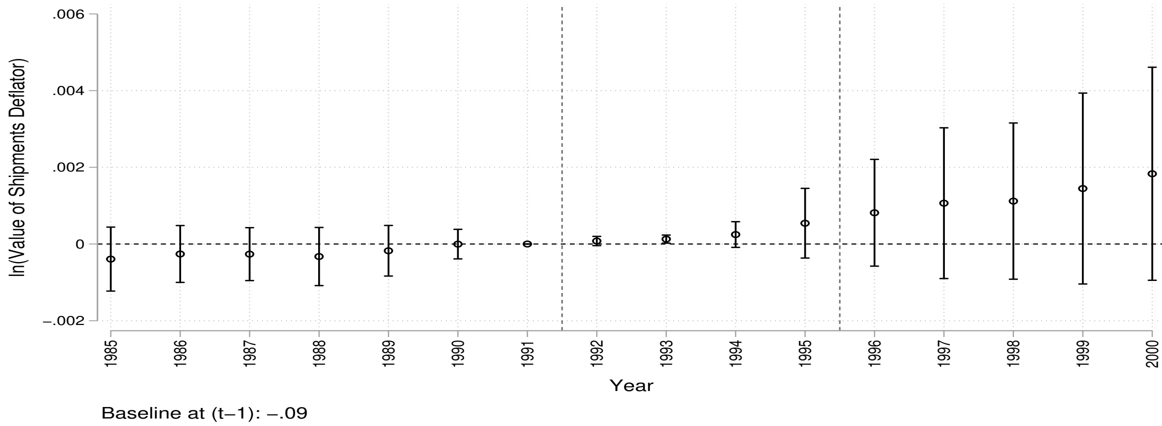
This subsection performs a difference-in-difference analysis by industry to study the dynamic effect of the policy on welfare and TFP. Measures of welfare and productivity and the sectoral treatment are the same as in the previous subsection. The policy-relevant time-window is 1985-2000 and the specification of the regression is

$$y_{s,t} = \alpha_s + \sum_{k=1985}^{2000} \gamma_k \mathbf{1}_{(t=k)} + \sum_{k=1985}^{2000} \beta_k \mathbf{1}_{(t=k)} T_s + \Xi X_{s,t} \varepsilon_{s,t} \tag{45}$$

where the dependent variable $y_{s,t}$ for sector s and year t is the natural logarithm of either of the two outcomes described above, α_s are industry fixed effects, $\mathbf{1}_{(t=k)}$ denotes yearly dummies, $X_{s,t}$ is matrix of controls that include: 4-digit NAICS industry $\times$ year effects, and the natural logs of the energy price deflator and material costs deflator. T_s is the sectoral treatment and $\varepsilon_{s,t}$ is the error term. (45) is estimated by weighted least squares. Weights are the mass of patents produced in the sector in 1985, to take into account heterogeneous innovation-related industry sizes.

Figure B.45 plots the difference-in-difference coefficients capturing the dynamic effect of a change in patent protection length on the logarithm of sectoral TFP. For $T_s < 0$, the dynamics of the point-estimates–which are very close to zero at first and grow larger as time goes–is consistent with a slow positive impact of higher policy-induced innovation on the level of productivity.

Figure B.46: **Marginal effect of 1 more day of protection on sectoral value of shipments deflator**



The plot shows the β_k coefficients of regression (45) having as dependent variable the natural logarithm of the values of shipments deflator in sector s and year t. Clustered 95% confidence bands by 3-digit NAICS industry and year are plotted. The first vertical line lies just before the news year (1992) and the second vertical line lies just before the implementation year (1995).

Figure B.46 plots the difference-in-difference coefficients capturing the dynamic effect of a change in patent protection length on the logarithm of the value of shipments deflator, which is the (inverse) measure of welfare. Welfare gains from innovation take time to be achieved, exactly because TFP gains are slow.

Appendix C Additional theoretical results

C.1 Model derivations

This section presents the details and the derivations of the model of Section 5.

C.1.1 Consumers

The consumer has linear utility $u(c(t)) = c(t)$ in per-capita consumption $c(t)$, invests in real-assets $a(t)$, and inelastically supplies labor. The maximization problem of the representative agent is

$$\max_{c(t),a(t)} \int_0^\infty e^{-\rho} c(t) dt \tag{46}$$

$$\text{subject to} \quad \dot{a}(t) = r(t)a(t) - c(t) + w(t) \tag{47}$$

where is defined as aggregate consumption divided by population, i.e. $c(t) \equiv C(t)/L(t)$, and $a(t) \equiv A(t)/L(t)$ are total assets per capita. Aggregate real assets are and $A(t) \equiv K(t) + \int_0^\infty v(t-s,t)S(t-s,t)ds$. The first term is the total stock of physical capital. The second term represent the total value of firms owning patents on profit-generating intermediate capital varieties.

In particular, the last term is defined by the following expressions

$$\upsilon(t-s,t) = \int_{t}^{t-s+T} \pi(t')e^{-\int_{t'}^{t-s+T}(\rho+\lambda(z))dz}dt' \text{ if } s \leq T$$
$$\upsilon(t-s,t) = 0 \text{ if } s \geq T$$

is the residual value at time t of a patent generated at time $t-s$. While the term $S(t-s,t)$ represents the mass of patents generates at time $t-s$ that have not been creatively destroyed up to time t. This is defined by the expression

$$S(t-s,t) = (1+\psi)\dot{V}(t-s)e^{-\int_{t-s}^{t}\lambda(t')dt'}$$

No arbitrage conditions ensure that all the real assets give a net return equal to t. The solution of problem (46) gives the Euler equation $r(t) = \rho$.

C.1.2 Final good production

The final good is produced by a competitive firm that optimally chooses labor and each of the intermediates to maximize profits. The problem is

$$\max_{\{X(i,t)\}_{i\in[0,V(t)]},L(t)\}} \Big[h(t)L(t)\Big]^{1-\alpha}\Big[\int_{0}^{V(t)} X^{\alpha}(i,t)di\Big] - \int_{0}^{V(t)} z(i,t)X(i,t)di - w(t)L(t)$$

$$Y(t) = \Big[h(t)L(t)\Big]^{1-\alpha}\Big[\int_{0}^{V(t)} X^{\alpha}(i,t)di\Big] \tag{48}$$

(48) is the production function. The first order conditions of the problem are

$$w(t) = (1-\alpha)h(t)^{1-\alpha}L(t)^{-\alpha}\Big[\int_{0}^{V(t)} X^{\alpha}(i,t)di\Big] \tag{49}$$

$$z(i,t) = \alpha h(t)^{1-\alpha}L(t)^{1-\alpha}X^{\alpha-1}(i,t) \;\; \forall\, i \in [0,V(t)] \tag{50}$$

(49) is the inverse labor demand and determines the equilibrium wage rate. (50) is the inverse demand for intermediate i.

C.1.3 Monopolistic intermediate goods production

A share $\zeta(t)$ of the existing $V(t)$ intermediate good varieties are protected by a monopoly, granted by a valid patent. The monopolistic producer of variety i chooses the quantity to produce in order to maximize profits subject to the inverse demand given by (50), and subject to the production function: One unit of each of the intermediate goods can be produced by using one unit of raw capital $K(t)$. The latter is rented from households at a rate $r_K(t) = r(t)+\delta$, where δ is the depreciation rate of physical capital. The maximization problem is

$$
\begin{aligned}
&\max_{X(i,t),z(i,t)} \Big\{ z(i,t)X(i,t) - (r(t)+\delta)X(i,t) \Big\} \\
&\text{s.t.} \quad z(i,t) = \alpha h(t)^{1-\alpha} L(t)^{1-\alpha} X^{\alpha-1}(i,t)
\end{aligned}
$$

and the first order condition implies

$$
z(i,t) = \alpha (h(t)L(t))^{1-\alpha} X(i,t)^{\alpha-1} = \frac{1}{\alpha}(r(t)+\delta) \tag{51}
$$

i.e. the price is a constant markup $1/\alpha$ over the marginal cost $(r(t)+\delta)$. The produced quantity and the profits are symmetric across monopolistic i's and satisfy

$$
X(i,t) = X_p(t) = \alpha^{\frac{2}{1-\alpha}} (r(t)+\delta)^{-\frac{1}{1-\alpha}} h(t)L(t) \quad \forall\, i \in [0, \zeta(t)V(t)] \tag{52}
$$

$$
\pi(i,t) = \pi(t) = \Big(\frac{1}{\alpha} - 1\Big)(r(t)+\delta)X_p(t) \tag{53}
$$

C.1.4 Non-monopolistic intermediate goods production

A fraction $1-\zeta(t)$ of intermediates are competitively produced because legal patent protection on them has expired after the maximum patent length T. These non-monopolistic varieties are produced in a regime of Bertrand competition and therefore the price $z(i,t)$ is driven to the marginal cost of production $(r(t)+\delta)$. It follows from the inverse demand function (50) that the production of these competitively-produced intermediate varieties is symmetric and given by

$$
X_{np}(t) = \alpha^{\frac{1}{1-\alpha}} (r(t)+\delta)^{-\frac{1}{1-\alpha}} h(t)L(t) \quad \forall\, i \in (\zeta(t)V(t), V(t)] \tag{54}
$$

which implies that $X_p(t) = \alpha X_{np}(t)$. Since $\alpha \in (0,1)$ by assumption, this implies that the quantity produced of monopolistic varieties is lower than the one of competitive varieties, which is the main distortion from monopoly in the model.

C.1.5 Physical capital market clearing condition

Physical capital market clearing requires that the quantity of capital supplied by households $K(t)$ is equal to the quantity of capital demanded by firms to produce the intermediate capital goods, i.e.

$$
\begin{aligned}
K(t) =& \zeta(t)V(t)X_p(t) + (1-\zeta(t))V(t)X_{np}(t) \\
=& [\alpha\zeta(t) + (1-\zeta(t))]V(t)X_{np}(t)
\end{aligned} \tag{55}
$$

C.1.6 Research investment to discover new projects

The model features an unit mass of identical firms that invest in research. The output of research investment is new ideas that need subsequent development by successful firms. The research investment problem of the representative research firm is

$$\max_{I_R(t)} \left\{P(t)E(t)^{\chi}V(t)^{\phi_1}I_R(t)^{\phi_2} - I_R(t)\right\}$$

$P(t)$ is the economic value of a new idea, or, alternatively, it can be thought as the exclusivity value of a development project. The optimal research investment is given by

$$I_R(t) = \left[\phi_2 P(t)E(t)^{\chi}V(t)^{\phi_1}\right]^{\frac{1}{1-\phi_2}}$$

C.1.7 Investment in development of projects

Development occurs independently on each existing project, even in the case of a single firm running multiple projects. The project-specific maximization problem can be written in recursive form as

$$r(t)P(t) - \dot{P}(t) = \max_{\iota_D(t)} \left\{\iota_D(t)\left[\upsilon(t) - P(t)\right] - \mu\iota_D(t)^{\theta}\upsilon(t)\right\} \tag{56}$$

where the equation captures the fact that if the project is successful with instantaneous probability $\iota_D(t)$, the firm receives a value $\upsilon(t)$ for the intermediate variety obtained but it loses the value of the project $P(t)$, which expires after completion. $\upsilon(t)$ is the value of a patent on a variety, and it is defined by (11) in the paper. The optimal development project completion rate is

$$\iota_D(t) = \left[\frac{1}{\theta\mu}\left(1 - \frac{P(t)}{\upsilon(t)}\right)\right]^{\frac{1}{\theta-1}} \tag{57}$$

The process of creative destruction captured by the $\lambda(t)$ term is endogenous, and it is driven by the rate of growth of the number of varieties $V(t)$. It is defined as $\lambda(t) \equiv \psi\frac{\dot{V}(t)}{V(t)}$, i.e. in times when the rate of growth of varieties is higher, the rate of creative destruction is higher.

C.1.8 Evolution of aggregate quantities

Previous optimal policies determine the evolution of aggregate quantities. First, the number of varieties $V(t)$ evolves according to

$$(1+\psi)\dot{V}(t) = \iota_D(t)N(t) \tag{58}$$

where $\psi\dot{V}(t)$ is by how much creative destruction reduces the mass of intermediate goods available, and $\iota_D(t)N(t)$ is the number of development projects successfully turned into a variety. $\iota_D(t)$ is the instantaneous probability that each of the existing projects $N(t)$ is successfully completed. Since it is identical and independent across projects, a suitable law of large numbers applies and the aggregate representation provided holds. Second, the evolution of projects is given by

$$\dot{N}(t) = E(t)^{\chi}V(t)^{\phi_1}I_R(t)^{\phi_2} - \iota_D(t)N(t) \tag{59}$$

where the first term captures the mass of new projects generated by research investment, and the second term captures the destruction of projects due to successful completion.

The evolution of the share of existing varieties that are covered by monopoly, i.e. $\zeta(t)$, is given by

$$\dot{\zeta}(t) = (1 - \zeta(t))\frac{\dot{V}(t)}{V(t)} - (1 + \psi)\frac{\dot{V}(t - T)}{V(t)}e^{-\int_{t-T}^{t}\lambda(t')\mathrm{d}t'} \tag{60}$$

where the first term captures the additions to monopolistic varieties due to new patented innovations, and the second term captures the fact that all those varieties that have not already been creatively destroyed become competitive when the maximum patent term T expires.

The derivation of equation (60) is the following. Let $V_p(t)$ be the mass of existing varieties covered by monopoly. Then, $\zeta(t) \equiv \frac{V_p(t)}{V(t)}$. Reshuffling the definition of $\zeta(t)$ and taking time derivatives, we get

$$\dot{\zeta}(t)V(t) + \zeta(t)\dot{V}(t) = \dot{V}_p(t)$$

$\dot{V}_p(t)$ is given by the inflow of new varieties in the stock of monopolistic ones, minus the outflow from this stock, due to the expiration of the maximum patent term T. This is what needs to be derived. At every instant t, the gross production of new varieties–which are obviously monopolistic at their creation–is given by $(1 + \psi)\dot{V}(t) = \iota_D(t)N(t)$. I assume for simplicity that all creatively destroyed varieties $\psi\dot{V}(t)$ come from the pool of monopolistic ones. This simplifies things because it implies that the addition to the stock of existing varieties, net of creative destruction, $\dot{V}(t)$ also coincides with the net addition to the stock of monopolistic varieties $V_p(t)$. Therefore, the net inflow component of $\dot{V}_p(t)$ is simply $\dot{V}(t)$. As to the outflow, we need to consider that the mass of varieties of vintage $t - T$, i.e. $(1 + \psi)\dot{V}(t - T)$, which go out of monopoly at instant t, has been eroded by creative destruction over time. Let $S(t_0)$ be the stock of such patents issued at time t_0. In practice, take $S(t_0) = (1 + \psi)\dot{V}(t_0)$. Due to creative destruction, the evolution of this stock responds to the following law of motion: $S(t + dt) = S(t) - (\lambda(t)dt)S(t)$, which can be re-written as a first order differential equation $\dot{S}(t) = \lambda(t)S(t)$. Its solution, for two generic points in time t_0 and t_1, is $S(t_1) = S(t_0)e^{-\int_{t_0}^{t_1}\lambda(t')dt'}$. Now, the outflow from the mass of monopolistic varieties is given by the residual mass of gross varieties produced at $t - T$ and survived from $t - T$ up to t. Therefore, replacing $S(t_0) = (1 + \psi)\dot{V}(t_0)$, $t_0 = t - T$, and $t_1 = t$, we get that the outflow from the mass of monopolistic varieties is the right hand side of the last equation. i.e. $(1 + \psi)\dot{V}(t - T)e^{-\int_{t-T}^{t}\lambda(t')dt'}$. Therefore,

$$\dot{\zeta}(t)V(t) + \zeta(t)\dot{V}(t) = \dot{V}(t) - (1 + \psi)\dot{V}(t - T)e^{-\int_{t-T}^{t}\lambda(t')dt'}$$

Moving the second LHS addend to the right, and dividing everything by $V(t)$, we get exactly (60).

The evolution of aggregate capital satisfies $\dot{K}(t) = I_K(t) - \delta K(t)$, where $I_K(t)$ is the investment in physical capital done by the households out of the final good, and $\delta K(t)$ is the depreciation of the existing stock.

