

Böhme, Enrico; Frank, Severin; Kerber, Wolfgang

Conference Paper

Optimal Incentives for Patent Challenges in the Pharmaceutical Industry

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016: Demographischer Wandel - Session: Patents, No. E11-V1

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Böhme, Enrico; Frank, Severin; Kerber, Wolfgang (2016) : Optimal Incentives for Patent Challenges in the Pharmaceutical Industry, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2016: Demographischer Wandel - Session: Patents, No. E11-V1, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/145637>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Optimal Incentives for Patent Challenges in the Pharmaceutical Industry

Enrico Böhme^{*}, Jonas Severin Frank^{**}, Wolfgang Kerber^{***}

Abstract

Patent settlements in the pharmaceutical industry between originator and generic firms have been scrutinized critically by competition authorities for delaying the market entry of generics and being therefore potentially anticompetitive. In this paper we present a model that analyzes the tradeoff between limiting the delaying of generic entry through patent settlements and giving generic firms more incentives for challenging weak patents of the originator firms. We can show that under general assumptions allowing patent settlements with a later market entry of generics than the expected market entry under patent litigation would increase consumer welfare. We introduce a policy parameter for determining the optimal additional period for collusion that would maximize consumer welfare and show that the size of this policy parameter depends on the size of the challenging costs, the intensity of competition, and the duration between the market entries of the first and second generic.

Keywords: patent settlements, collusion, patent challenges

JEL classification: L10, L40, O34

^{*} Post-Doc at Philipps-University Marburg, School of Business & Economics, Research Group Institutional

^{**} Doctoral Candidate at Philipps-University Marburg, School of Business & Economics, Research Group Economic Policy.

^{***} Professor of Economics at Philipps-University Marburg, School of Business & Economics, Research Group Economic Policy.

The authors thank Laura Werner and the participants of the PHD course Law and Economics of Intellectual Property Rights (Marburg, Oct. 5-8 2015) and Jan Voßwinkel and the participants of the 45. Hohenheimer Oberseminar (Marburg, Oct. 30-31 2015) for valuable comments and suggestions.

1. Introduction

It is a much discussed phenomenon that patent offices grant many patents that are later found invalid (Lemley/Shapiro 2005, Lemley 2001, Allison/Lemley 1998). Thus, patents, in reality, do not grant an ironclad “right to exclude” but a weaker “right to try to exclude” facing potentially re-examination proceedings in courts and - invalidation (Shapiro 2003, p. 395). This is why patents can be seen as “probabilistic” intellectual property (Lemley/Shapiro 2005.). If patents can be seen as “probabilistic” which means that they entail a chance to be invalid, society has an interest to let them be challenged and eventually be removed by courts (see Ayres/Klemperer 1999, Shapiro 2003 or Lemley/Shapiro 2005). From an economic perspective there exists a general tradeoff between having the benefits of the challenging of intellectual property rights and accepting the costs for incentivizing these patent challenges. This is because on one hand earlier market entry through patent challenges could increase consumer welfare but on the other hand firms need to have incentives to challenge which entails costs for society. Challenging a patent is costly for firms and it might well be that only a narrow range of patents are worth challenging. Since consumers ultimately benefit from patent challenges with the chance of removing invalid patents, it makes sense that consumers bear a share of the challenge costs. This means to let consumers take a share of challenging costs to maintain firm’s incentive to challenge (Frank/Kerber 2015).

Following this rationale, we model in our paper the implementation of a policy parameter granting the right to parties in a framework of a patent settlement to share profits for a period of time (i.e. to agree on later generic market entry) to give optimal incentives to challenge probabilistic patents. Hence the used framework of a patent settlement can be seen as an application for the more general tradeoff problem described before which we model by implementation of our policy parameter. Since challenging incentives could also be too high and market entry after a patent challenge too late, our policy parameter can also become negative.

We show that there exists an optimal policy parameter which, if implemented, yields higher consumer welfare compared to the litigation solution under the constellation that the policy parameter is non-negative, which is given in case challenging costs are sufficiently high. Under constellations where the policy parameter becomes negative, all established settlements are welfare decreasing. Also we derive that the optimal policy parameter should increase in

case challenge costs for patents increase. Using a specific example we can show that in case the second generic in our model has a later decision on market entry or in case the market is more competitive, the policy parameter should also increase.

The paper is structured as follows: Section 2 describes the basic patent settlement problem and shows the relevance of challenging incentives in this context in the literature and current discussion. Section 3 explains the model framework. In Section 4 we derive the model equilibrium and the optimal policy parameter as well as specific conditions for a feasible solution. Also we deliver an intuition of the tradeoff leading to the inner solution for the optimal parameter. In Section 5 we are deriving the consumer welfare difference between a situation with the policy parameter and the prohibition of collusive patent settlement agreements and explain the underlying effects. Section 6 provides comparative statics where we analyze the behavior of the optimal policy parameter in case of a change in our model parameters. Section 7 summarizes our results and concludes their relevance for policy recommendation.

2. Patent Settlements and Incentives for Patent Challenges

Patent Settlements in the pharmaceutical sector have caused vast discussions in the area of antitrust-/ competition law and intellectual property in the US and the EU. Agreements between originator and generic firms have been accused of violating competition rules through the sharing of supracompetitive profits in the context of patent disputes where generics challenge the validity of so-called weak patents (patents with a high probability that they are found invalid in courts) (Shapiro 2003, Lemley/Shapiro 2005). The core allegation is that originators induce generics to drop patent challenges, so stopping them from attacking these weak patents, by agreeing on a specific generic market entry date before regular patent expiry and/or granting a value transfer. Society is deprived of the chance that the patent is invalidated in a court proceeding and of the possibility of subsequent earlier market access by generics. Antitrust- and competition authorities in the US and the EU as well as courts, especially the US Supreme Court in its Actavis ruling, have taken action in several cases and have used the argument that particularly a high reverse payment from the originator to the generic can be a sign for parties cloaking the weakness of a patent and delaying generic

market entry (570 U. S. ____ (2013) *FTC v. Actavis*, FTC Study 2002, p. vii, European Commission Competition DG 5th Report on the Monitoring of Patent Settlements, p. 16 at para 50, Case At.39226 – *Lundbeck*). The litigation solution where both parties litigate the patent can be interpreted as a benchmark in the sense that consumers should in expectation not be worse off in a settlement agreement (this is referred to as “Shapiro criterion”, see Shapiro 2003). Thus the right to reap monopoly profits through a patent should be discounted according to the patent strength (the chance that the patent is held valid in court) which e.g. could translate into a 50% reduction of patent duration in case the patent is valid with 50% probability. So on one hand there is the argument that generic market entry in a patent settlement should not be later than expected under litigation and on the other, that it is beneficial for society if invalid patents are removed, i.e. costly patent challenges are incentivized.

The positive effect of patent challenges has been discussed in various forms (e.g. Dickey et al. 2010, p. 399, Gratz 2012, p.15.). However, only Gratz 2012 (implicitly) models the tradeoff between incentivizing patent challenges and allowing higher firm profits and later market entry in patent settlements. In her model she finds that a specifically tailored rule-of-reason assessment of patent settlements, where reverse payments from the originator to the generic are allowed, leads courts to a certain extent to erroneously uphold anticompetitive patent settlements. This at first is bad for consumers but results in higher settlement profits for the parties and higher challenging incentives overcompensating the negative effect. Gratz 2012 crucial assumption is that under a rule of reason regime courts make errors in assessing whether settlement agreements violate competition rules. An important implication from her model is, that firm’s profits from entering into a settlement agreement and also their incentive to challenge patents increase in case the error interval increases as well. This is because the error has an asymmetrical effect for firm profits: Courts wrongly accept anticompetitive patent settlements (which increase expected profits of firms) and wrongly prohibit legal settlements (which theoretically decreases expected profits of firms), but in the latter case parties could still choose litigation (which prevents the loss in expected profits for the settlement parties from the wrong court decision). As a result Gratz 2012 derives an optimal error to maximize consumer welfare (*ibid.*, pp. 13).

Thus reverse payments can be beneficial for consumers when incentives to challenge weak patents increase which have previously been inefficiently low. In the EU the guidelines for the assessment of patent settlements state that the removal of intellectual property which is invalid and thereby constitutes a blocking of innovations is beneficial for society which points

at the public interest of incentivizing patent challenges and properly assess patent settlements in face of this goal (Guidelines on the application of Article 101 of the Treaty on the Functioning of the European Union to technology transfer agreements, p. 44). The European Commission finds that regulating prices for generics (e.g. price caps) results in a decreasing speed of market entries by generics (European Commission 2009, p. 72 at para. 197). In the US there exists legislation called the Hatch-Waxman Act which allows for 180 days of marketing exclusivity for the first generic challenger of a patent as an incentive mechanism (United States Code U.S.C. Title 21 at § 355(j), Hemphill 2006).¹ Also, a cash-bounty program for successful patent challengers has been proposed in the literature (Miller 2004). Fischmann 2014 notes that a different allocation of procedural costs, alternatively to a direct bonus for patent challenges, could induce additional patent challenges.² The Actavis ruling of the US Supreme Court, with the majority opinion holding that large, unjustified reverse payments are problematic (570 U. S. ____ (2013) FTC, pp. 14) entails a dissenting opinion by Chief Justice Roberts stating that putting limits on the possibility to engage in patent settlements with certain entry dates reduces the incentives to challenge patents (570 U. S. ____ (2013) FTC v. Actavis, Roberts, C.J. dissenting pp.17).