C.1.9 Market clearing in the goods market

Given the production decisions of intermediate producers and final good producers, GDP for this economy can be rewritten as

$$Y(t) = [\alpha^\alpha \zeta(t) + (1 - \zeta(t))]V(t)h(t)^{1-\alpha}L(t)^{1-\alpha}X_{np}^\alpha(t) \tag{61}$$

where $[\alpha^\alpha \zeta(t) + (1 - \zeta(t))]V(t)h(t)^{1-\alpha}$ is the measured TFP. The productivity of the economy grows with the number of varieties available, and decreases with the share of monopolistic varieties, as $\alpha^\alpha < 1$. On the other hand, the total production of the final good must also satisfy the resource constraint

$$Y(t) = C(t) + I_K(t) + I_R(t) + \mu\iota_D(t)^\theta \upsilon(t)N(t) \tag{62}$$

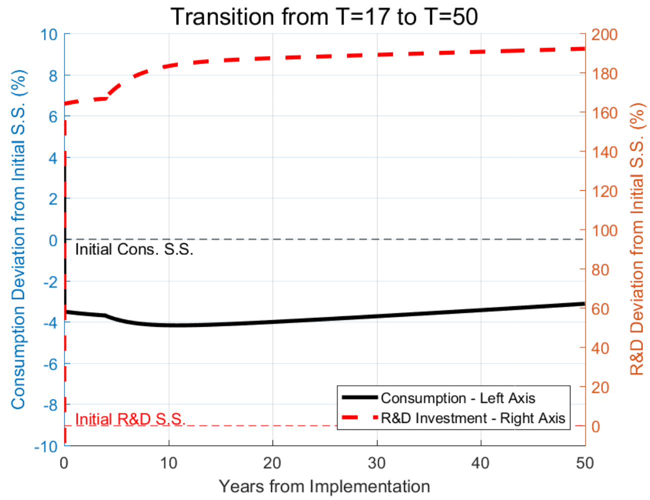
C.1.10 Balanced growth path

Population $L(t)$ and the productivity term $h(t)$ exogenously grow at constant rate n and g_h, respectively. Since $r(t) = \rho$, the real interest rate is constant. From equations (52), (54), and (53) the growth rate of $X_p(t)$, $X_{np}(t)$, and profits is identical in the b.g.p and equal to $g_h + n$. From the definition of $\upsilon(t)$, the patent value must grow at the same rate of profits. In addition, the rate of creative destruction $\lambda(t)$ is constant along the balanced growth path. From the value function of the development investment problem, $P(t)$ must grow at the same rate of $\upsilon(t)$, i.e. $g_P = n+g_h$, and the development speed $\iota_D(t)$ must be constant. A constant $\iota_D(t)$ also implies that the externality term $E(t)$ is constant in the b.g.p. The evolution of $V(t)$ in (58) implies that $g_V = g_N$, and the evolution of $N(t)$ in (59) requires that $g_N = \phi_1 g_V + \phi_2 g_{I_R}$. From (62), the rate of growth of $C(t)$, $I_K(t)$, and $I_R(t)$ must be the same as output, i.e. $g_Y = g_C = g_{I_K} = g_{I_R}$. In addition, from (60), $\zeta(t)$ is constant in the b.g.p., and therefore the equilibrium production function (61) requires $g_Y = (1 - \alpha)(g_h + n) + g_V + \alpha g_X$. Since $g_X = n + g_h$, we then obtain $g_Y = g_V + n + g_h$. Using $g_{I_R} = g_Y$ and $g_V = g_N$, and plugging the last expression into $g_N = \phi_1 g_V + \phi_2 g_{I_R}$, g_V can be solved as $g_V = \frac{\phi_2}{1-\phi_1-\phi_2}(n + g_h)$. The latter can be used to solve explicitly for all the other growth rates.

C.2 Transitional dynamics computational algorithm

I solve the stationary version of the system, which can be obtained by re-scaling each variable by its growth rate along the b.g.p, as computed in subsection C.1.10. The stationary version of the variables of the model is denoted with a tilde, as in the main text. In any solution of the model $r(t) = \rho$, which gives an explicit solution for the full dynamic path of $\tilde{X}_p(t)$, $\tilde{X}_{np}(t)$,

Figure C.1: **Transition dynamics of a change from $T = 17$ to $T = 50$**



The figure shows the path of consumption and total R&D investment $I_R + \mu\iota_D^{\theta}\upsilon N$ in the transition from the status quo $T = 17$ to $T = 50$, compared to the pre-change steady state.

and $\tilde{\pi}(t)$. For the other variables, I setup a mesh that goes from $t_0 = 0$ to $t_{max} = 2000$, and I assume that *i*) just before t_0, the stationary version of the system is in the pre-policy news steady state, and *ii*) by t_{max} it has reached the post-policy change steady state.

I start from a guess of $\tilde{\lambda}(t)$ from $t_0 = 0$ to $t_{max} = 2000$, which I initially fix to be equal to $\psi g_V = \psi \frac{g_h+n}{1-\phi_1-\phi_2}$ at any time. Given $\tilde{\lambda}(t)$, I can solve for the full dynamic path of $\tilde{\upsilon}(t)$ using equation (11). I impose the terminal condition on $\tilde{P}(t)$ that it must be at the post-policy steady state at t_{max}. Then, for each $\tilde{P}(t + dt)$ I solve the development investment problem given $\tilde{\upsilon}(t)$, obtaining $\iota_D(t)$ and $\tilde{P}(t)$. I use the full sequence of $\iota_D(t)$ to build the delayed externality term and, given the computed $\tilde{P}(t)$, I solve for the optimal $\tilde{I}_R(t)$ at every instant using the fact that both $\tilde{N}(t)$ and $\tilde{V}(t)$ are assumed to be at the old steady state at t_0, as they are state variables. With all previous objects, I solve forward (14) and (13) obtaining $\tilde{N}(t + dt)$ and $\tilde{V}(t + dt)$ $\forall t$. Also, given the full series of $\tilde{\lambda}(t)$, I solve forward for $\zeta(t)$, again assuming that this state variable is at the pre-policy steady state at t_0. With $\tilde{V}(t)$, $\zeta(t)$, and $X_{np}(t)$, I use the capital market clearing condition to compute the aggregate series for $\tilde{K}(t)$ and, subsequently, the series $\tilde{I}_K(t)$ that is required to sustain $\tilde{K}(t)$, assuming that at t_0 the level of physical capital is at the old steady state. Using the exogenous $\tilde{L}(t)$ and $\tilde{h}(t)$ with $\tilde{X}_{np}(t)$, $\tilde{V}(t)$ and $\zeta(t)$, I solve for $\tilde{Y}(t)$ and $\tilde{C}(t)$ using the resource constraint. As a final step, I use the series $\tilde{V}(t)$ to update the guess for $\tilde{\lambda}(t)$ according to $\tilde{\lambda}(t) = \psi\left(g_V + \frac{\dot{\tilde{V}}(t)}{\tilde{V}(t)}\right)$, and I iterate the previous steps until convergence of the $\tilde{\lambda}(t)$ series.

C.3 Computation of standard errors

The quadratic loss function used for simulated method of moments estimation is $F = \mathbf{g}'(\mathbf{\Gamma})W\mathbf{g}(\mathbf{\Gamma})$, where $\mathbf{\Gamma}$ is the vector of estimated parameters and $\mathbf{g}(\mathbf{\Gamma})$ is the vector of the deviation of model-based moments computed at $\mathbf{\Gamma}$ from the empirically estimated moments. Overall, there are 69 moment restrictions. 33 are the post-announcement reduced-form estimates of the effect of the reform on patenting activity, 33 are the post-announcement reduced-form estimates of the effect of the reform on patent-read R&D effort, and 3 are the long-run moment restrictions on the capital-output ratio, the consumption-output ratio, and the R&D spending-output ratio. In estimation, W is a diagonal matrix giving unit weight to the first 66 moments, and weights 1, 10, and 100, to the K/Y, C/Y, and $R\&D/Y$ long-run restrictions, to correct for their respective scale. The variance-covariance matrix of the estimated parameters for the resulting GMM estimator is

$$\hat{\mathbf{V}} = (D(\hat{\mathbf{\Gamma}})WD'(\hat{\mathbf{\Gamma}}))^{-1}D(\hat{\mathbf{\Gamma}})W\mathbf{g}(\mathbf{\Gamma})\mathbf{g}'(\mathbf{\Gamma})W'D'(\hat{\mathbf{\Gamma}})(D(\hat{\mathbf{\Gamma}})WD'(\hat{\mathbf{\Gamma}}))^{-1}$$

where W was defined above and $D'(\hat{\mathbf{\Gamma}})$ is defined as $D'(\hat{\mathbf{\Gamma}}) = \left.\frac{\partial \mathbf{g}(\mathbf{\Gamma})}{\partial \mathbf{\Gamma}'}\right|_{\mathbf{\Gamma}=\hat{\mathbf{\Gamma}}}$. The latter is computed numerically around the optimal $\hat{\mathbf{\Gamma}}$. The standard errors of the parameters are computed as the square root of the main diagonal elements of $\hat{\mathbf{V}}$.

C.4 The mechanism at work in the model

The plots show the response of various theoretical objects to an anticipated patent term increase of 100 days, starting from a patent term of 17 years. The anticipation is assumed to be 2 years and 8 months, consistently with the TRIPs anticipation. The news is assumed to happen at $t = 0$, and the vertical line refers to the policy implementation. The red horizontal line represents the steady state in the old regime ,and the blue horizontal line is the steady state in the new regime. The black solid lines show the response of the variables of interest.

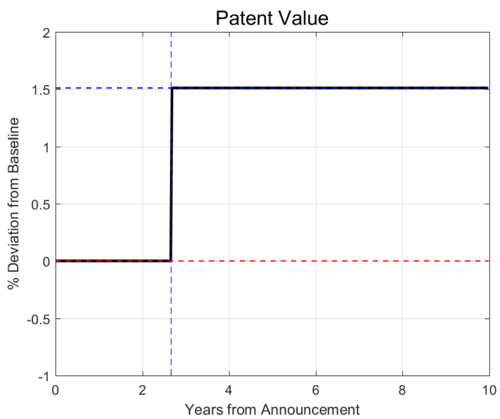
C.5 Identification of key structural parameters

The Figures reported in this subsection support the discussion of subsection 6.1.3 of the paper about the identification of key structural parameters of the innovation process through specific moments of the empirical estimates of Section 3.

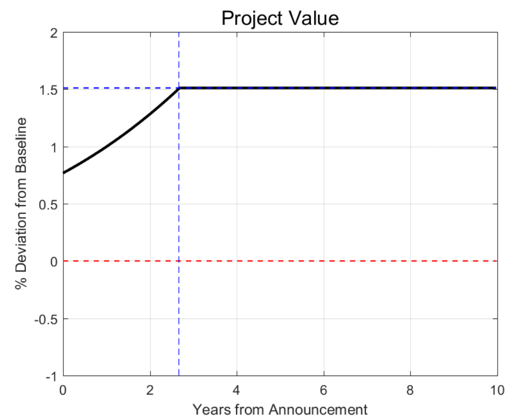
Figure C.3 shows the response of innovation and patent-read R&D effort, in the model and in the data, fixing $\theta = 2$ or $\theta = 1.02$. Raising θ from 1.01, which is the optimal estimate found in Section 6, to $\theta = 2$ implies moving from very mild to quite severe cost convexity of the pace of development, so that the adjustment of the latter is much more costly for innovators. As argued in subsection 6.1.3, this leads to a much more muted response of the innovation and R&D to the news shock in the model. Therefore, the strong response of both variables observed in the data is informative about the mild cost convexity estimated in the model.

Figure C.4 shows the response of innovation and patent-read R&D effort, in the model and in the data, fixing $\chi = 2$ or $\chi = 8$. Lowering χ from 11, which is the optimal estimate

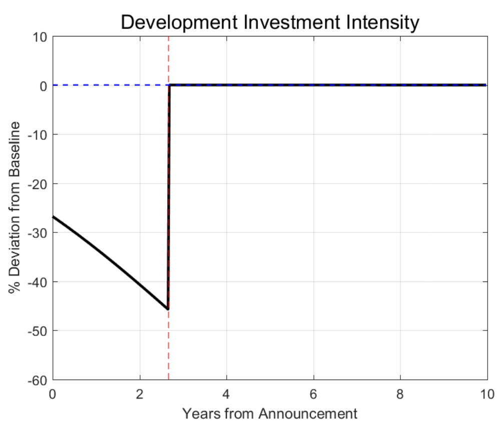
Figure C.2: **Evolution of aggregates**



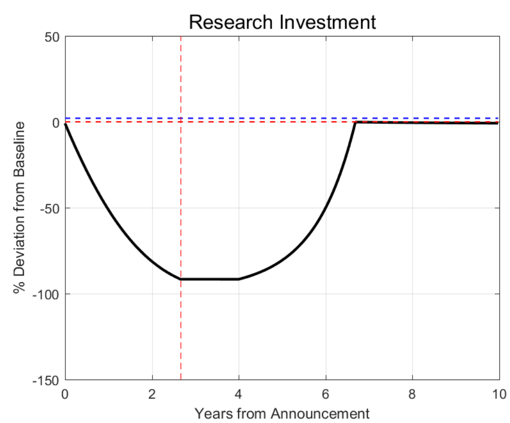
(a) Patent value: $v(t)$



(b) Project value: $P(t)$

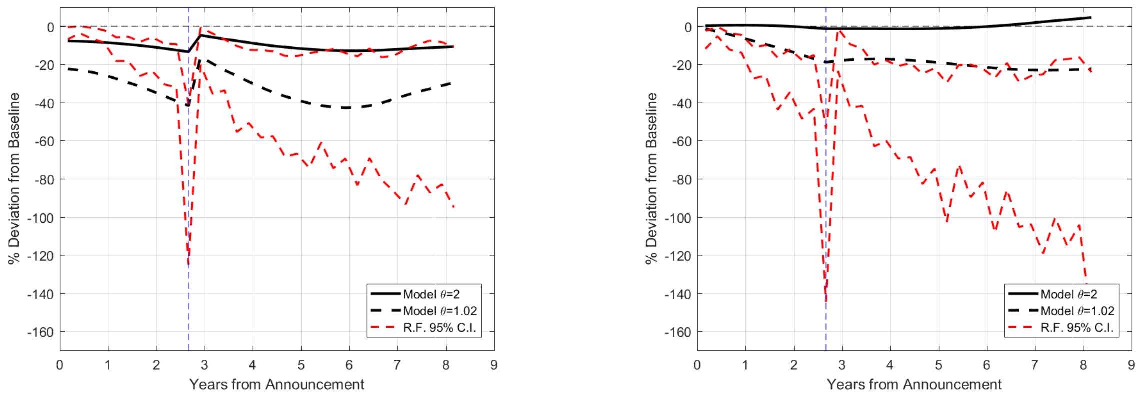


(c) Development speed: $\iota(t)$



(d) Research investment: $I_R(t)$

Figure C.3: **Policy simulation in a model with $\theta = 2$ or $\theta = 1.02$ *vs.* empirical estimates**



(a) Innovation (Patents)

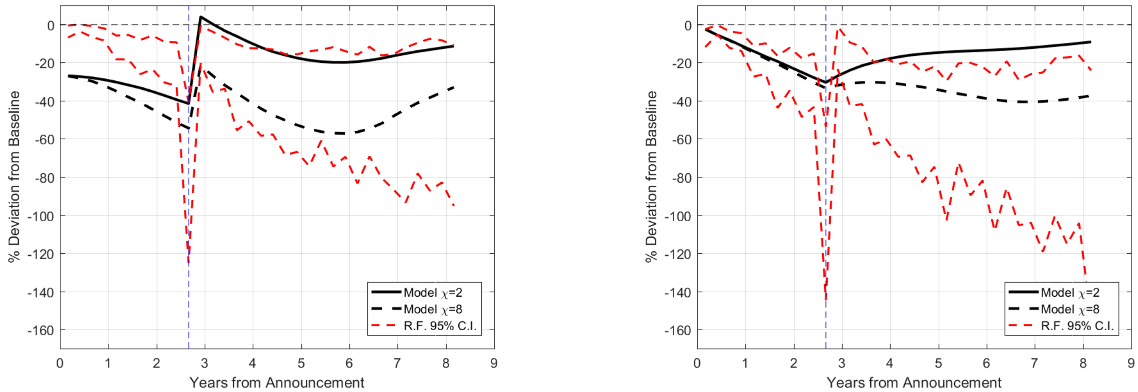
(b) Patent-read R&D effort

The black solid (dashed) lines are the model-based responses of the model with parameter values reported in Table 2, but $\theta = 2$ ($\theta = 1.02$). The red dashed lines are 95% confidence bands of the reduced form estimates of Section 3. The system is assumed to be at the pre-policy change steady state at $t = 0$, when the news of 100-days increase in protection time implemented after 2 years and 8 months (blue vertical line) happens.

found in Section 6, to $\chi = 2$ or $\chi = 8$ implies moving from a very strong externality of development pace on research productivity to a milder effect. As argued in subsection 6.1.3, this leads to a much weaker persistence effect–at the maximum drop after $d = 4$ years–of innovation and R&D in the model during the post-implementation period. Therefore, the strong response of both variables observed in the data is informative about the strength of the spillover, leading to the estimation of a high χ.

Figure C.5 shows the response of innovation and patent-read R&D effort, in the model and in the data, fixing $\phi_1 = 0.08$ and $\phi_2 = 0.69$ or $\phi_1 = 0.5$ and $\phi_2 = 0.26$. This implies changing the estimates for ϕ_1 and ϕ_2 without modifying total returns to research. This imposes severe decreasing returns to the contribution of existing varieties in the production function of ideas–i.e. a weaker "standing on the shoulders of giants" effect–and relaxes the decreasing returns to research investment, while keeping fixed the aggregate returns to all factors. As argued in subsection 6.1.3, weakening the "standing on the shoulders of giants" effect implies a much faster recovery of innovation and R&D from the maximum drop during the post-implementation, *persistence* phase. This is because the rich dynamics observed during the news phase and immediately after the implementation depress the total mass of varieties. As the latter are a slow-moving state variable, their adjustment up to the new, higher steady state takes time and, therefore, it affects productivity of research for long. The higher the importance of existing varieties in research activity, the slower the recovery from low-V. Therefore, the slow empirical recovery from the maximum drop is informative about a relatively high ϕ_1.