In our model, the implemented policy parameter is constructed to allow parties of a patent settlement additional periods of collusion to broaden the range of patent challenges. Thus deviations from the benchmark litigation solution, i.e. the expected entry under litigation, can be positive for consumer welfare. This is true for generic entry *before* and *after* expected entry under litigation. It is important to note that we discuss patent challenges that potentially lead to an invalidation of a patent and not to a mere non-infringement decision by courts. This is crucial since we want to examine patent challenges which have an effect on one other generic as well. In addition, we discuss the sequential entry of two generics. There is evidence that multiple generic patent challenges happen, but are limited (Grabowski/Kyle 2007, pp. 500). From a theoretical perspective it is also difficult for an originator to settle with many

¹ The challenging of weak patents can also be described as a free rider problem since the first generic challenger has costs to invalidate the patent and all other generics profit from that. This is referred to as a public good problem with the risk to cause under provision of patent challenges (See Farrell/Merges (2004, pp. 952), Hemphill (2006, pp. 150). This different effect is discussed in the literature and also can be seen as motivation to introduce the 180 day rule in the Hatch Waxman framework. It should be noted that establishing a reverse payment patent settlement is also difficult in case there exist multiple challengers, since the patent holder needed to settle with each of them (570 U. S. ____ (2013) FTC v. Actavis, Roberts, C.J. dissenting p.16). In case for subsequent challengers there are smaller, insufficient incentives to challenge patents this problem could be solved endogenously (Fischmann 2014, pp. 423). In our analysis we do not focus on the multiple challenger problem entailing the mentioned public good character of patent challenges as we do not model endogenous interaction between different generics.

² E.g. a patent challenger could get a refund on the litigation costs independent of the outcome of litigation in case certain requirements for the patent challenge are met (Fischmann 2014, pp. 436).

generic challengers, since each of them needed to be paid off (Court of Appeals for the Eleventh Circuit, p. 38 citing Hovenkamp 2004, p 25). Also, incentives to additionally challenge might decrease if already other challengers exist and if the value of the market is not relatively high. However, especially in the US the occurrence of multiple generic challengers is not rare, which seems quite intuitive, because the first filer obtains 180 days of marketing exclusivity (Federal Trade Commission 2011, p. 138). Hence, in our model the notion of two generics seeking entry seems realistic, although we do not model the specific conditions of the US Hatch-Waxman pharmaceutical sector regulation.

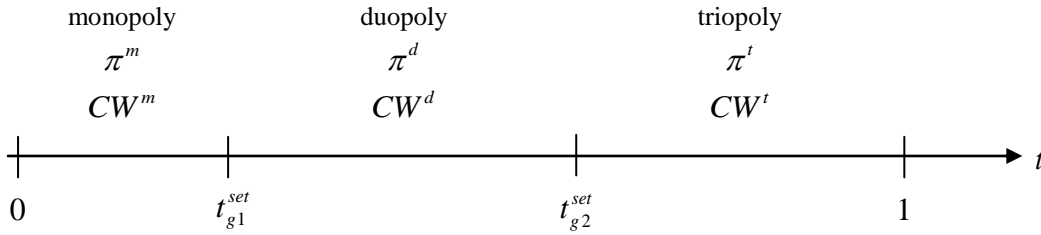
3. Model Framework

Our model follows the general framework introduced by Gratz (2012). In her model an originator (O) holds a patent with remaining patent duration $t \in [0,1]$ (whereby patent duration ends at time $t=1$) and patent strength γ , which is a random variable following a continuous uniform distribution, i.e. $\gamma \sim U[0,1]$ (whereby γ reflects the probability that the patent is found valid in court). Our model framework implies that the value of the patent for society is fixed for all patents. Two generics (G_1 and G_2) are potential entrants in the market by challenging the patent at a specific time, while facing challenge costs f_g . The value of γ is common knowledge for originator and generics. G_1 and G_2 are (potentially) entering in a fixed sequence (G_1 always enters before G_2). However, there can be situations where both enter, only G_1 enters or none enters. The decision whether the generics challenge the patent and enter the market are based on the rationale if the generated net profits from such a patent challenge, respecting challenge costs, are larger than zero. The firm's profits from a settlement consist of an even share of the settlement surplus, which is the additional profit compared to litigation, as well as the profits the firm can reap from ongoing litigation. It is assumed that firms maintain the option to litigate and hence reap, in addition to the settlement surplus, their expected litigation profits. G_1 decides on challenging the patent at time $t=0$, whereas G_2 decides to challenge at time $t=\lambda$, where $\lambda \in [0,1]$. In case the patent is held invalid, entry by the first generic occurs at time $t=0$ and entry of the second generic at time $t=\lambda$. In case the patent is held valid, entry by both generics occurs at time $t=1$. This implies that the expected entry under litigation of G_1 is $t_{g1}^{lit} = \gamma \cdot 1 + (1-\gamma) \cdot 0 = \gamma$, while the expected entry under litigation of G_2 is $t_{g2}^{lit} = \gamma \cdot 1 + (1-\gamma) \lambda > t_{g1}^l$. It is important to note that the

expected entry dates under litigation for the first- and second generic correspond to the benchmark solution or Shapiro criterion.