Figure C.4: **Policy simulation in a model with $\chi = 2$ or $\chi = 8$ *vs.* empirical estimates**

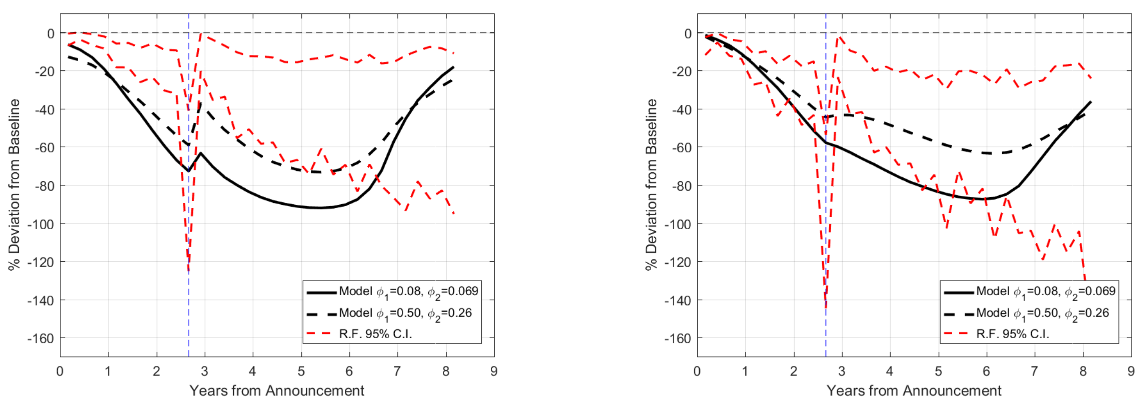


(a) Innovation (Patents)

(b) Patent-read R&D effort

The black solid (dashed) lines are the model-based responses of the model with parameter values reported in Table 2, but $\chi = 2$ ($\chi = 8$). The red dashed lines are 95% confidence bands of the reduced form estimates of Section 3. The system is assumed to be at the pre-policy change steady state at $t = 0$, when the news of 100-days increase in protection time implemented after 2 years and 8 months (blue vertical line) happens.

Figure C.5: **Policy simulation in a model with ($\phi_1 = 0.08$; $\phi_2 = 0.69$) or ($\phi_1 = 0.50$; $\phi_2 = 0.26$) *vs.* empirical estimates**



(a) Innovation (Patents)

(b) Patent-read R&D effort

The black solid (dashed) lines are the model-based responses of the model with parameter values reported in Table 2, but $\phi_1 = 0.08$ and $\phi_2 = 0.69$ ($\phi_1 = 0.50$ and $\phi_2 = 0.26$). The red dashed lines are 95% confidence bands of the reduced form estimates of Section 3. The system is assumed to be at the pre-policy change steady state at $t = 0$, when the news of 100-days increase in protection time implemented after 2 years and 8 months (blue vertical line) happens.

Supplementary Materials

Appendix D Variables construction

D.1 Number of granted patents - Technical field level

Using PATSTAT table `tls201`, which contains information on patent applications, I select "patent of invention" applications for which the reported application authority is the USPTO, and for which the application filing date is the same as the priority date, i.e. the earliest filing date in PATSTAT. This is because, in the main analysis, I want to focus on innovations that primarily refer to the US, excluding technologies that are developed and protected elsewhere at first, and subsequently try to obtain protection in the US too. In addition, I just keep applications tha are subsequently granted. Then, using PATSTAT table `tls209`, I attach to each patent application information on the IPC classes associated to the invention, and I truncate the IPC codes to the 4-digit level. A patent to which multiple 4-digit IPC codes are associated is counted once for each of them in my dataset. Finally, in order to compute the quarterly measure, I build a "customized" quarterly date that better fits the timing of the TRIPs implementation. In particular, since the TRIPs was formally adopted in the US system on December 8, 1994, and the new patent law entered into force on June 8, 1995, I define quarters starting from the eighth day of the month. Hence, for example, 1995Q1 starts on January 8, 1995 and ends on March 7, 1995. A patent is counted in quarter t if its priority date falls in that quarter. Finally, the variable $Pat_{j,t}$ is the total count of granted patent applications satisfying the conditions outlined above, i.e. patents classified in IPC class j, and whose priority date falls in quarter t.

D.2 Number of citations-weighted granted patents - Technical field-level

Citations-weighted patent counts are usually employed to weigh patents by their scientific quality, as measured by their relevance for subsequent technological developments. In order to build this measure, I follow the same steps described in subsection D.1, and I stop before the IPC-quarterly aggregation step. I assign to the selected patent applications the associated patent publications using PATSTAT table `tls211`, which contains publication information. PATSTAT table `tls212`, instead, reports for each publication the list of applications and publications that cite it. I use information in `tls212` to assign to each patent application the publications or applications that cite publications associated with it. I select as a citation date the publishing date of the citing publication, and I just keep citations that occur within 5 years from the application date. This is done to avoid truncation bias in the citation-weighted patent measure. In a robustness check, not reported, I keep citations that occur within 3 years from the application date, and results are fully robust. Finally, I count for each patent application the number of forward citations received, and I build the citation-weighted patent measure by summing this citations count by IPC and quarter of priority date of the focal patent application.

D.3 Pending period and treatment - Technical field-level

In order to build this measure, I follow the same steps described in subsection D.1, and I stop before the IPC-quarterly aggregation step. When I build the treatment variable, I restrict the sample to patents whose priority date is between January 1, 1990 and May 31, 1992, in order to focus on a time-window close enough to the news of the policy change, but also unaffected by it. For this sample of patents, I compute the pending period by counting the number of days between the grant date, i.e. the publication date of the official document granting the patent, and the priority date reported in PATSTAT, which, given my sample restriction, coincides with the application date. I compute an average of such patent-level pending time at the IPC level, and I subtract it to 1095, which is 3 years in number of days. Therefore, the treatment variable is negative if the average pending period computed for applications filed between 1/1/1990 and 5/31/1992 is longer than 3 years, and positive otherwise. When I build the quarterly version of the pending period underlying Figure B.8, I still compute the patent-specific pending time in the same way described above, and I compute its average at the IPC$\times$quarter level, with quarters defined as in subsection D.1. Finally, as a measure of treatment precision that I use to conduct a triple difference analysis, I compute the standard deviation of the average pending period by technical field.

D.4 Patent renewal rate - Technical field-level

To build the patent renewal rate, I use information on legal events related to patents reported in the PATSTAT LEGAL section of PATSTAT and, specifically, in table `tls803`. This dataset reports, for each US granted patent application, whether maintenance fees at 3.5 years, 7.5 years, and 11.5 years since patent grant have been paid in order to maintain the patent active. Therefore, for each patent selected according to the criteria explained in subsection D.1, I can compute whether or not the maintenance fees at 11.5 years since grant have been paid. This indicates whether, for the specific patent, the maximum patent term was binding or not. In order to compute the IPC-specific pre-policy change measure of incidence of the maximum patent term, I focus again only on patents whose priority date is between January 1, 1990 and May 31, 1992, and I average out at the 4-digit IPC level the indicator variable that takes value 1 if the 11.5 years maintenance fees have been paid for a patent and 0 otherwise. The resulting IPC-specific measure is the ratio of patents classified in the IPC for which the maximum patent length was binding.

D.5 Unique number of inventors - Technical field-level

PATSTAT table `tls207` associates to each application a list of personal id's that correspond to the inventors and to the applicants listed on the patent. Table `tls206`, instead, reports, for each of these personal id's, details such as the full name listed on the patent, the address of the inventor or the applicant, and other information. Since these personal id's assigned by PATSTAT do not uniquely identify a person or a firm, a substantial harmonization effort has been done by the EPO, the OECD, and other researchers. Among the harmonized id's available in table `tls206`, I chose the STAN harmonized applicant's identifiers developed starting from the EPO Worldwide Bibliographic Database. Hence, combining `tls207` and `tls206` with patent application information as selected in subsection D.1, I can assign to each patent the unique (up to harmonization errors) identifiers of the inventors listed on the patent. In order to build the quarterly measure of unique inventors working in a given IPC, I simply count the number of id's that are associated to a patent classified in the IPC and with priority date in the quarter, dropping from this count multiple records of the same inventor in multiple patents in the same IPC-quarter.

D.6 Number of new applicants, patents granted to new applicants and their share on the total - Technical field-level

In order to compute the number of new applicants and measures associated to this concept, I follow a similar approach as the one just described in subsection D.5, and I attach to each patent application selected according to the rules of subsection D.1 the harmonized identifiers of the applicants associated to the patent according to table `tls207`. To determine whether an applicant is a new or an incumbent one, for each quarter and IPC I build a list of applicant's ids that have already appeared at least once in the specific IPC and, for each applicant, I check whether the id belongs to this list or not. If the id does not belong to the list, the applicant is assigned a flag of 1 as a new entrant for that IPC-quarter pair. The unique number of new applicants is computed by counting the unique number of ids for which the flag is 1 by IPC and quarterly priority date of the application. Similarly, the number of granted patents attributable to new applicants is computed by assigning a value of 1 to a dummy variable in case at least one of the applicants is an entrant, and 0 otherwise. Then, the number of patents with such dummy equal to 1 is counter by IPC and quarter. Finally, the share of patents attributable to new applicants is simply computed by dividing the absolute figure just described by the total number of patents filed in the corresponding IPC-quarter.

D.7 Herfindahl-Hirschman Index of concentration - Technical field-level

In order to compute the HHI by technological field, I follow a similar approach as in subsection D.5, and I attach to each patent application, selected according to the rules of subsection

D.1, the harmonized identifiers of the applicants associated to the patent, taken from table `tls207`. Then, I compute the total number of patents (or citations-weighted patents) made by a specific applicant in a given technical field and quarter, and the total number of patents (or citations-weighted patents) generated in the same technical field and quarter by any applicant. Let $s_{i,j,t}$ be the share of patents made by applicant i over the total number of patents in field j and quarter t. Then, the concentration index is

$$HHI_{j,t} = \sum_i s_{i,j,t}^2 100^2$$

D.8 Within-field backward citations and within-field backward citation intensity - Technical field-level

To compute the number of backward citations by IPC and quarter, I start from the pool of patents selected according to the criteria of subsection D.1, and I follow subsection D.2 to attach, to each patent application, the associated publications and the citations information of table `tls212`. However, in this case, rather than keeping the list of citing publications, I keep the list of documents that each application (or publications associated to each application) cite. Also, I separately keep track of citations directly made by the applicant (`citn_origin`='APP') rather than added by examiners or during search. This distinction may be important because previous literature has pointed out that only backward citations made by applicants are representative of genuine knowledge flows. Therefore, for each patent, I compute the overall number of patent documents backward cited (overall and by applicants only), *and* the number of backward-cited patents (overall and by applicants only) that are classified in the same IPC of the patent considered. I aggregate both variables by IPC and quarterly priority date of the citing patent application. The aggregation of the latter variable is called in the main text within-IPC backward citations. The intensity measure is computed as the fraction of patents where the applicant makes at least one backward citation directed to another patent in the same field over the number of patents having at least one backward citation.

To compute the number and the intensity of within-field backward citations made by patents filed during the post-implementation period July 1995 - July 2000 and directed to patents filed during the pre-implementation period November 1992 - June 1995, I repeat the same steps described above, but I restrict my attention to patents satisfying previous timing criteria. Obviously, the steps are the same for the control group of patents filed during July 1985 - July 1990 and backward citing other patents classified in the same technical field and filed during the period November 1982 - June 1985.

D.9 Private economic value of innovation - Technical field-level

To compute a measure of economic value of patents by technical field and quarter, I start from the data provided in the replication package of Kogan et al. (2017). The variables relevant for the present analysis are: *i*) the 7-digit US patent number, *ii*) the private economic value of a patent ξ, and *iii*) the application date of the patent. Using the 7-digit US patent number, I merge the dataset with the NBER patent database and, specifically, with the dataset which contains information on International Patent Classification classes assigned to the patent. Then, using the original application filing date and the IPC classes from the NBER database, I add up the economic value of patents by quarterly application date and technical field.

D.10 Number of patents and citations-weighted patents - Firm-level

For the firm-level dataset, I rely on the NBER Patent Database merged with COMPUSTAT using the applicant-`gvey` cross-walk provided in the NBER Database itself. In particular, the NBER Database provides a list of `gvey` identifiers associated to each patent over its life. Multiple `gvey`'s over time indicate that the ownership of the patent has changed. However, here I am just interested in the firm which has originated the invention through its R&D effort, and this is why I just keep the `gvey` associated to the patent in the year of application. Then, I download from COMPUSTAT firm-level information, and I match this dataset with the `gvey`'s retrieved from the NBER Database. I build the patent count and the citations-weighted patent variable by summing for each `gvey` and year the number of patents applied for by the firm and the truncation-adjusted citations variable available in the database, respectively. Truncation adjustment is performed in the original dataset by applying to citations the weights proposed by Hall, Jaffe and Trajtenberg (2001).

D.11 R&D expenditure - Firm-level

Data on firm-level R&D expenditure is downloaded from COMPUSTAT and merged with information on patents using the `gvey` link. The name of the original R&D expenditure variable in COMPUSTAT is `xrd`.

D.12 Treatment - Firm-level

In order to build the treatment at the firm level, I start from the patent-level dataset of the NBER Database, which reports also information on the 4-digit IPC classes associated to the patents. Then, for each `gvey` identifier, I just keep those patents with application year between 1971 and 1991, in order to have enough patenting-related information for each firm and to exclude possible effects of the policy news. Then, for each firm, I compute the share of the total number of patents, filed during this period, that is classified in each of the 4-digit IPC. Let's call it $s_{i,j}$, where i indexes firms and j IPCs. I interpret this fraction as the exposure of firm i to the technical field j before the policy news. The firm-level treatment is then built as a weighted average of the field-specific treatment described in subsection D.3, i.e.

$$T_i = \sum_j s_{i,j} T_j$$

I take this approach because the main source of $ex-ante$ heterogeneity in pending period is linked to the different technical fields and, relatedly, to the different technical offices and examination difficulties. Therefore, I still want to use field-level heterogeneity, interacting it with heterogeneity in the technological location of firms. An alternative would be to compute the firm-level treatment by computing the average pending period of patents filed by firm i, i.e. a pending period based on the specific experience of the firm. I do not follow this route because I think this treatment variable would be more prone to endogeneity concerns that the one I propose: In this case, the treatment might be correlated with the quality of innovation performed by the firm, or with the responsiveness of the firm to the inquiries of the patent office.

D.13 Other COMPUSTAT variables - Firm-level

I compute firm age using the `begyr` NBER patent database variable, and I use 2-digit SIC industry code assigned to each firm in the database. Firm's yearly sales are taken from COMPUSTAT using the variable `sales`.

D.14 Aggregate investment externality - Firm-level

To compute the firm-level externality measure used in Section 4 of the paper, I try to follow what has been done in the literature on the topic. Therefore I compute, for the period 1971-1991, the total number of patents obtained by each of the firms in my sample in any 4-digit IPC. This information is included in firm-specific vector f_i, which stacks, in each entry, the number of patents obtained by firm i in IPC j in the above-mentioned period. Then, based on these vectors, I compute for every pair of firms (i,k) a technological distance measure proposed by Jaffe (1986)

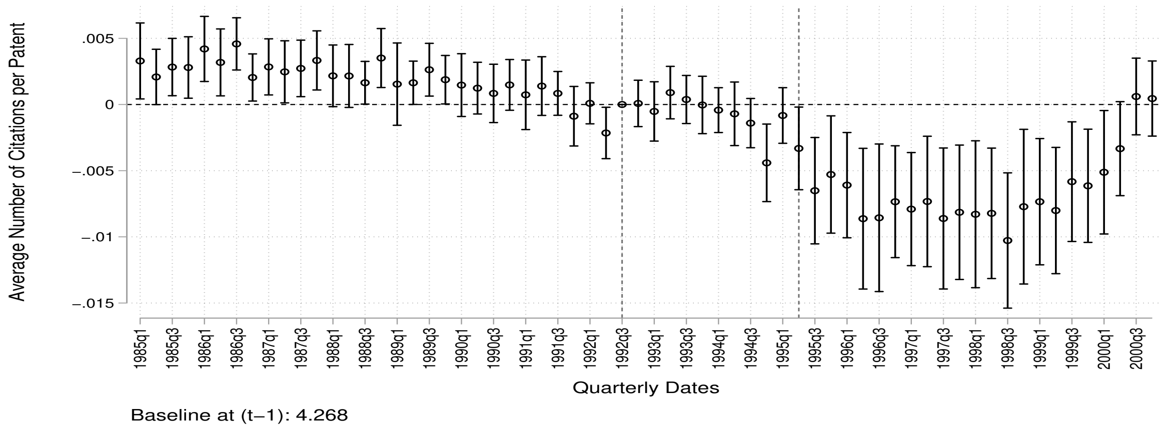
$$d_{i,k} = \frac{f_i f_k'}{\sqrt{(f_i f_i')(f_k f_k')}}$$

The externality measure for firm i at time t is then

$$E_{i,t} = \sum_{k \neq i} d_{i,k} R\&D_{k,t}$$

i.e. it is an aggregation of the firm-level R&D expenditure of other firms that uses as aggregation weights the Jaffe (1986) measure of technological distance across firms. The idea

Figure E.1: **Marginal effect of 1 more day of protection on average citations per patent**



The plot shows the β_k coefficients of the specification (27) having as dependent variable the average number of forward citations obtained by patents filed in quarter-t and field-j. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

underlying such externality variable is that the influence of other firms' R&D is stronger if such firms are technologically closer to the firm of interest.

D.15 Productivity and welfare - Sectoral level

The productivity and price variables used in the sectoral welfare analysis are directly taken from the NBER CES manufacturing database. Productivity is measured as 5-factors TFP, whose constructions is detailed in the technical paper Bartelsman and Gray (1996). Welfare is (inversely) measured by the value of shipments price deflator, which is built aggregating product-specific deflators computed by the Bureau of Economic Analysis.

Appendix E Additional empirical results

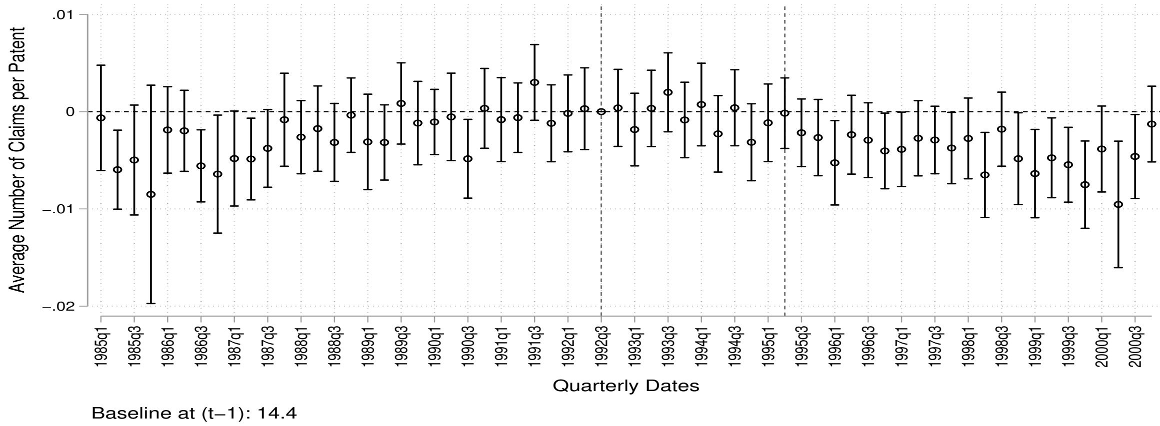
E.1 Results at the technical field level

E.1.1 Average number of citations per patent

Figure E.1 plots the β_k coefficients of regression (1) having as dependent variable the *average* number of forward citations per patent, for patents filed in quarter-t and classified in field-j.[104] Fields with zero patents in at least one quarter are excluded from the estimation sample because the average number of citations is not well-defined in such cases. Results are analogous, however, when just excluding the field-quarter observation which is not well-defined.

[104] I count citations obtained within 5 years from application, to avoid truncation bias.

Figure E.2: **Marginal effect of 1 more day of protection on average number of claims per patent**



The plot shows the β_k coefficients of the specification (27) having as dependent variable the average number of claims made by patents filed in quarter-t and field-j. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

E.1.2 Average number of claims per patent

Figure E.2 plots the β_k coefficients of regression (1) having as dependent variable the *average* number of claims per patent, for patents filed in quarter-t and classified in field-j. Fields with zero patents in at least one quarter are excluded from the estimation sample because the average number of claims is not well-defined in such cases. Results are analogous, however, when just excluding the field-quarter observation which is not well-defined.

E.1.3 Average originality and average generality

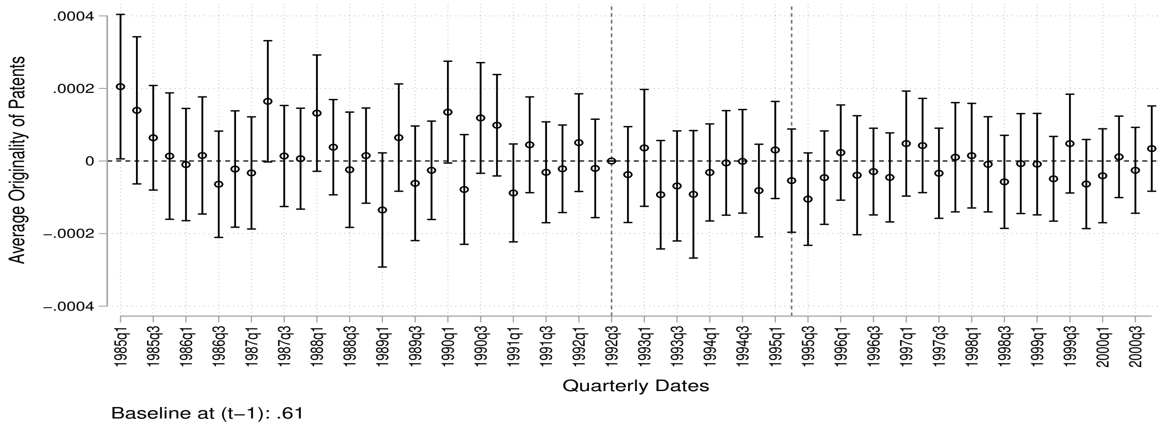
Figure E.3 plots the β_k coefficients of regression (1) having as dependent variable the *average* originality of patents filed in quarter-t and classified in field-j. The originality index of each patent i is taken from the NBER patent database and it is computed as

$$O_i = 1 - \sum_{j=1}^{n} s_{i,j}^2$$

where $s_{i,j}$ denotes the percentage of citations made by patent i that belong to patent class j, out of n patent classes. The results show that the policy does not affect the average originality of patents, which is often taken as a proxy of patent quality.

Figure E.4 plots the β_k coefficients of regression (1) having as dependent variable the *average* generality of patents filed in quarter-t and classified in field-j. The generality index

Figure E.3: **Marginal effect of 1 more day of protection on average originality of patents**



The plot shows the β_k coefficients of the specification (1) having as dependent variable the average originality of patents filed in quarter-t and field-j. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

of each patent i is taken from the NBER patent database and it is computed as

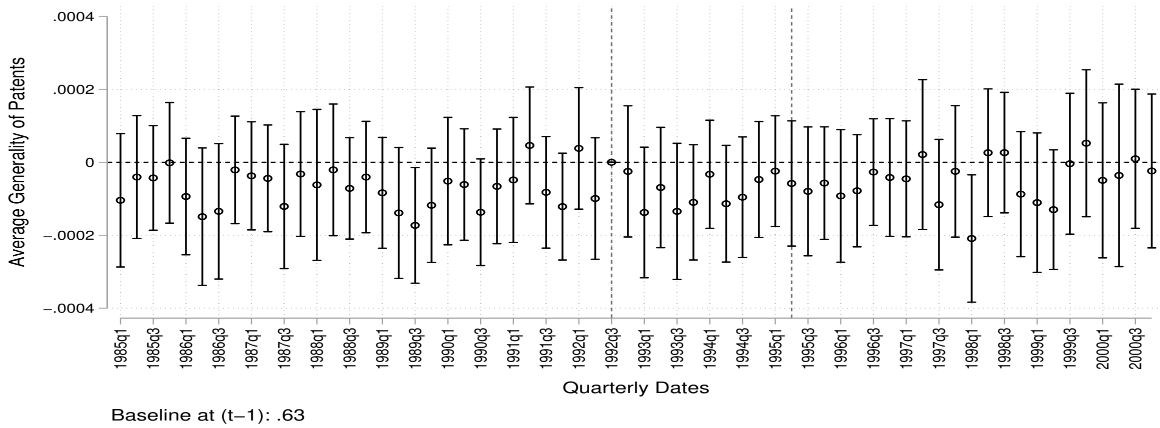
$$G_i = 1 - \sum_{j=1}^{n} s_{i,j}^2$$

where $s_{i,j}$ denotes the percentage of citations received by patent i that belong to patent class j, out of n patent classes. The results show that the policy does not affect the average generality of patents, which is often taken as a proxy of patent quality.

E.1.4 Maintenance fee payment probability

In order to keep patent protection active, patent owners must pay fees after 3.5 years, 7.5 years, and 11.5 years from the grant. The payment of renewal fees is commonly linked to the quality of patents–i.e., higher quality patents are renewed for longer–and to the rate of creative destruction. If a technology is competed away by a new invention, it is pointless to pay fees to keep alive the patent on an old technology that will not generate profits. In this subsection I examine whether a patent term extension has any effects on the average renewal rate of patents at later stages of their maintenance. Figure E.5 plots the β_k coefficients of regression (1) having as dependent variable the share of patents filed in quarter-t and classified in field-j that are renewed up to the maximum patent length. The results show that a patent term extension does not induce innovators to renew their patents for longer. Since other analyses showed that the average quality of patents was not changing due to the policy, I interpret this finding as suggestive of the fact that the pressure of creative destruction does not fall in fields getting a patent term extension. Figure E.6 shows that results are analogous when using as the outcome variable the share of patents filed in quarter-t and classified in

Figure E.4: **Marginal effect of 1 more day of protection on average generality of patents**



The plot shows the β_k coefficients of the specification (1) having as dependent variable the average generality of patents filed in quarter-t and field-j. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

field-j that are renewed until 11.5 years since the grant.

E.1.5 Alternative specifications

Citations and inventors with log. transformation Figures E.7 and E.8 report the results of specification (27) from Appendix subsection B.2.13–dependent variable taken in natural logarithms– using citations-weighted patents as the dependent variable. Figures E.9 and E.10 report the same coefficients using the number of unique inventors by quarterly and field as the dependent variable.

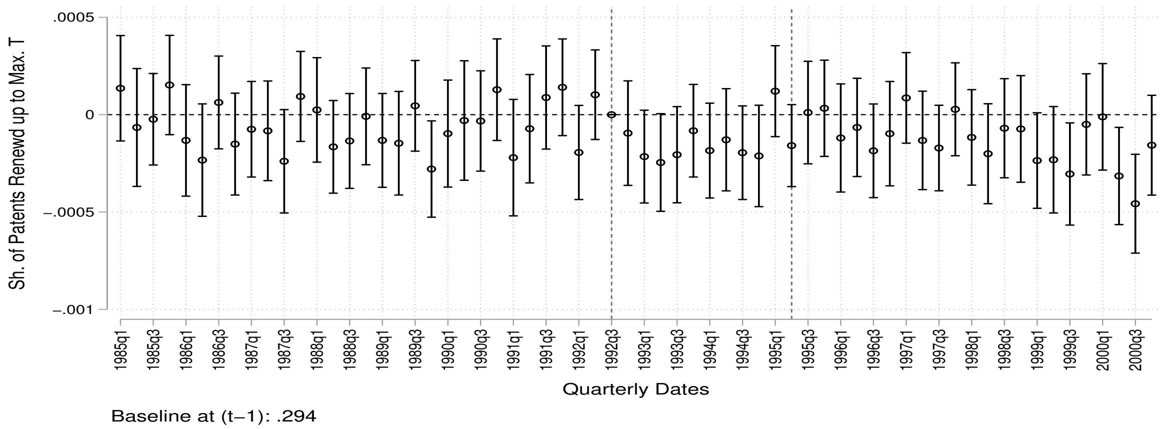
Citations and inventors with negative binomial model Figures E.11 and E.12 plot the β_k coefficients of the negative binomial model (28) having as dependent variable the count of quarter-t and field-j citations-weighted patents. Figures E.13 and E.14 plot the same coefficients for the number of unique inventors by quarter and field as the dependent variable.

Inverse sine transformation A further alternative transformation of the dependent variable is the inverse sine transformation, which is increasingly used as an alternative to the natural logs transformation because, differently from the natural logarithm, it is well-defined at 0. The specification is

$$ln(P_{j,t} + \sqrt{P_{j,t}^2 + 1}) = \alpha_j + \sum_{k=1985Q1}^{2000Q4} \gamma_k \mathbf{1}_{(t=k)} + \sum_{k=1985Q1}^{2000Q4} \beta_k \mathbf{1}_{(t=k)} T_j + \varepsilon_{j,t} \quad (63)$$

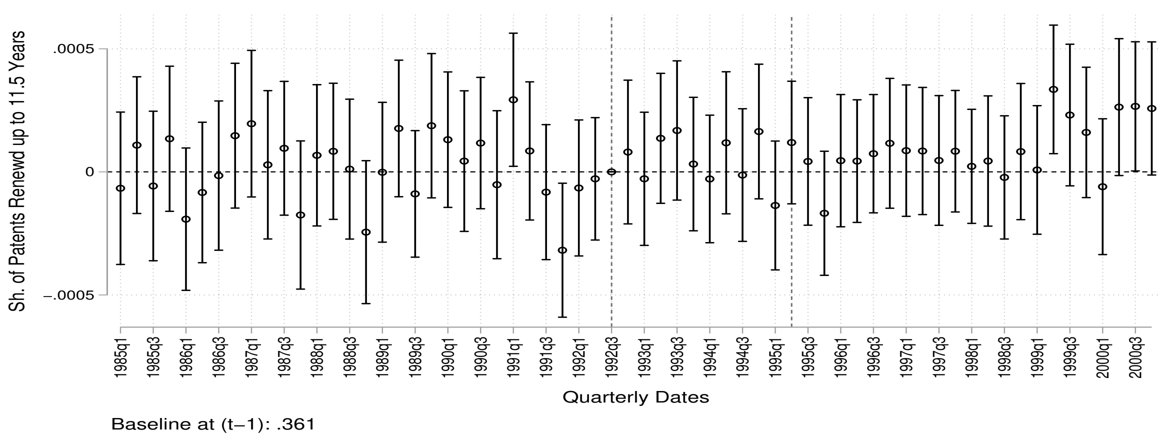
where all the variables have the same meaning as in subsection 3.1. Figure E.15 plots the β_k coefficients of specification (63) when the regression is run considering all technical

Figure E.5: **Marginal effect of 1 more day of protection on average renewal rate of patents up to the maximum term**



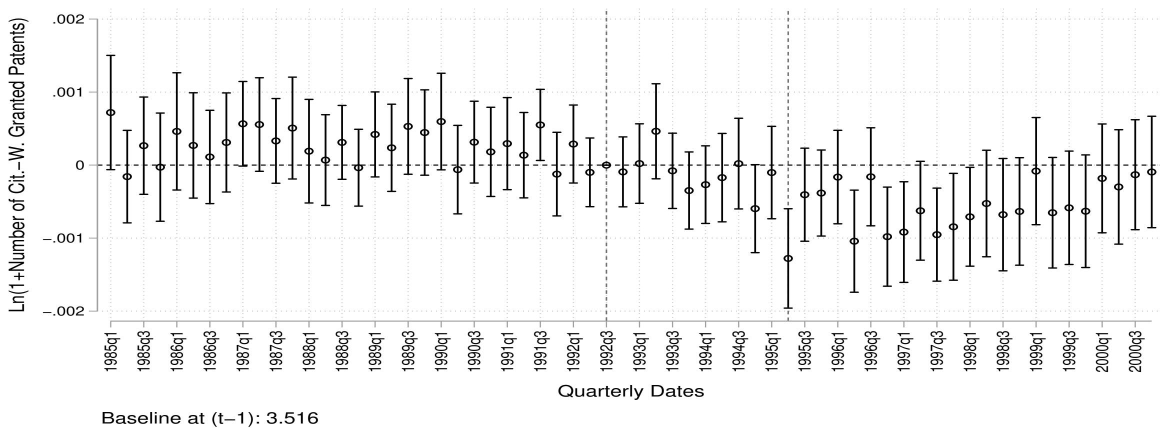
The plot shows the β_k coefficients of the specification (1) having as dependent variable the share of patents filed in quarter-t and classified in field-j that are renewed up to the maximum patent length. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure E.6: **Marginal effect of 1 more day of protection on average renewal rate of patents up to 11.5 years since the grant**



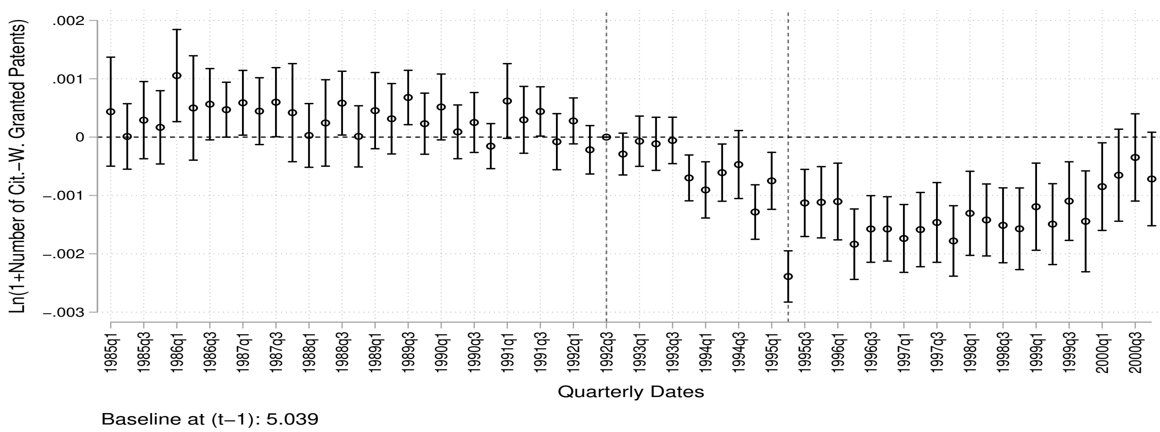
The plot shows the β_k coefficients of the specification (1) having as dependent variable the share of patents filed in quarter-t and classified in field-j that are renewed up to 11.5 years since the grant. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure E.7: **Marginal effect of 1 more day of protection on citations-weighted patents**