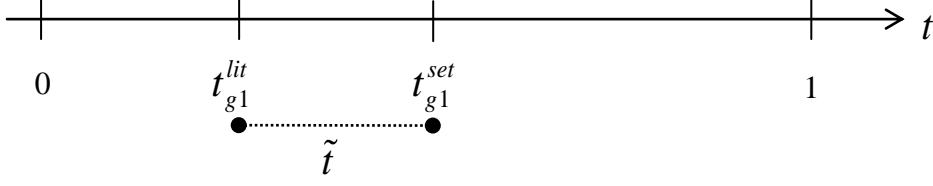
The actual entry dates of the parties under a settlement, t_{g1}^{set} and t_{g2}^{set} , are endogenous with $t_{g1}^{set} \in [0,1]$ and $t_{g2}^{set} \in [\lambda,1]$. Since market entry occurs sequentially, the originator can reap monopoly profits π^m for the period $[0, t_{g1}^{set})$, whereas for $[t_{g1}^{set}, t_{g2}^{set})$ originator and first generic generate duopoly profits $\pi_o^d + \pi_g^d = \pi^d$, where $\pi_o^d \geq \pi_g^d$. Finally, for $[t_{g2}^{set}, 1]$ the firms realize triopoly profits $\pi_o^t + 2\pi_g^t = \pi^t$, where $\pi_o^t \geq \pi_g^t$. The resulting consumer welfare is given by CW^m , CW^d , and CW^t , respectively. We assume that $\pi^m > \pi^d > \pi^t$ as well as $CW^m < CW^d < CW^t$, which corresponds to standard conditions under competitive markets. Figure 1 summarizes this situation.

Fig 1: Illustration of market structure depending on firms' market entry dates



In a patent settlement parties are able to share profits (e.g. by paying a reverse-payment from the originator to the generic). However, settlements are only possible in case they do not violate the benchmark litigation solution since they otherwise would result in later entry than expected under litigation and violate antitrust rules. This implies that they are not allowed to specify entry dates beyond $t_{g1}^{lit} = \gamma$ and $t_{g2}^{lit} = \gamma + (1-\gamma)\lambda$. Deviating from the Gratz model we introduce a policy parameter \tilde{t} in our model framework, which explicitly allows the competition authorities (and eventually the court) to grant the parties an additional time period for collusion. In particular, O and G_1 can agree to share monopoly profits until $t_{g1}^{set} = t_{g1}^{lit} + \tilde{t} = \gamma + \tilde{t}$, where t_{g1}^{set} denotes the entry of G_1 under this policy parameter. The policy parameter \tilde{t} is an *additional* percentage share of the remaining patent duration granted for collusion. This additional collusion period \tilde{t} is added to the share of the remaining patent duration in which firms could previously collude based on the benchmark litigation solution t_{g1}^{lit} which is illustrated in Figure 2.

Fig 2: Illustration of the policy parameter as share of the remaining patent duration



4. Equilibrium Analysis: The optimal Policy Parameter

Since the competition authority will not challenge all settlements where $t_{g1}^{set} \leq t_l + \tilde{t}$, firms will optimally choose the corner solution, i.e. they choose $t_{g1}^{set} = t_{g1}^{lit} + \tilde{t} = \gamma + \tilde{t}$. This implies that the corresponding entry date of a second settlement is $t_{g2}^{set} = t_{g1}^{set} + (1 - t_{g1}^{set})\lambda$. Hence, under policy \tilde{t} the joint settlement profits are given by

$$(1) \quad \Pi^{set} = \begin{cases} (\gamma + \tilde{t})\pi^m + (1 - (\gamma + \tilde{t}))[\lambda(\pi_o^d + \pi_g^d) + (1 - \lambda)(\pi_o^t + 2\pi_g^t)] & \text{for } \gamma \in [0, \gamma_{g2}^{set}] \\ (\gamma + \tilde{t})\pi^m + (1 - (\gamma + \tilde{t}))(\pi_o^d + \pi_g^d) & \text{for } \gamma \in (\gamma_{g2}^{set}, \gamma_{g1}^{set}] \end{cases}$$

Therefore, the generated surplus compared to litigation is

$$(2) \quad s_1 = \tilde{t} \left(\pi^m - \lambda(\pi_o^d + \pi_g^d) - (1 - \lambda)(\pi_o^t + 2\pi_g^t) \right) \text{ for } \gamma \in [0, \gamma_{g2}^{set}]$$

$$(3) \quad s_2 = \tilde{t} \left(\pi^m - (\pi_o^d + \pi_g^d) \right) \text{ for } \gamma \in (\gamma_{g2}^{set}, \gamma_{g1}^{set}]$$

which allows us to determine the critical levels of patent strength for which generic companies are indifferent between challenging a patent or not. Respecting that the relevant expected litigation profits for G_1 and G_2 are given by $\pi_{g1}^{lit}(\gamma) = (1 - t_{g1}^{lit})\pi_g^d = (1 - \gamma)\pi_g^d$ and $\pi_{g2}^{lit}(\gamma) = (1 - t_{g2}^{lit})\pi_g^t = (1 - \gamma)(1 - \lambda)\pi_g^t$, the critical values are determined by

$$(4) \quad \pi_{g1}^{lit}(\gamma) + \frac{S_2}{2} - f_g = 0 \Leftrightarrow \gamma_{g1}^{set} = 1 - \frac{f_g - \frac{S_2}{2}}{\pi_g^d}$$

$$(5) \quad \pi_{g2}^{lit}(\gamma) + \frac{S_1}{3} - f_g = 0 \Leftrightarrow \gamma_{g2}^{set} = 1 - \frac{f_g - \frac{S_1}{3}}{(1 - \lambda)\pi_g^t} < \gamma_{g1}^{set}$$