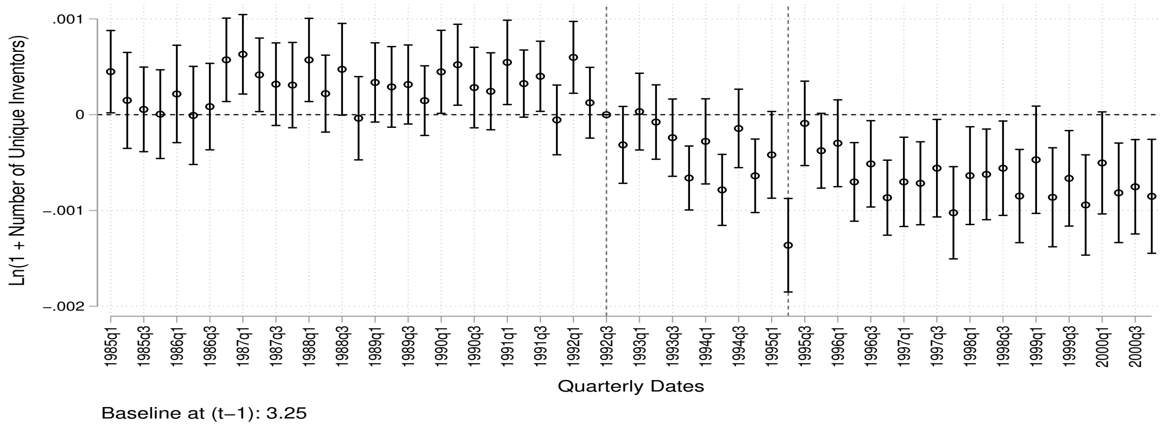
The plot shows the β_k coefficients of the specification (27) having as dependent variable the log. of one plus quarter-t and field-j number of citations-weighted granted patents. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure E.8: **Marginal effect of 1 more day of protection on citations-weighted patents**



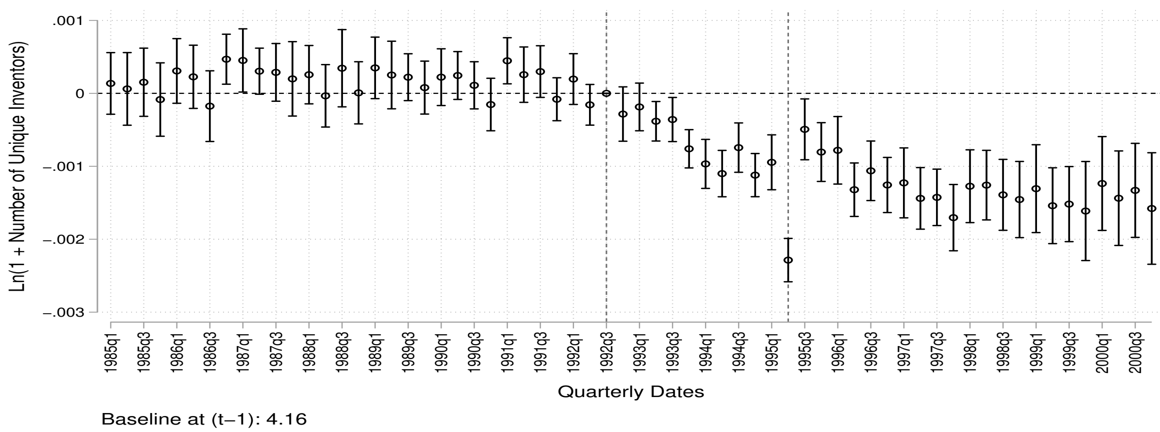
The plot shows the β_k coefficients of the specification (27) having as dependent variable the log. of one plus quarter-t and field-j number of citations-weighted granted patents. The sample covers only technical fields with a number of total patents above the sample median. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure E.9: **Marginal effect of 1 more day of protection on unique inventors**



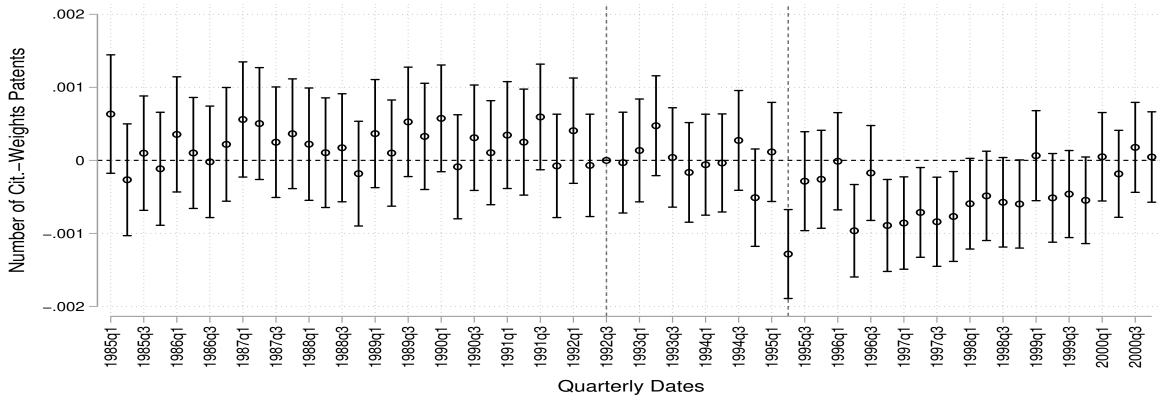
The plot shows the β_k coefficients of the specification (27) having as dependent variable the log. of one plus quarter-t and field-j number of unique inventors. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure E.10: **Marginal effect of 1 more day of protection on unique inventors**



The plot shows the β_k coefficients of the specification (27) having as dependent variable the log. of one plus quarter-t and field-j number of unique inventors. The sample covers only technical fields with a number of total patents above the sample median. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

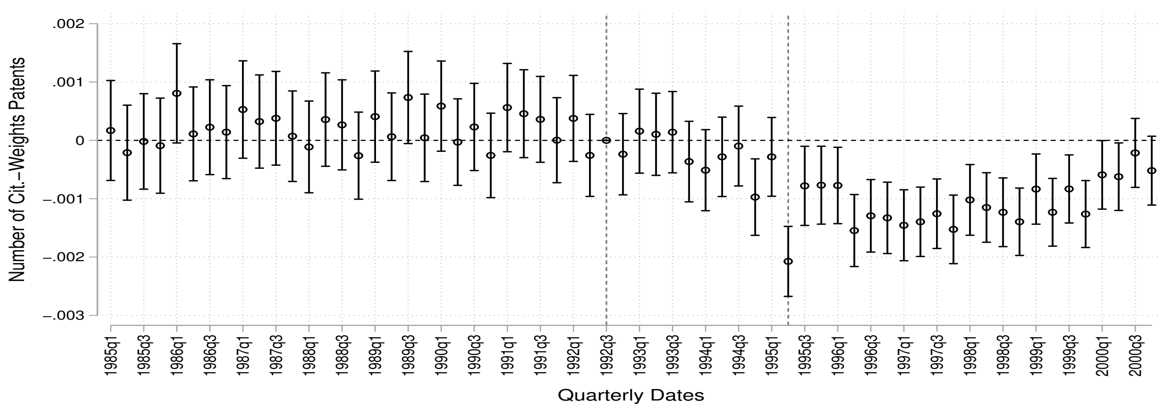
Figure E.11: **Estimated treatment coefficients of a negative binomial model - Number of citations-weighted patents as outcome**



Baseline at (t–1): 158.647

The plot shows the β_k coefficients of the specification (28) having as dependent variable quarter-t and field-j number of citations-weighted granted patents. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

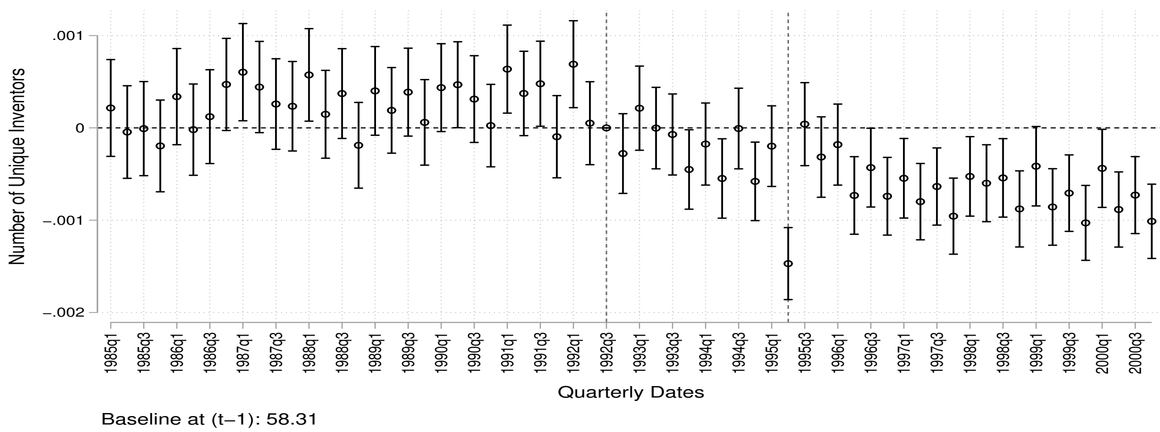
Figure E.12: **Estimated treatment coefficients of a negative binomial model - Number of citations-weighted patents as outcome**



Baseline at (t–1): 292.56

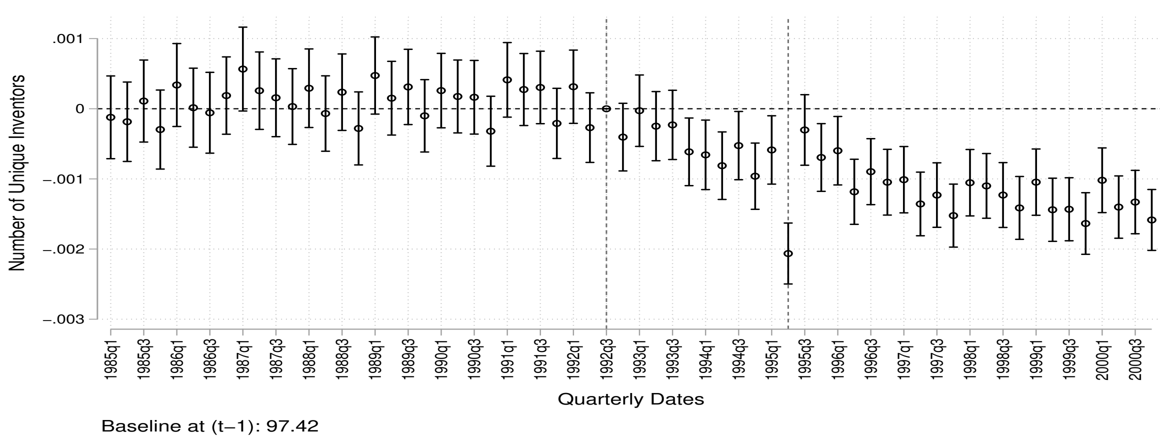
The plot shows the β_k coefficients of the specification (28) having as dependent variable quarter-t and field-j number of citations-weighted granted patents. The sample covers only technical fields with a number of total patents above the sample median. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure E.13: **Estimated treatment coefficients of a negative binomial model - Number of unique inventors as outcome**



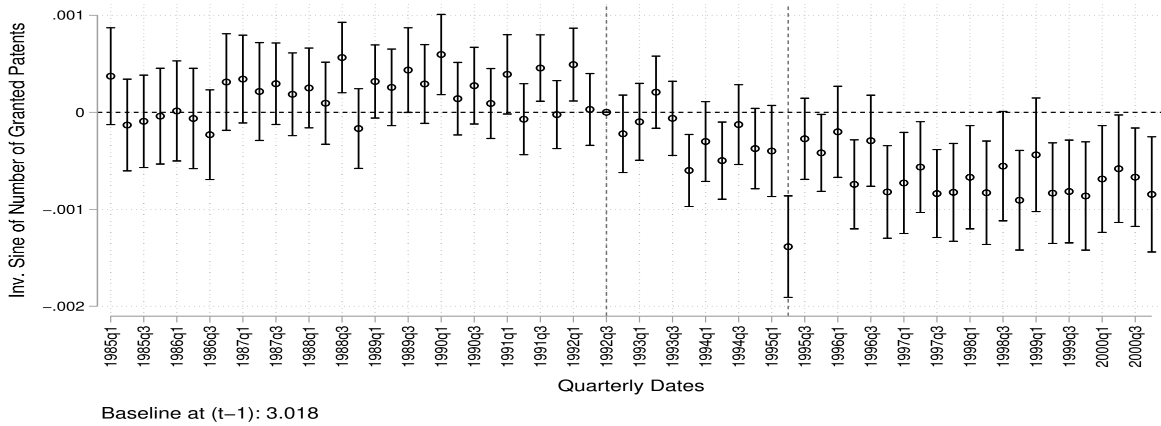
The plot shows the β_k coefficients of the specification (28) having as dependent variable quarter-t and field-j number of unique inventors. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure E.14: **Estimated treatment coefficients of a negative binomial model - Number of unique inventors as outcome**



The plot shows the β_k coefficients of the specification (28) having as dependent variable quarter-t and field-j number of unique inventors. The sample covers only technical fields with a number of total patents above the sample median. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure E.15: **Marginal effect of 1 more day of protection on granted patents - Inverse sine transformation**



The plot shows the β_k coefficients of the specification (63) having as dependent variable the inverse sine transformation of quarter-t and field-j number of granted patents. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

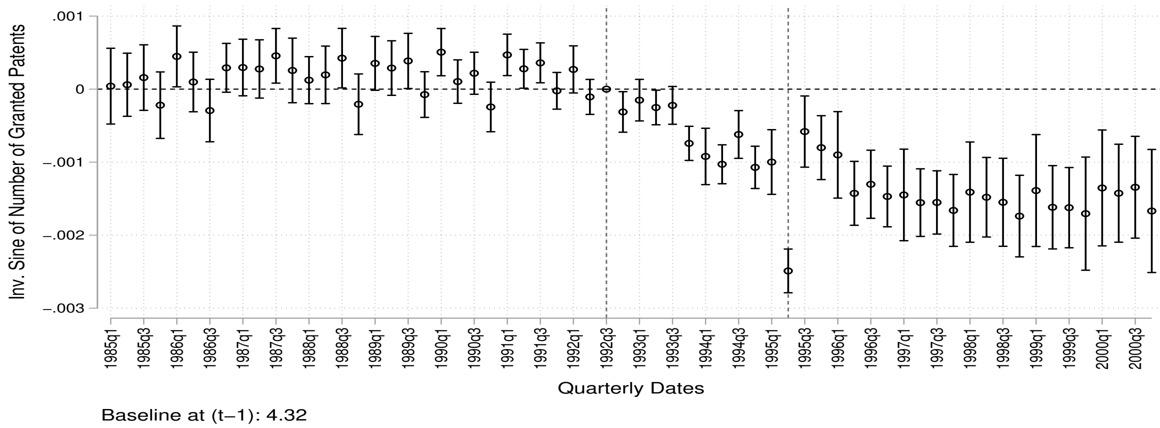
fields. However, consistently with Appendix subsection B.2.13, I also run the same regression excluding from the sample the fields with a total number of patents in the sample below the median. The remaining half of the fields generates more than 90% of total patents in the sample. Figure E.16 plots the β_k coefficients of specification (63) on such restricted sample, showing that results of section 3 of the paper are largely confirmed. Also, the results are fully analogous when using citations-weighted patents or the unique number of inventors as outcomes of interest.

Percentage deviations from 1985Q1 patents As a final transformation, I run specification (1) using as dependent variable the percentage deviation of patents in quarter-t and field-j from the number of patents filed in the same field in 1985Q1. Therefore, the dependent variable is $P^d_{j,t} = \frac{P_{j,t}}{P_{j,1985Q1}}$. Figure E.17 plots the results, which are consistent with the main evidence. The results for citations-weighted patents and the number of inventors–not reported–are fully consistent.