Given Equations (4) and (5) we can conclude that G_1 challenges patents for $\gamma \in [0, \gamma_{g1}^{set}]$, while G_2 challenges for $\gamma \in [0, \gamma_{g2}^{set}]$, which leads to the following market structures: If

neither G_1 nor G_2 challenges any patents, i.e. for $\gamma \in (\gamma_{g1}^{set}, 1]$, the originator's monopoly covers the entire remaining patent duration. For $\gamma \in (\gamma_{g2}^{set}, \gamma_{g1}^{set}]$ only G_1 enters the market, so the originator company holds a monopoly for $t_{g1}^{set} - 0 = \gamma + \tilde{t}$. Then, G_1 enters at t_{g1}^{set} , creating a duopoly for the time period $1 - t_{g1}^{set} = 1 - \gamma - \tilde{t}$. On the other hand, for $\gamma \in [0, \gamma_{g2}^{set}]$ we find that G_2 additionally enters the market. Here, monopoly lasts for $t_{g1}^{set} - 0 = \gamma + \tilde{t}$, duopoly for $t_{g2}^{set} - t_{g1}^{set} = (1 - \gamma - \tilde{t})\lambda$, and triopoly for $1 - t_{g2}^{set} = (1 - \gamma - \tilde{t})(1 - \lambda)$. Hence, respecting CW^m , CW^d , and CW^t , we know that consumer welfare under \tilde{t} is described by

$$\begin{aligned}
(6) \quad CW^{set}(\tilde{t}) &= \int_0^{\gamma_{g2}^{set}(\tilde{t})} [(\gamma + \tilde{t})CW^m + (1 - \gamma - \tilde{t})[\lambda CW^d + (1 - \lambda)CW^t]] d\gamma \\
&+ \int_{\gamma_{g2}^{set}(\tilde{t})}^{\gamma_{g1}^{set}(\tilde{t})} [(\gamma + \tilde{t})CW^m + (1 - \gamma - \tilde{t})CW^d] d\gamma + \int_{\gamma_{g1}^{set}(\tilde{t})}^1 CW^m d\gamma \\
&= \left[(1 - \tilde{t})\gamma_{g1}^{set}(\tilde{t}) - \frac{\gamma_{g1}^{set}(\tilde{t})^2}{2} \right] (CW^d - CW^m) + CW^m \\
&+ \left[(1 - \tilde{t})\gamma_{g2}^{set}(\tilde{t}) - \frac{\gamma_{g2}^{set}(\tilde{t})^2}{2} \right] (1 - \lambda)(CW^t - CW^d).
\end{aligned}$$

Maximizing Equation (6) with respect to \tilde{t} yields

$$\begin{aligned}
(7) \quad \frac{\partial CW^{set}}{\partial \tilde{t}} &= \underbrace{\frac{\partial \gamma_{g1}^{set}}{\partial \tilde{t}} (1 - \gamma_{g1}^{set} - \tilde{t}) (CW^d - CW^m)}_{\text{Incentive Effect - } G_1} - \underbrace{\gamma_{g1}^{set} (CW^d - CW^m)}_{\text{Entry Delay Effect - } G_1} \\
&+ \underbrace{\frac{\partial \gamma_{g2}^{set}}{\partial \tilde{t}} (1 - \gamma_{g2}^{set} - \tilde{t}) (1 - \lambda) (CW^t - CW^d)}_{\text{Incentive Effect - } G_2} - \underbrace{\gamma_{g2}^{set} (1 - \lambda) (CW^t - CW^d)}_{\text{Entry Delay Effect - } G_2}.
\end{aligned}$$

The interpretation of (7) is straightforward: If the competition authority is more generous with patent settlements, the joint settlement profits increase and hence generic companies challenge more patents. This is reflected by the incentive effects. In particular, G_1 additionally challenges $\partial \gamma_{g1}^{set} / \partial \tilde{t}$ patents, for which we have an increase of consumer welfare by $(CW^d - CW^m)$ for the period after market entry, i.e. for $1 - t_{g1}^{set} = 1 - \gamma_{g1}^{set} - \tilde{t}$. On the other hand, due to G_1 's delayed market entry, consumer welfare decreases by $(CW^d - CW^m)$ for all γ_{g1}^{set} patents, which would have been challenged anyway. The same logic applies to G_2 . Hence, the optimal policy is determined by the point where these two opposing effects marginally compensate each other, i.e. it is determined by $\partial CW^{set} / \partial \tilde{t} = 0$. Obviously, we

find that a marginal increase in \tilde{t} creates a tradeoff between incentivizing more patent challenges and creating more collusion. The next step is to find the optimal policy, which is denoted by \tilde{t}_{opt} .

Since we have that

$$(8) \frac{\partial \gamma_{g1}^{set}}{\partial \tilde{t}} = \frac{(\pi^m - (\pi_o^d + \pi_g^d))}{2\pi_g^d} \text{ and } (9) \frac{\partial \gamma_{g2}^{set}}{\partial \tilde{t}} = \frac{(\pi^m - \lambda(\pi_o^d + \pi_g^d) - (1-\lambda)(\pi_o^t + 2\pi_g^t))}{3(1-\lambda)\pi_g^t},$$

we find that \tilde{t}_{opt} can be explicitly described by

$$(10) \quad \tilde{t}_{opt} = \left[CW^d - CW^m - (CW^d - CW^t)(1-\lambda) + \frac{\varphi_2(CW^m - CW^d)f_g}{2(\pi_g^d)^2} + \frac{(CW^m - CW^d)f_g}{\pi_g^d} - \frac{\varphi_1(CW^t - CW^d)f_g}{3(1-\lambda)(\pi_g^t)^2} + \frac{(CW^m - CW^d)f_g}{\pi_g^t} \right] / \left[-\frac{\varphi_2(CW^d - CW^m)(\varphi_2 + 4\pi_g^d)}{4(\pi_g^d)^2} - \frac{(\varphi_1)^2(CW^t - CW^d)}{9(1-\lambda)(\pi_g^t)^2} - \frac{2\varphi_1(CW^t - CW^d)f_g}{3\pi_g^t} \right],$$

where $\varphi_1 = \pi^m - \lambda(\pi_o^d + \pi_g^d) - (1-\lambda)(\pi_o^t + 2\pi_g^t)$ and $\varphi_2 = \pi^m - (\pi_o^d + \pi_g^d)$. As we know that $CW^m < CW^d < CW^t$, $\lambda \in [0,1)$ and $\varphi_1, \varphi_2, \pi_g^d, \pi_g^t > 0$, we can show that

$$\begin{aligned} \partial(CW^{rr}) / \partial(\tilde{t})^2 &= (CW^d - CW^m) \left(-(\varphi_2)^2 / (4(\pi_g^d)^2) - \varphi_2 / \pi_g^d \right) \\ &+ (CW^t - CW^d)(1-\lambda) \left(-(\varphi_1)^2 / (9(1-\lambda)^2(\pi_g^t)^2) - 2\varphi_1 / (3(1-\lambda)\pi_g^t) \right) < 0. \end{aligned}$$

Hence, we can conclude that \tilde{t}_{opt} is a unique maximum solution.

Since firms would never accept any settlement profits that are lower than profits under litigation, we cannot have $\tilde{t}_{opt} < 0$ as this would imply an entry date prior to that under litigation and therefore a negative surplus. In addition, market entry cannot be delayed beyond the original patent duration, so that we must have $\tilde{t}_{opt} \leq 1$. Hence, in our model framework feasible solutions require $\tilde{t}_{opt} \in [0,1]$, which holds for $f_g \in [\underline{f}_g, \bar{f}_g]$, where

$$\underline{f}_g = \left[-6(1-\lambda)(CW^m - CW^t(1-\lambda) - CW^d\lambda)(\pi_g^d)^2(\pi_g^t)^2 \right] / \left[2\varphi_1(CW^t - CW^d)(\pi_g^d)^2 \right]$$

$$\begin{aligned}
& -3(1-\lambda)\pi_g^t \left(2(CW^d - CW^t)(\pi_g^d)^2 - (CW^d - CW^m)(\varphi_2 + 2\pi_g^d)\pi_g^t \right) \Big], \\
& \bar{f}_g = \left[(CW^d - CW^t)(\pi_g^d)^2 \left(4(\varphi_1)^2 + 24\varphi_1(1-\lambda)\pi_g^t \right) - 9(1-\lambda) \right. \\
& \quad \left. \left(\varphi_2(CW^d - CW^m)(\varphi_2 + 4\pi_g^d) + (CW^t - CW^m + \lambda(CW^d - CW^t))(\pi_g^d)^2 \right)(\pi_g^t)^2 \right] / \\
& \quad \left[6 \left(2\varphi_1(CW^d - CW^t)(\pi_g^d)^2 - 3(1-\lambda)\pi_g^t \left(2(CW^t - CW^d)(\pi_g^d)^2 + (CW^d - CW^m)(\varphi_2 + 2\pi_g^d)\pi_g^t \right) \right) \right].
\end{aligned}$$

Note that we can show that $0 < \underline{f}_g < \bar{f}_g$, which implies that $\tilde{t}_{opt} > 0$ requires challenge costs to be strictly positive and sufficiently high. In case that challenge costs are below the lower threshold level, the optimal policy would result in $\tilde{t}_{opt} < 0$, so that firms would strictly prefer to go for litigation. On the other hand, if challenge costs are above the upper threshold, we would have $\tilde{t}_{opt} > 1$. Hence, we end up with the corner solution where $\tilde{t}_{opt} = 1$. Therefore, our model is particularly relevant for those cases where challenge costs are sufficiently high, i.e. where the incentives to challenge patents are (too) small.