E.1.6 Alternative sample restrictions

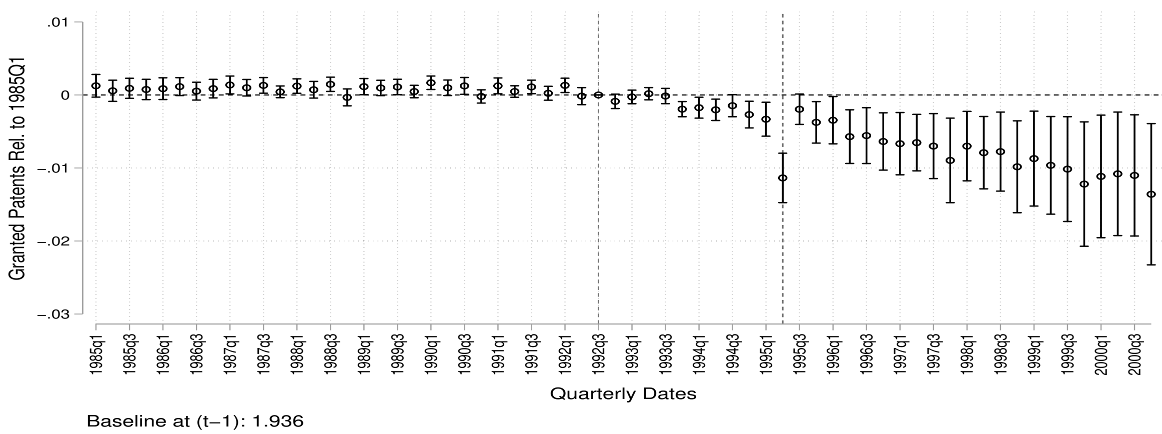
In order to reduce the skewedness of the distribution of patenting outcomes across fields, an alternative to performing a transformation of the dependent variable–such as taking natural logs or the inverse sine of quarterly granted patents or citations weighted patents–is to drop technical fields which are either very small or very big compared to the average. In this subsection, I show the β_k coefficients of the linear-in-levels specification (1) used in the main analysis when the sample is restricted to technical fields that, in all quarters, have not less

Figure E.16: **Marginal effect of 1 more day of protection on granted patents - Inverse sine transformation**



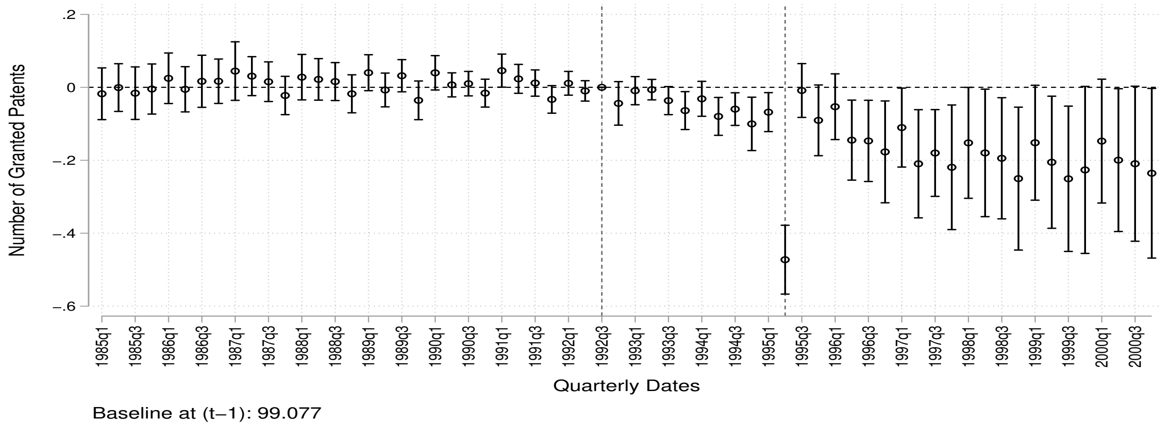
The plot shows the β_k coefficients of the specification (63) having as dependent variable the inverse sine transformation of quarter-t and field-j number of granted patents. The sample covers only technical fields with a number of total patents above the sample median. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure E.17: **Marginal effect of 1 more day of protection on granted patents**



The plot shows the β_k coefficients of the specification (1) having as dependent variable $P^d_{j,t} = \frac{P_{j,t}}{P_{j,1985Q1}}$, i.e. the percentage deviation of patents in quarter-t and field-j from the number of patents filed in the same field in 1985Q1. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

Figure E.18: **Marginal effect of 1 more day of protection on granted patents**



The plot shows the β_k coefficients of the specification (1) having as dependent variable granted patents in quarter-t and field-j. The sample is restricted to technical fields that, in all quarters, have not less than 25 patents and not more than 500. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

than 25 patents and not more than 500.[105] Figures E.18 and E.19 show the results obtained on such restricted sample, when using granted patents and citations-weighted granted patents as outcome variables, respectively.

E.2 Results at the sector-level

E.2.1 Evidence on innovation outcomes

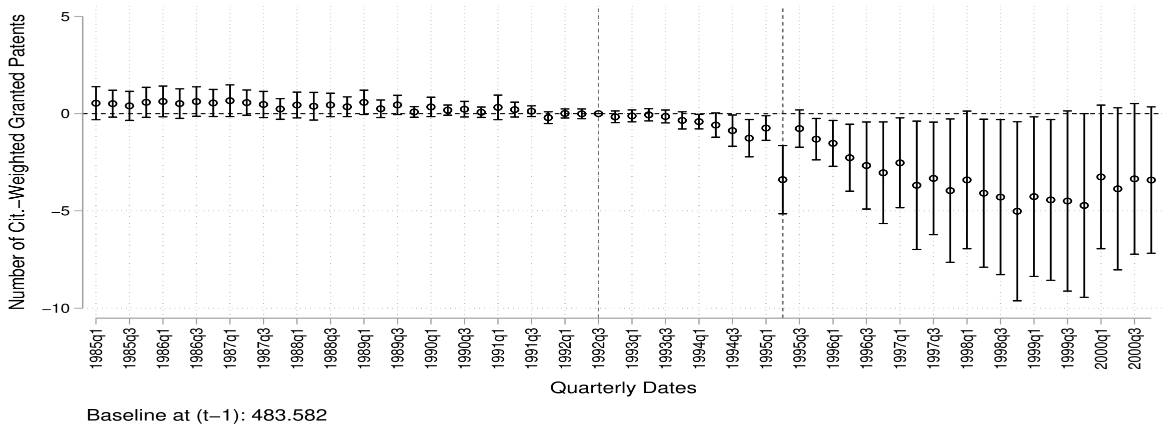
In this subsection I provide evidence on the effect of the policy on innovation outcomes, as measured by patents, citations-weighted patents, and patent value, at the NAICS 6-digit industry level.[106] This layer of analysis requires aggregation of innovation measures by industry, and the adaptation of the technical-field level treatment variable, i.e. the policy-induced change in patent protection time, at the industry-level.

To build measures of innovation by 6-digit NAICS and year, I start from measures of innovation, i.e. number of granted patents, number of citations-weighted patents, and private economic value of patents, by technical field and quarter. The first step is to aggregate previous innovation measures at the yearly level. The second step involves mapping them into 6-digit NAICS. This is done through the following formula

[105] For sake of clarity, these figures refer to the number of applications that are subsequently granted. As in the other parts of the paper, the count of patents is done based on the quarter when the applications is filed, irrespective of the subsequent grant quarter.

[106] An example of the depth of the sectoral classification I use in the analysis is the following. *31-33* is the aggregate 2-digit classification for *Manufacturing*; *324* is the 3-digit *Petroleum and Coal Products Manufacturing*, *3241* is the 4-digit *Petroleum and Coal Products Manufacturing*; which includes the 5-digit *32412 Asphalt Paving, Roofing, and Saturated Materials Manufacturing*, which includes the 6-digit sectors 324121 *Asphalt Paving Mixture and Block Manufacturing* and 324122 *Asphalt Shingle and Coating Materials Manufacturing*.

Figure E.19: **Marginal effect of 1 more day of protection on citations-weighted patents**



The plot shows the β_k coefficients of the specification (1) having as dependent variable citations-weighted granted patents in quarter-t and field-j. The sample is restricted to technical fields that, in all quarters, have not less than 25 patents and not more than 500. Standard errors are clustered by technical field. 95% confidence bands are plotted. The first vertical line refers to the quarter before the policy news (1992Q3) and the second vertical line refers to the quarter before the policy implementation (1995Q2).

$$I_{s,t} = \sum_j I_{j,t}\pi_{s|j}$$

$I_{s,t}$ is innovation in 6-digit NAICS sector s and year t, $I_{j,t}$ is innovation in 4-digit IPC field j and year t, and $\pi_{s|j}$ is the probability that a patent classified in technical field j is linked to sector s or, alternatively, contributes to innovation in sector s. $\pi_{s|j}$ is directly taken from the 'Algorithmic Links with Probabilities' crosswalk by Goldschlag, Lybbert and Zolas (2019), which exactly compute these conditional probabilistic links between sectors and technical field based on text analysis.

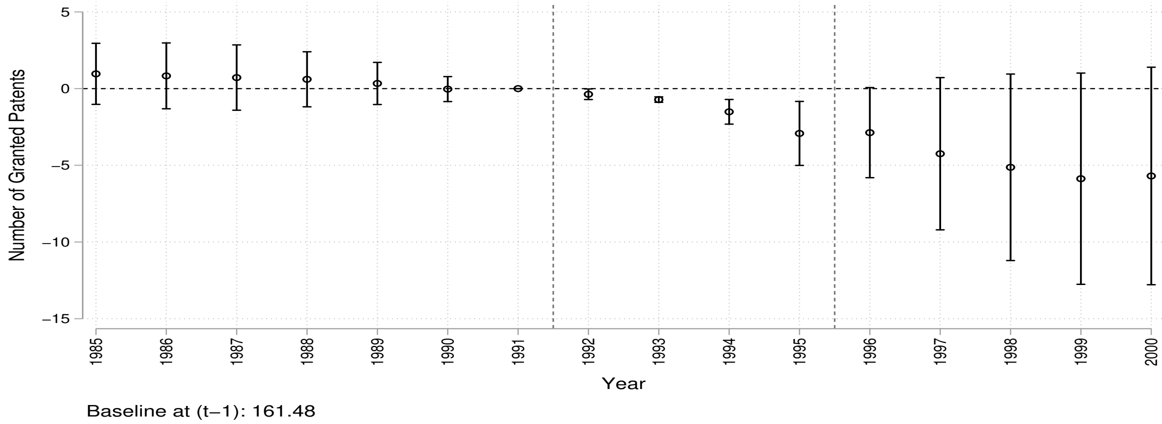
To convert the technical field-level treatment into a 6-digit NAICS sectoral treatment, I rely again on probabilistic links between 4-digit IPC classes and 6-digit NAICS industries computed by Goldschlag, Lybbert and Zolas (2019). Specifically,

$$T_s = \sum_j T_j\pi_{j|s}$$

The treatment T_s for sector s is the sum of technical field-level treatments T_j's, weighted by the probability that, given that a patent is assigned NAICS s, it comes from technical field j.

The specification of the difference-in-difference regression at the industry-level is analogous to the one by technical field

Figure E.20: **Marginal effect of 1 more day of protection on sectoral granted patents**



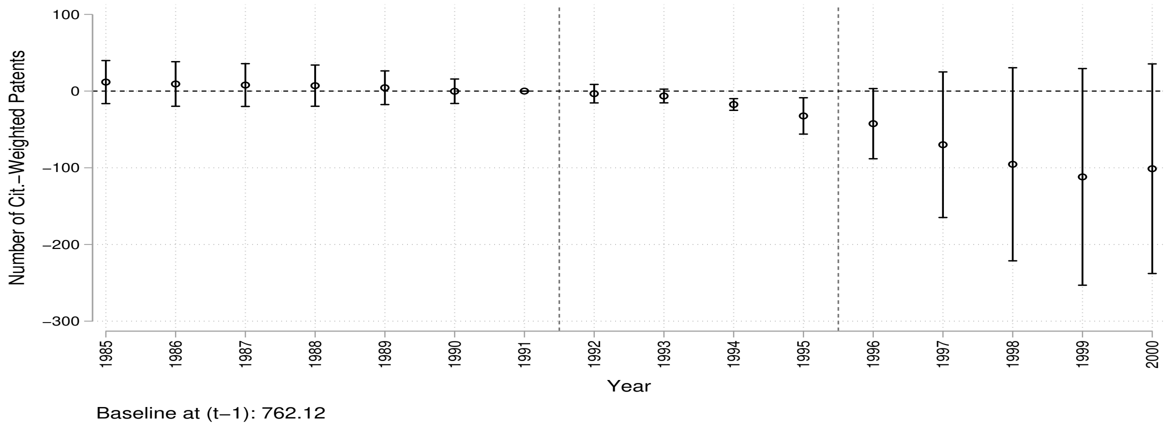
The plot shows the β_k coefficients of regression (64) $P_{s,t} = \alpha_s + \sum_{k=1985}^{2000} \gamma_k \mathbf{1}_{(t=k)} + \sum_{k=1985}^{2000} \beta_k \mathbf{1}_{(t=k)} T_s + \varepsilon_{s,t}$. $P_{s,t}$ is the number of patents attributable to sector s and filed for in year t, and T_s is the industry-specific treatment. I omit the dummy for 1991, which is the pre-treatment year. 95% confidence bands are plotted. Standard errors are clustered by 3-digit NAICS industry and year. The first vertical line lies just before the news year (1992) and the second vertical line lies just before the implementation year (1995).

$$Y_{s,t} = \alpha_s + \sum_{k=1985}^{2000} \gamma_k \mathbf{1}_{(t=k)} + \sum_{k=1985}^{2000} \beta_k \mathbf{1}_{(t=k)} T_s + \Xi X_{s,t} + \varepsilon_{s,t} \tag{64}$$

where α_s are industry fixed effects, $\mathbf{1}_{(t=k)}$ are yearly dummy variables, T_s is the sectoral treatment, $X_{s,t}$ is matrix of controls that includes 4-digit NAICS industry $\times$ year effects, the natural logarithm of the energy price deflator, and the natural logarithm of the material costs deflator, $\varepsilon_{s,t}$ is an idiosyncratic error term. Standard errors are clustered by 3-digit NAICS industry$\times$ year in this case and the regressions.

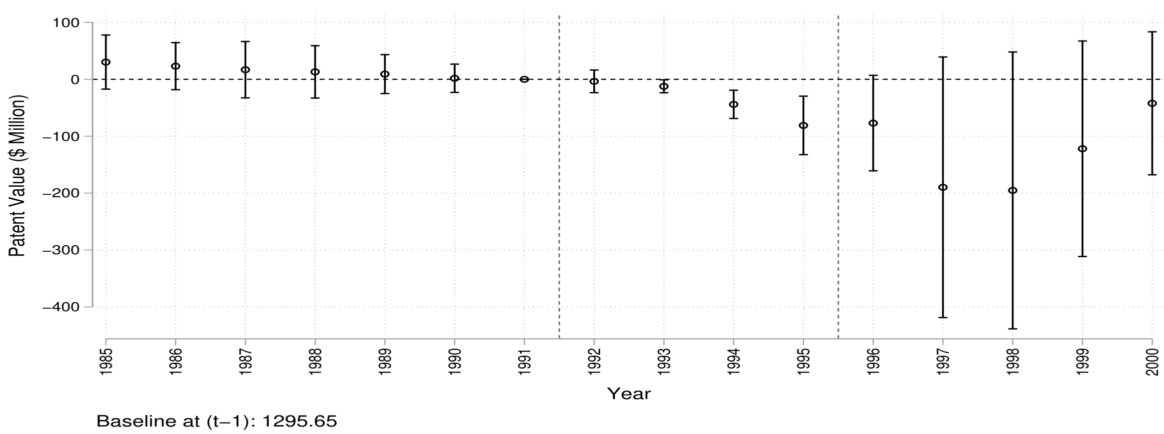
Figure E.20 plots the difference-in-difference β_k coefficients of interest, together with their 95% confidence bands, for specification (64) run having the number of granted patents by industry and year as dependent variable. The pre-treatment coefficients are remarkably close to 0, confirming the absence of pre-trends, and the pattern of post-treatment estimated marginal effects is similar to the evidence by technical field presented in Section 3 of the paper. Figures E.21 and E.22 plot the same coefficients for citations-weighted patents and patent value as dependent variables , respectively. Again, the evidence is very consistent with previous one, even though confidence bands are larger in this case.

Figure E.21: **Marginal effect of 1 more day of protection on sectoral citations-weighted granted patents**



The plot shows the β_k coefficients of regression (64) $C_{s,t} = \alpha_s + \sum_{k=1985}^{2000} \gamma_k \mathbf{1}_{(t=k)} + \sum_{k=1985}^{2000} \beta_k \mathbf{1}_{(t=k)} T_s + \varepsilon_{s,t}$. $C_{s,t}$ is the number of citations-weighted patents attributable to sector s and filed for in year t, and T_s is the industry-specific treatment. I omit the dummy for 1991, which is the pre-treatment year. 95% confidence bands are plotted. Standard errors are clustered by 3-digit NAICS industry and year. The first vertical line lies just before the news year (1992) and the second vertical line lies just before the implementation year (1995).

Figure E.22: **Marginal effect of 1 more day of protection on sectoral patent value**



The plot shows the β_k coefficients of regression (64) $V_{s,t} = \alpha_s + \sum_{k=1985}^{2000} \gamma_k \mathbf{1}_{(t=k)} + \sum_{k=1985}^{2000} \beta_k \mathbf{1}_{(t=k)} T_s + \varepsilon_{s,t}$. $V_{s,t}$ is the private economic value of patents attributable to sector s and filed for in year t, and T_s is the industry-specific treatment. I omit the dummy for 1991, which is the pre-treatment year. 95% confidence bands are plotted. Standard errors are clustered by 3-digit NAICS industry and year. The first vertical line lies just before the news year (1992) and the second vertical line lies just before the implementation year (1995).

Appendix F Additional theoretical results

F.1 A model with R&D in one stage: Jones (1995)

In this subsection, I sketch the key elements of an adaptation of Jones (1995)'s model of semi-endogenous growth to a setting with finite patent length, which I use the show that, without the distinction of research and development activity in the way they are modelled in Section 5 of the paper, the theory cannot match the empirical responses. The consumption side and the production side are exactly equal to what presented in subsection 5.2.2 of the paper. What is different is the R&D block of the model. In the standard setup, R&D is done in one step, and investment of $I_R(t)$ units of the final good generates $V(t)^{\phi_1} I_R(t)^{\phi_2}$ new intermediate good varieties. So, the R&D maximization problem is simply

$$\max_{I_R(t)} \left\{ \upsilon(t) V(t)^{\phi_1} I_R(t)^{\phi_2} - I_R(t) \right\}$$

and, given the optimal investment $I_R^*(t)$ solving previous problem, the evolution of varieties over time is ruled by

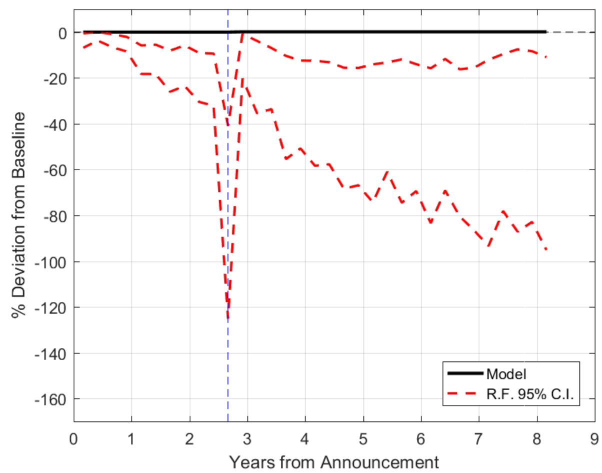
$$(1+\psi)\dot{V}(t) = V(t)^{\phi_1} I_R^*(t)^{\phi_2}$$

Clearly, $N(t)$, $P(t)$, and other development-related variables disappear from the model. So, the counterpart of new patents in the model is simply $V(t)^{\phi_1} I_R^*(t)^{\phi_2}$, and the patent-read R&D coincides with investment $I_R^*(t)$.