5. Welfare Analysis

In order to show that the implementation of \tilde{t}_{opt} is beneficial for consumers, we have to compare our results to the benchmark case where settlements are not allowed and firms go for litigation. Using Equation (6) we can easily compute consumer surplus under litigation by evaluating $CW^{set}(\tilde{t})$ at $\tilde{t} = 0$. Hence, consumer welfare under litigation is given by $CW^{set}(0)$, whereas consumer welfare under the optimal policy \tilde{t}_{opt} is determined by $CW^{set}(\tilde{t}_{opt})$. Our results are summarized in Proposition 1:

Proposition 1: For $\tilde{t}_{opt} \in (0,1]$ the implementation of \tilde{t}_{opt} strictly increases consumer welfare, i.e. we have that $CW^{set}(\tilde{t}_{opt}) > CW^{set}(0)$. For the special case where $\tilde{t}_{opt} = 0$ we obviously have that $CW^{set}(\tilde{t}_{opt}) = CW^{set}(0)$.

Proof: See Appendix A.

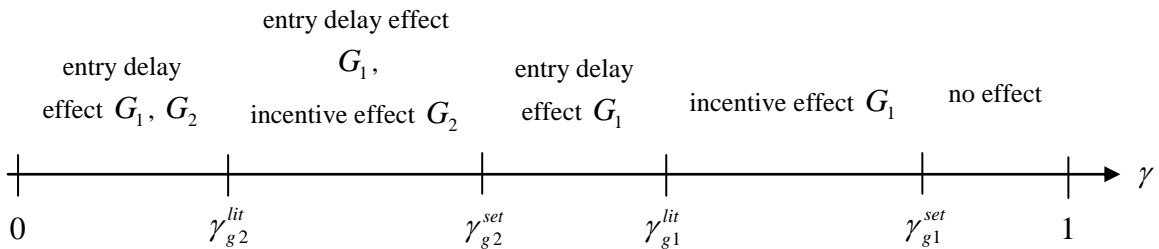
Proposition 1 shows that the implementation of \tilde{t}_{opt} is indeed welfare increasing for consumers. Hence, we can conclude that under the optimal policy the incentive effect at least partially overcompensates the entry delay effect. Note that, except for the case of $\tilde{t}_{opt} = 0$, this result holds in general, i.e. it holds for $\tilde{t}_{opt} > 0$ as well as for $\tilde{t}_{opt} < 0$. However, in the latter case firms will optimally choose to go for litigation, so that for $\tilde{t}_{opt} < 0$ consumer welfare

under \tilde{t}_{opt} is equivalent to consumer welfare under litigation. In addition, note that our results neither require any additional assumptions with respect to the mode of competition nor with respect to the intensity of competition. As a robustness check we can easily reproduce the results of Proposition 1 for the case of one generic entrant (Proof: See Appendix B).

One important implication from Proposition 1 is that in our model patent settlements can only be procompetitive for $f_g \in [\underline{f}_g, \bar{f}_g]$, i.e. for $\tilde{t}_{opt} > 0$. Here, a settlement under the optimal policy strictly increases consumer welfare. However, for $f_g < \underline{f}_g$, i.e. for $\tilde{t}_{opt} < 0$, the welfare maximizing policy would require the generic(s) to enter the market earlier than under litigation, resulting in firms to litigate. In this situation, we can conclude that each settlement that specifies an entry date where $t_{g1}^{lit} < t_{g1}^{set}$ or $t_{g2}^{lit} < t_{g2}^{set}$ is in any case anticompetitive, since we know from the uniqueness of \tilde{t}_{opt} that we are on the downward sloping part of $CW^{set}(\tilde{t})$. Hence, regulating authorities should strictly reject all settlements under these conditions. In the remaining part of this paper we exclusively refer to the case of procompetitive settlements, i.e. to $\tilde{t}_{opt} > 0$.

In order to explain the intuition behind the results of Proposition 1, we have to analyze the impact of \tilde{t}_{opt} on the range of challenged patents at first. Since we focus on $\tilde{t}_{opt} > 0$, we know that the optimal policy allows for more collusion, i.e. for a delayed market entry of both generics. Indeed, we can show that the market entry dates $t_{g1}^{set} = t_{g1}^{lit} + \tilde{t} = \gamma + \tilde{t}$ and $t_{g2}^{set} = t_{g1}^{set} + (1 - t_{g1}^{set})\lambda$ are extended beyond their counterparts under litigation, i.e. we have that $t_{g1}^{set} > t_{g1}^{lit}$ as well as $t_{g2}^{set} > t_{g2}^{lit}$. Therefore, the settlement surpluses are strictly positive, which induces both generics to challenge more patents. Hence, we have that the critical level of patent strength under such a settlement is larger than under litigation, i.e. $\gamma_{g1}^{set} > \gamma_{g1}^{lit}$ and $\gamma_{g2}^{set} > \gamma_{g2}^{lit}$, which is illustrated in Figure 3.

Fig 3: Entry Delay- and Incentive Effects in intervals of patent strength under $\tilde{t}_{opt} > 0$.