Then, I simulate the policy episode in the model, and I compare its response to the reduced-form estimates of Section 3, in the same way this is done for the estimation of the benchmark model's structural parameters. For sake of comparability between the settings, I fix the parameter of the model without development lags to the values estimated for the benchmark model of Section 5 (parameters reported in Table 2 of the paper). Figure F.1 shows the response of new innovations in the model and in the data. The standard model fails to capture the pre- and post-implementation response shown by the data, and the responses of innovation and R&D, plotted in Figure F.2, are flat at the news and drop at the policy implementation, slowly adjusting to toward the new steady-state.

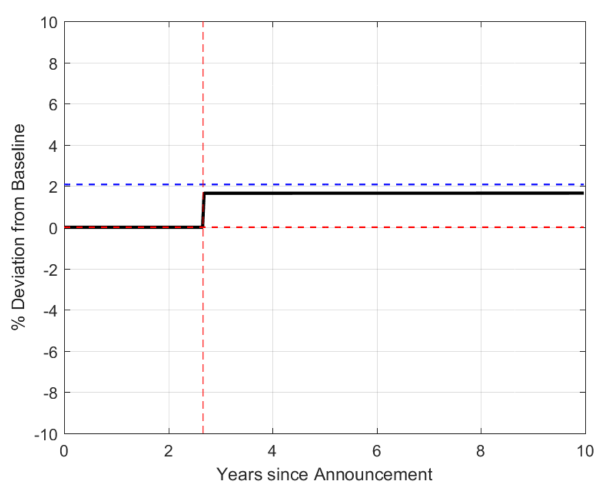
As mentioned in the paper, the cause of this dynamics is that the value of a patent, at the announcement of a future patent term increase, does not jump much, barely changing innovation incentives until the actual policy implementation. When the increase is implemented, however, patent value and investment increase. Investment does not increase immediately to the new steady state because $V(t)$ is slow-moving, and the transition features a productivity of research, given by $V(t)^{\phi_1}$, that is temporarily lower than in the new steady-state V, keeping investment lower during the transition.

Figure F.1: **New patents response to simulated policy: Jones (1995) model**



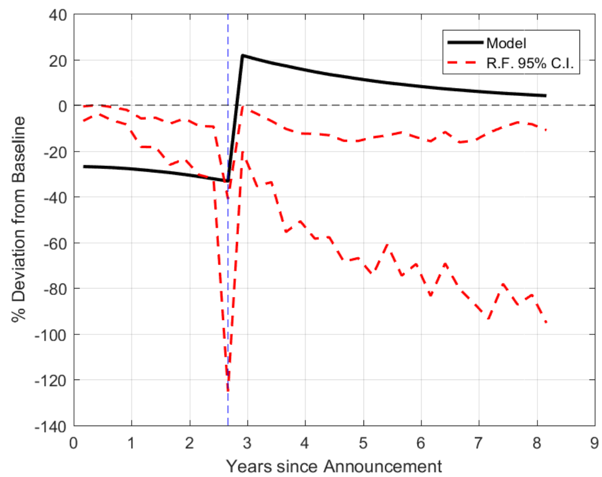
The graph shows the response of new patents $V(t)^{\phi_1} I_R(t)^{\phi_2}$ to the news of a patent term increase of 100 days, starting from $T_0 = 17$ years, implemented 2 years and 8 months after the announcement. The vertical blue dashed line refers to the implementation moment and the system is assumed to be in the steady state associated with $T_0 = 17$ years before the news at time 0. The black solid line is the response implied by the standard Jones (1995)'s model of semi-endogenous growth. All the parameters are set to the values estimated in Table 2. The red dashed lines are the 95% confidence bands of the empirical reduced form estimated of Section 3.

Figure F.2: **R&D effort response to simulated policy: Jones (1995) model**



The graph shows the response of R&D investment $I_R(t)$ to the news of a patent term increase of 100 days, starting from $T_0 = 17$ years, implemented 2 years and 8 months after the announcement. The vertical blue dashed line refers to the implementation moment and the system is assumed to be in the steady state associated with $T_0 = 17$ years before the news at time 0. The black solid line is the response implied by the standard Jones (1995)'s model of semi-endogenous growth. All the parameters are set to the values estimated in Table 2.

Figure F.3: **New patents response to simulated policy: Model with no spillover**



The graph shows the response of new patents $\iota_D(t)N(t)$ to the news of a patent term increase of 100 days, starting from $T_0 = 17$ years, implemented 2 years and 8 months after the announcement. The vertical blue dashed line refers to the implementation moment and the system is assumed to be in the steady state associated with $T_0 = 17$ years before the news at time 0. The black solid line is the response implied by the model of Section 5 shutting down the externality channel by setting $\chi = 0$. All the other parameters are set to the values estimated in Table 2. The red dashed lines are the 95% confidence bands of the empirical reduced form estimated of Section 3.

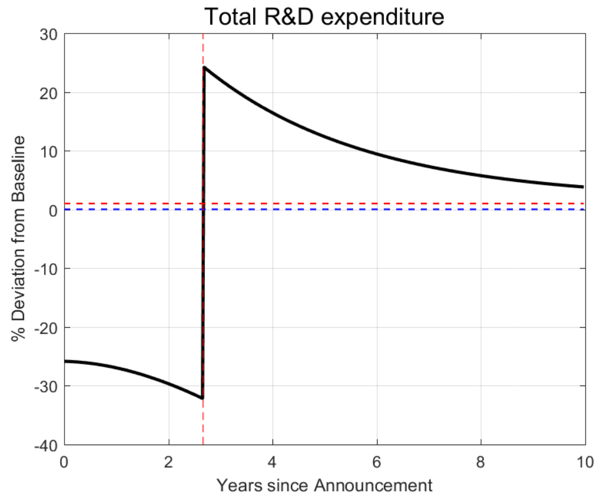
F.2 Performance of the model with no spillover

In this subsection, I illustrate what would be the behavior of patenting and of the model counterpart of patent-read R&D effort if the externality term of the benchmark model is muted, i.e. if the parameter χ is set to zero. In response to the new of a patent term increase, the pre-implementation deceleration, which is driven by the drop in development intensity, is unaffected. However, absent the externality, the post-implementation persistence vanishes, and patenting jumps to a higher level than the old steady state, converging to the new one from above. This occurs because $V(t)$ is slow moving, and its drop in the pre-implementation phase is not strong enough to counteract the effect of the implementation of the longer term on research investment.

F.3 Performance of the model if creative destruction depends on patent length

This subsection shows how the positive and normative performance of the model changes if the rate of creative destruction ψ is assumed to be a function of patent length T. I consider two cases. First, $\lambda(t) = \Big(\max\{\psi + 0.001(T-17)*365; 0\}\Big)\frac{\dot{V}(t)}{V(t)}$, i.e. the rate of creative destruction is a linear function of deviations of patent length from the status quo of 17 years. In this case, a longer patent length linearly increases creative destruction. Figure F.5 shows the positive performance of the modified model in matching the empirical response of inno-

Figure F.4: **R&D expenditure (flow) response to simulated policy: Model with no spillover**



The graph shows the response of the flow of R&D expenditure to the news of a patent term reduction of 100 days, starting from $T_0 = 17$ years, implemented 2 years and 8 months after the announcement. The vertical blue dashed line refers to the implementation moment and the system is assumed to be in the steady state associated with $T_0 = 17$ years before the news at time 0. The black solid line is the response implied by the model of Section 5 shutting down the externality channel by setting $\chi = 0$. All the other parameters are set to the values estimated in Table 2. The red dashed lines are the 95% confidence bands of the empirical reduced form estimated of Section 3.

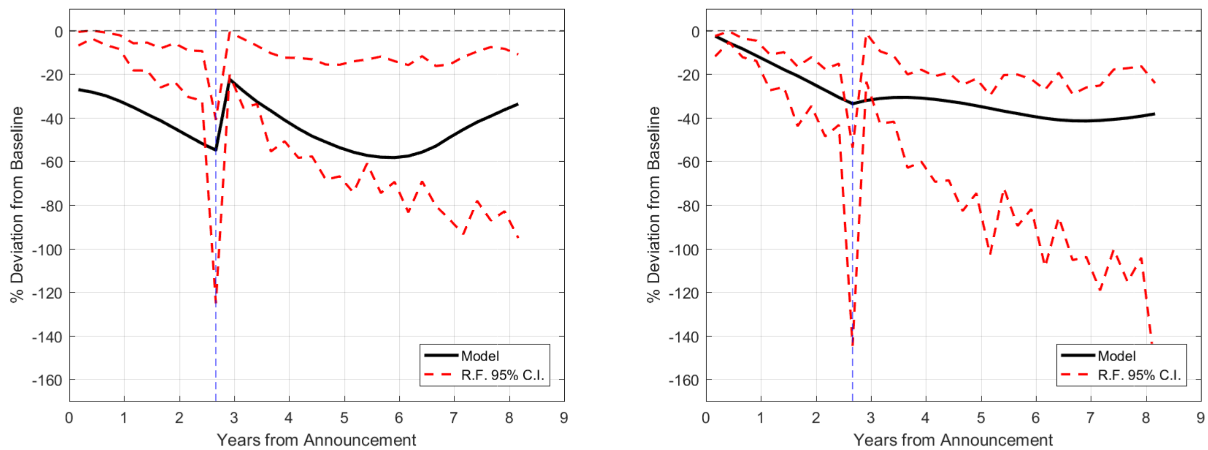
vation and R&D to a 100-days patent term increase anticipated by 2 years and 8 months. The performance is unaffected relative to the benchmark model. Figure F.6 shows the welfare change from the *unanticipated* implementation of a new patent length (left panel) and the implied rate of creative destruction ad a function of T (right panel).

In the second case I consider, creative destruction depends quadratically on deviations of T from 17 years, i.e. $\lambda(t) = \Big(\max\{\psi + 0.001(T-17)*365 - 0.0001[(T-17)*365]^2; 0\}\Big)\frac{\dot{V}(t)}{V(t)}$. Figure F.7 shows the positive performance of the modified model in matching the empirical response of innovation and R&D to a 100-days patent term increase anticipated by 2 years and 8 months. The performance is unaffected relative to the benchmark model. Figure F.8 shows the welfare change from the *unanticipated* implementation of a new patent length (left panel) and the implied rate of creative destruction ad a function of T (right panel).

F.4 Model with labor as R&D input

In this Section, I will present the details of an alternative specification of the model presented in Section 5: Research and development activities are not carried out by employing units of the final good, but they are carried out by hiring labor at the competitive wage $w(t)$. As in subsection C.1, I will start from the maximization problem faced by agents, the derivation of the aggregate laws of motion, and the derivation of the balanced growth path. The consumer

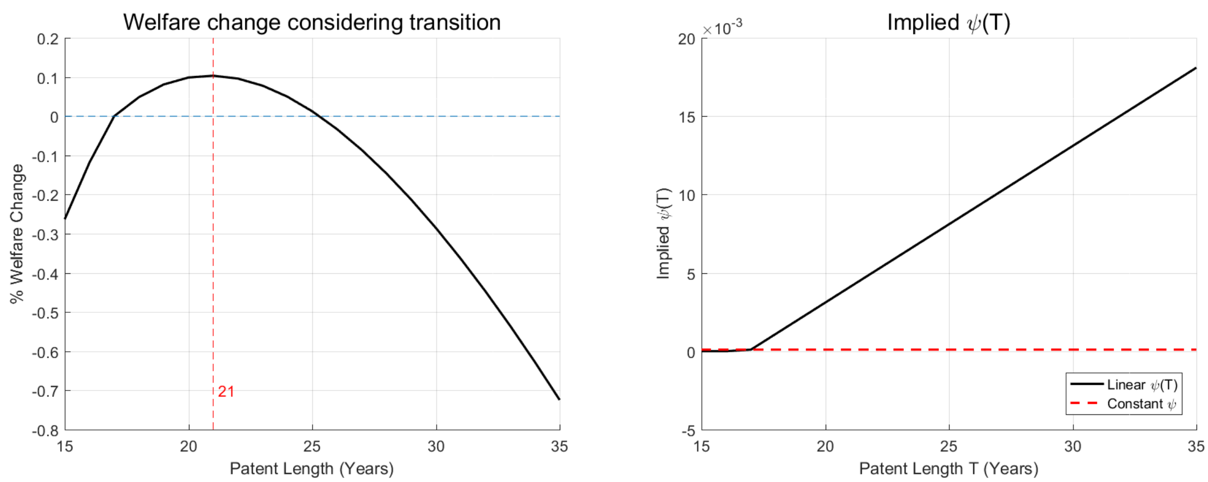
Figure F.5: **Model-based simulation of the policy and targeted reduced-form estimates**



(a) New patents (b) R&D Investment

The black solid lines are the model-based responses of the model with parameter values reported in Table 2–with creative destruction linearly depending on patent length–and the red dashed lines are 95% confidence bands of the reduced form estimates of Section 3. The system is assumed to be at the pre-policy change steady state at $t = 0$, when the news of 100-days increase in protection time implemented after 2 years and 8 months (blue vertical line) happens.

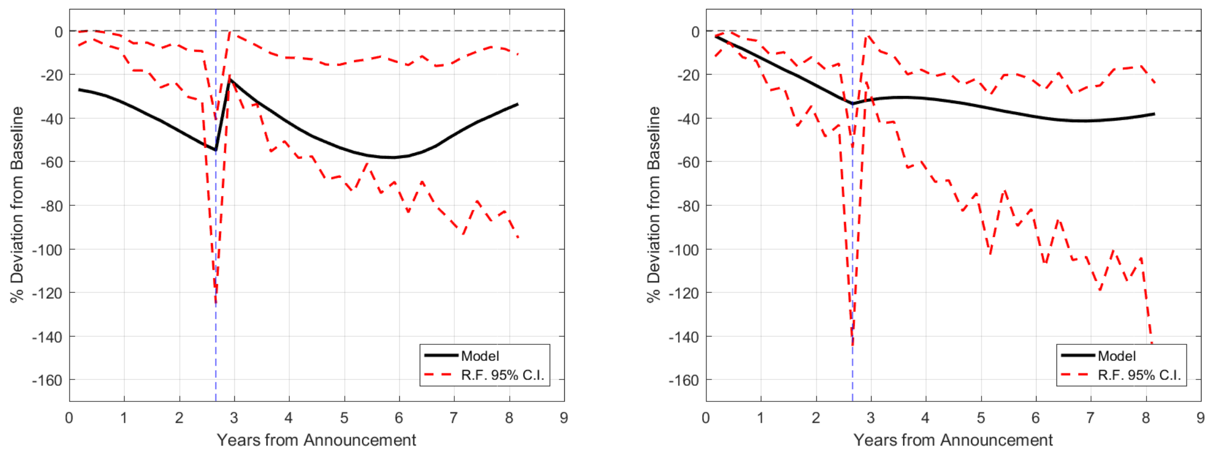
Figure F.6: **Welfare change from an unanticipated policy change**



(a) Welfare (b) Creative destruction

The left panel shows the change in the welfare index (19)–relative to the status quo–from the unanticipated implementation of a new patent length. The right panel shows how the rate of creative destruction is assumed to change with patent length T (black solid line) and it plots it against the rate of creative destruction assumed in the benchmark model (red dashed line).

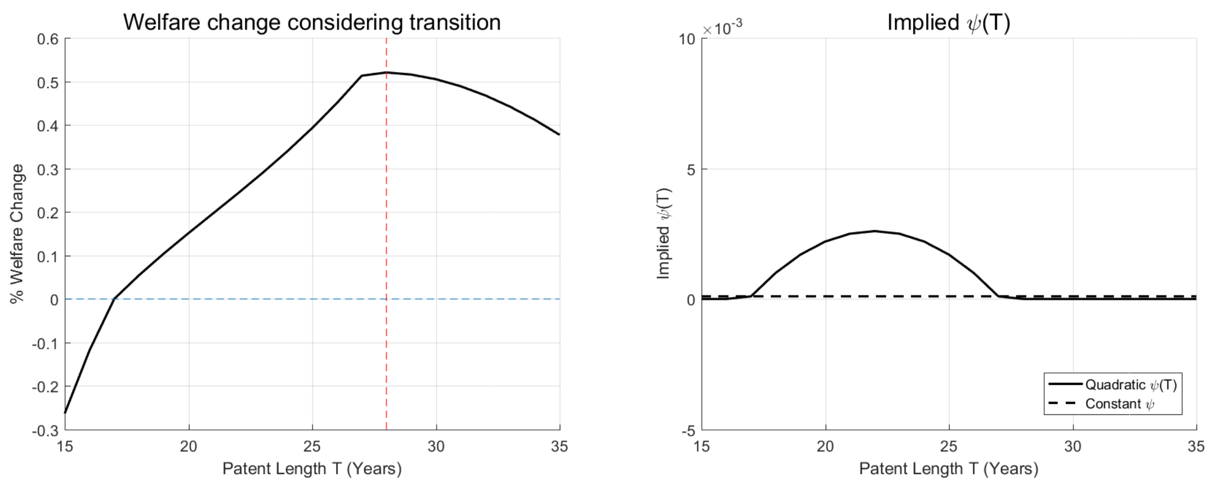
Figure F.7: **Model-based simulation of the policy and targeted reduced-form estimates**



(a) New patents (b) R&D Investment

The black solid lines are the model-based responses of the model with parameter values reported in Table 2–with creative destruction quadratically depending on patent length–and the red dashed lines are 95% confidence bands of the reduced form estimates of Section 3. The system is assumed to be at the pre-policy change steady state at $t = 0$, when the news of 100-days increase in protection time implemented after 2 years and 8 months (blue vertical line) happens.

Figure F.8: **Welfare change from an unanticipated policy change**



(a) Welfare (b) Creative destruction

The left panel shows the change in the welfare index (19)–relative to the status quo–from the unanticipated implementation of a new patent length. The right panel shows how the rate of creative destruction is assumed to change with patent length T (black solid line) and it plots it against the rate of creative destruction assumed in the benchmark model (red dashed line).

side is identical to the main model. A risk-neutral representative agent maximizes utility by choosing the optimal consumption-saving behavior, and supplies a total of $L(t)$ units of labor, which are used for the production of the final good in quantity $L_P(t)$, for research in quantity $L_R(t)$, and for development in quantity $L_D(t)$. The wage is the same for all three types of labor, in equilibrium. The agent can save in physical capital or in the shares of intermediate good firms, but both assets must deliver the same net rate of return under no-arbitrage condition. Therefore, the dynamic consumer's problem delivers the equilibrium condition $r(t) = \rho$, stating that the net return of savings must be equal to the discount rate of the consumer.

F.4.1 Final good production

The final good is produced by a competitive firm that chooses labor and the optimal quantity of each of the intermediate goods in the economy to maximize profits. The problem is

$$\max_{\{X(i,t)\}_{i\in[0,V(t)]},L_P(t)\}} \Big[h(t)L_P(t)\Big]^{1-\alpha}\Big[\int_0^{V(t)} X^{\alpha}(i,t)di\Big] - \int_0^{V(t)} z(i,t)X(i,t)di - w(t)L_P(t)$$

where output is

$$Y(t) = \Big[h(t)L_P(t)\Big]^{1-\alpha}\Big[\int_0^{V(t)} X^{\alpha}(i,t)di\Big] \tag{65}$$

and it is determined by $h(t)$, which is an exogenous labor-augmenting technology term that grows exponentially at a given constant rate g_h, $L_P(t)$, which is the hours devoted to production, exponentially growing at constant rate n, and a mass of $V(t)$ intermediate capital goods varieties. $w(t)$ is the wage rate and $z(i,t)$ is the instant-t price of intermediate variety i. The first order conditions of the problem are

$$w(t) = (1-\alpha)h(t)^{1-\alpha}L_P(t)^{-\alpha}\Big[\int_0^{V(t)} X^{\alpha}(i,t)di\Big] \tag{66}$$

and

$$z(i,t) = \alpha h(t)^{1-\alpha}L_P(t)^{1-\alpha}X^{\alpha-1}(i,t) \;\; \forall\, i \in [0, V(t)] \tag{67}$$

The former equation determines the equilibrium wage rate and it is the inverse demand for production labor, while the latter equation is the inverse demand for intermediate i.

F.4.2 Monopolistic intermediate goods production

A share $\zeta(t)$ of the existing $V(t)$ intermediate goods varieties are protected by a monopoly, granted by a valid patent. The monopolistic producer of variety i chooses the quantity to produce in order to maximize profits subject to the inverse demand given by (67), and subject to the production function. In particular, one unit of each of the intermediate goods can be produced by using one unit of raw capital $K(t)$, which can be rented from households at a rate $r_K(t) = r(t) + \delta$, where δ is the depreciation rate of physical capital. Therefore, the maximization problem is

$$\max_{X(i,t),z(i,t)} \Big\{ z(i,t)X(i,t) - (r(t)+\delta)X(i,t) \Big\}$$
$$\text{s.t.} \quad z(i,t) = \alpha h(t)^{1-\alpha} L_P(t)^{1-\alpha} X^{\alpha-1}(i,t)$$

and the first order condition implies

$$z(i,t) = \alpha(h(t)L_P(t))^{1-\alpha} X(i,t)^{\alpha-1} = \frac{1}{\alpha}(r(t)+\delta) \tag{68}$$

i.e. the price is a constant markup $1/\alpha$ over the marginal cost $(r(t)+\delta)$. This implies that the price of each of the monopolistically-produced intermediate capital varieties is the same and, therefore, also the produced quantity and the profits will be symmetric. In particular, these will satisfy

$$X(i,t) = X_p(t) = \alpha^{\frac{2}{1-\alpha}}(r(t)+\delta)^{-\frac{1}{1-\alpha}} h(t)L_P(t) \quad \forall\, i \in [0, \zeta(t)V(t)] \tag{69}$$

$$\pi(i,t) = \pi(t) = \Big(\frac{1}{\alpha} - 1\Big)(r(t)+\delta)X_p(t) \tag{70}$$

F.4.3 Non-monopolistic intermediate goods production

A fraction $1-\zeta(t)$ of intermediates are not monopolistically produced because legal patent protection on it has expired. These non-monopolistic varieties are produced in a regime of Bertrand competition, and therefore the price $z(i,t)$ is equal to the marginal cost of production $(r(t)+\delta)$. It follows from the inverse demand function (67) that the production of these competitively-produced intermediate varieties is symmetric and given by

$$X_{np}(t) = \alpha^{\frac{1}{1-\alpha}}(r(t)+\delta)^{-\frac{1}{1-\alpha}} h(t)L_P(t) \quad \forall\, i \in (\zeta(t)V(t), V(t)] \tag{71}$$

which implies that $X_p(t) = \alpha X_{np}(t)$. Since $\alpha \in (0,1)$ by assumption, this implies that the quantity produced of monopolistic varieties is lower than the one of competitive varieties, which is what the distortion of monopoly consists in.

F.4.4 Physical capital market clearing condition

The equilibrium in the physical capital market requires that the quantity of capital supplied by households $K(t)$ is equal to the quantity of capital demanded by firms to produce the intermediate capital goods, i.e.

$$\begin{aligned} K(t) =& \zeta(t)V(t)X_p(t) + (1-\zeta(t))V(t)X_{np}(t) \\ =& [\alpha\zeta(t) + (1-\zeta(t))]V(t)X_{np}(t) \end{aligned} \tag{72}$$

F.4.5 Research investment to discover new projects

The model features an unit mass of identical firms that invest in research. The output of research investment is new ideas, on which the successful firm can exclusively invest in order to develop the idea into a new intermediate variety. The research investment problem of the representative research firm is

$$\max_{L_R(t)} \left\{ P(t)E(t)^{\chi}V(t)^{\phi_1}L_R(t)^{\phi_2} - w(t)L_R(t) \right\}$$

Research requires $L_R(t)$ units of labor for the production of $E(t)^{\chi}V(t)^{\phi_1}L_R(t)^{\phi_2}$ new ideas, where $E(t)^{\chi}$ is the delayed externality term already discussed in Sections 4 and 5, and $V(t)^{\phi_1}$ is an externality from existing varieties that is common in endogenous growth models. Parameters are constrained so that $\phi_1 + \phi_2 < 1$. $\phi_1 < 1$ captures the fact that ideas become harder to find as the knowledge frontier expands, and $\phi_2 < 1$ captures the degree of decreasing returns to scale in research investment. Finally, $P(t)$ is the economic value of a new idea, or, alternatively, it can be thought as the exclusivity value of a development project. The optimal research investment is given by

$$L_R(t) = \left[\frac{\phi_2}{w(t)}P(t)E(t-d)^{\chi}V(t)^{\phi_1}\right]^{\frac{1}{1-\phi_2}}$$

F.4.6 Development of projects

Development occurs independently on each existing project, even in the case when a single firm is running multiple projects. Therefore, for each project firms hire labor to obtain a patentable intermediate variety, and firms are successful with a Poisson arrival rate of

$$\left(\frac{l_D(t)V(t)}{L(t)}\right)^{\frac{1}{\theta}}$$

where $\theta > 1$ still captures the cost-convexity of the intensity with which development is carried out. $l_D(t)$ can be interpreted as development labor intensity. The innovation arrival rate is re-scaled by the total labor force, so that $l_D(t)/L(t)$ can be interpreted as the share of the labor force on each development project, and it is increasing in the number of existing varieties $V(t)$, to make sure that a balanced growth path is admissible for this economy. Then, the development problem can be written in recursive form as

$$r(t)P(t) - \dot{P}(t) = \max_{l_D(t)} \Big\{ \Big(\frac{l_D(t)V(t)}{L(t)} \Big)^{\frac{1}{\theta}} \Big[v(t) - P(t) \Big] - w(t) l_D(t) \Big\} \tag{73}$$

where the equation captures the fact that if, with instantaneous probability $\left(\frac{l_D(t)V(t)}{L(t)}\right)^{\frac{1}{\theta}}$ the project is successful, the investing firm receives a value $v(t)$ for the intermediate variety obtained, but it loses the value of the project $P(t)$, which expires after completion. The expected value of a newly patented variety, which is what the developing firm cares about when working on the project, is

$$v(t) = \int_t^{t+T} e^{-\int_t^s (r(t')+\lambda(t'))dt'} \pi(s) ds \tag{74}$$

where $\pi(t)$ is the flow of profits at instant t, $r(t)$ is the real interest rate, and $\lambda(t)$ is an endogenous Poisson rate at which, depending on aggregate innovation intensity, a monopoly can be creatively destroyed. Therefore, $v(t)$ is the expected net present discounted value of profits on a variety. The optimal labor hiring decision on each development project is

$$l_D(t) = \left[\frac{1}{\theta} \Big(\frac{V(t)}{L(t)} \Big)^{\frac{1}{\theta}} \Big(v(t) - P(t) \Big) \frac{1}{w(t)} \right]^{\frac{1}{1-\frac{1}{\theta}}} \tag{75}$$

The dynamic spillover term must be re-defined here as

$$E(t) \equiv d^{-1} \int_{t-d}^{t} \Big(\frac{l_D(s)V(s)}{L(s)} \Big)^{\frac{1}{\theta}} ds$$

which is identical in spirit the the expression of the benchmark model of Section 5, because the externality is simply a function of the development completion probability.

The process of creative destruction captured by the $\lambda(t)$ term is endogenous and it is driven by the rate of growth of the number of varieties $V(t)$. Specifically, it is defined as

$$\lambda(t) \equiv \psi \frac{\dot{V}(t)}{V(t)}$$

i.e. in times when the rate of growth of varieties is higher, the rate of creative destruction is higher. This strategy to model entry and creative destruction is motivated by the fact that, in the data, I observe that innovation by new applicants increases when overall innovation rate increases, but their relative weight does not change.

F.4.7 Evolution of aggregate quantities

The decisions resulting from the previous optimization problems shape the evolution of aggregate quantities as follows. First, the number of varieties $V(t)$ evolves according to

$$(1+\psi)\frac{\dot{V}(t)}{V(t)} = \Big(\frac{l_D(t)V(t)}{L(t)}\Big)^{\frac{1}{\theta}} N(t) \tag{76}$$

where $\psi\frac{\dot{V}(t)}{V(t)}$ is by how much creative destruction reduces the mass of intermediate goods available, while $\left(\frac{l_D(t)V(t)}{L(t)}\right)^{\frac{1}{\theta}} N(t)$ is the number of development projects successfully turned into a variety. This is the case because $\left(\frac{l_D(t)V(t)}{L(t)}\right)^{\frac{1}{\theta}}$ is the instantaneous probability that each of the existing projects $N(t)$ is successfully completed, generating varieties. Since this instantaneous probability is identical and independent across projects, a suitable law of large numbers applies, and the aggregate representation provided holds.

The evolution of projects is instead given by

$$\dot{N}(t) = E(t-d)^{\chi} V(t)^{\phi_1} L_R(t)^{\phi_2} - \Big(\frac{l_D(t)V(t)}{L(t)}\Big)^{\frac{1}{\theta}} N(t) \tag{77}$$

where the first term captures the mass of new projects generated by research investment and the second term captures the destruction of projects due to successful completion.

The evolution of the share of existing varieties that are covered by monopoly, i.e. $\zeta(t)$, is given by

$$\dot{\zeta}(t) = (1-\zeta(t))\frac{\dot{V}(t)}{V(t)} - (1+\psi)\frac{\dot{V}(t-T)}{V(t)} e^{-\int_{t-T}^{t}\lambda(t')\mathrm{d}t'} \tag{78}$$

where the first term captures the additions to the monopolistic varieties due to current innovation, and the second term captures the fact that all those varieties that have not already been creatively destroyed become competitive when the maximum patent term T expires.

The evolution of aggregate capital satisfies

$$\dot{K}(t) = I_K(t) - \delta K(t) \tag{79}$$

where $I_K(t)$ is the investment in physical capital done by the households out of the final good, and $\delta K(t)$ is the depreciation of the existing stock.

F.4.8 Market clearing in the goods market

Given the production decisions of the intermediate varieties producers and of the final good producer, GDP for this economy can be rewritten as

$$Y(t)=[\alpha^{\alpha}\zeta(t)+(1-\zeta(t))]V(t)h(t)^{1-\alpha}L_P(t)^{1-\alpha}X_{np}^{\alpha}(t) \tag{80}$$

where $[\alpha^{\alpha}\zeta(t)+(1-\zeta(t))]V(t)h(t)^{1-\alpha}$ is the measured TFP. Notice that the productivity of the economy grows with the number of varieties available, and decreases with the share of monopolistic varieties, as $\alpha^{\alpha}<1$.

On the other hand, the total production of the final good must also satisfy

$$Y(t)=C(t)+I_K(t) \tag{81}$$

as consumption and capital investment are funded out of the final good.

F.4.9 Market clearing in the labor market

Market clearing of labor market requires that the exogenous amount of labor available $L(t)$ equals the sum of labor used in production, research, and development in equilibrium.

$$L(t)=L_P(t)+L_R(t)+L_D(t) \tag{82}$$

where $L_D(t)=l_D(t)N(t)$ is the total labor used in development, given by the labor $l_D(t)$ optimally hired on each project times the number of projects.

F.4.10 Balanced growth path

Population $L(t)$ and the productivity term $h(t)$ exogenously grow at constant rate n and g_h, respectively. Also, since $r(t)=\rho$, the real interest rate is constant. From the labor market clearing condition, it follows that production labor $L_P(t)$ and research labor $L_R(t)$ must also grow at rate n, while labor employed in each single project $l_D(t)$ must grow at less than n, namely at n minus the growth rate of projects. From equations (69), (71), and (70), it follows that the growth rate of $X_p(t)$, $X_{np}(t)$, and profits along the balanced growth path is identical and equal to g_h+n. Also, from the definition of $\upsilon(t)$, it follows that the patent value must grow at the same rate of profits, i.e. g_h+n, and that, as a consequence, the rate of creative destruction $\lambda(t)$ is constant along the balanced growth path. In addition, from the value function of the development problem, it follows that, for a b.g.p. to be possible, $P(t)$ must grow at the same rate of $\upsilon(t)$, i.e. $g_P=n+g_h$, and that the arrival rate of innovations $\left(\frac{l_D(t)V(t)}{L(t)}\right)^{\frac{1}{\theta}}$ must be constant. This implies that the rate of growth of $l_D(t)$ must be equal to population growth n minus the rate of growth of varieties. Notice that a constant $\left(\frac{l_D(t)V(t)}{L(t)}\right)^{\frac{1}{\theta}}$ is consistent with the optimality condition (75), and it implies that also the externality term $E(t)$ is constant in the b.g.p.. The evolution of the stock varieties in equation (76) implies that $g_V=g_N$, and the evolution of the stock of projects in equation (77) requires that

$$g_N = \phi_1 g_V + \phi_2 n \tag{83}$$

Since $g_V = g_N$, it follows that the rate of growth of endogenous productivity is $g_V = g_N = \frac{\phi_2}{1-\phi_1}n$, and the rate of growth of labor devoted to each development project is

$$g_{l_D} = n - g_V = \left(1 - \frac{\phi_2}{1-\phi_1}\right)n = \left(\frac{1-\phi_1-\phi_2}{1-\phi_1}\right)n$$

which is smaller than population growth as long as $\phi_2 > 0$. From (81), the rate of growth of consumption and capital investment must be the same as output. Hence, $g_Y = g_C = g_{I_K} =$. In addition, from (78), it follows that $\zeta(t)$ is constant along the b.g.p. and, as a consequence, the equilibrium production function (80) requires

$$g_Y = (1-\alpha)(g_h + n) + g_V + \alpha g_X$$

But $g_X = n + g_h$, and therefore the last implies $g_Y = \frac{\phi_2}{1-\phi_2}n + n + g_h$, i.e.

$$g_Y = \frac{1-\phi_1+\phi_2}{1-\phi_2}n + g_h$$

which fully solves for balanced growth path growth rates, and shows that the previous model admits a balanced growth path.