In Figure 3 we see entry delay- and incentive effects for different intervals of patent strength in case \tilde{t}_{opt} is implemented and larger than zero. The intervals we are looking at are bound by critical levels of patent strengths in which generic are just indifferent to challenge or not with- and without the policy parameter. Thus we can derive for which patent strengths incentive- and entry delay effects are relevant and affect consumer welfare. As we can see from Figure 3, under litigation G_2 would challenge γ_{g2}^{lit} patents, while G_1 challenges $\gamma_{g1}^{lit} > \gamma_{g2}^{lit}$ patents, i.e. G_1 challenges more patents than G_2 . This result is an artefact from the assumption of sequential entry of the parties: Since G_1 always enters first, it can reap more profit and, taking challenging costs into account, a broader range of patent challenges is feasible. As we already know, the critical levels of patent strength under \tilde{t}_{opt} are γ_{g2}^{set} and γ_{g1}^{set} , where we find $\gamma_{g1}^{set} > \gamma_{g2}^{set}$ for the same reason.

Analyzing the incentive- and entry delay effects, we can see in Figure 3 that for $\gamma \in [0, \gamma_{g2}^{lit}]$ both generics would challenge these patents under litigation anyway. Hence, under this interval the implementation of \tilde{t}_{opt} induces an entry delay effect of G_1 and G_2 , which negatively influences consumer welfare. For $\gamma \in (\gamma_{g2}^{lit}, \gamma_{g2}^{set}]$ we see that G_2 additionally challenges patents that would not have been challenged under litigation, which describes G_2 's incentive effect. On the other hand, this patent range would still have been challenged by G_1 under litigation, so we have an additional entry delay effect for G_1 . Hence, for $\gamma \in (\gamma_{g2}^{lit}, \gamma_{g2}^{set}]$ we find two effects, which influence consumer welfare in opposite directions. For $\gamma \in (\gamma_{g2}^{set}, \gamma_{g1}^{lit}]$ G_1 's entry delay effect is still present, while there is no effect resulting from G_2 , since it doesn't challenge these (or stronger) patents. Considering higher values of γ where $\gamma \in (\gamma_{g1}^{lit}, \gamma_{g1}^{set}]$ we find all patents that are now additionally challenged by G_1 under the optimal policy. This describes G_1 's incentive effect, which positively affects consumer welfare. For the remaining values of γ , i.e. for $\gamma \in (\gamma_{g1}^{set}, 1]$, we do not observe any effects as these patents are neither challenged under litigation nor under \tilde{t}_{opt} . Here, the implementation of \tilde{t}_{opt} creates no impact on consumer welfare. Since we know from Proposition 1 that \tilde{t}_{opt} is consumer welfare increasing, we can conclude that the incentive effects outweigh the entry delay effects.

6. Comparative Statics

The Effects of an Increase in Challenging Costs

In this chapter we analyze the comparative statics of the optimal policy parameter \tilde{t}_{opt} . As we already know that challenge costs have a crucial impact on the optimal policy, we examine at first how \tilde{t}_{opt} changes with f_g . The challenge costs include firm investments for entering the market, i.e. technology- and infrastructure investments to overcome market entry barriers, as well as settlement costs, i.e. time and effort to prepare and arrange a settlement. The corresponding result is given in Proposition 2:

Proposition 2: The optimal policy parameter \tilde{t}_{opt} is strictly increasing in f_g .

Proof: See Appendix A.

In order to interpret the result of Proposition 2, we recall that \tilde{t}_{opt} was determined by Equation (7). Analyzing (7), we see by using (8) and (9) that a marginal increase in f_g doesn't have an impact on $\partial\gamma_{g1}^{set}/\partial\tilde{t}$ and $\partial\gamma_{g2}^{set}/\partial\tilde{t}$, as we have that $\partial\gamma_{g1}^{set}/\partial\tilde{t}\partial f_g = 0$ and $\partial\gamma_{g2}^{set}/\partial\tilde{t}\partial f_g = 0$. However, a marginal increase in f_g makes it more costly to challenge patents and hence firms would ceteris paribus only challenge relatively weaker patents, i.e. patents with a smaller γ . Therefore, we have that γ_{g1}^{set} and γ_{g2}^{set} are strictly decreasing in f_g , which in turn has an impact on the incentive effects: Formally, a marginal increase in \tilde{t} increases consumer welfare by $(\partial\gamma_{g1}^{set}/\partial\tilde{t})(1-\gamma_{g1}^{set}-\tilde{t})(CW^d - CW^m)$, which describes the incentive effect of G_1 . This incentive effect is higher in case f_g increases, since γ_{g1}^{set} decreases. The interpretation is intuitive: Since γ_{g1}^{set} decreases, $t_{g1}^{set}(\gamma_{g1}^{set}) = \gamma_{g1}^{set} + \tilde{t}$ decreases as well, i.e. firms would ceteris paribus agree on earlier market entry in a settlement. Hence, consumers benefit from CW^d for a longer period of time. As this holds for all $\partial\gamma_{g1}^{set}/\partial\tilde{t}$ additionally challenged patents as well, a marginal increase in \tilde{t} has a more positive impact on CW^{set} if f_g marginally increases. The same logic applies to the incentive effect of G_2 , where a marginal increase in \tilde{t} positively affects CW^{set} by $(\partial\gamma_{g2}^{set}/\partial\tilde{t})(1-\gamma_{g2}^{set}-\tilde{t})(1-\lambda)(CW^t - CW^d)$. This incentive effect for G_2 is also higher in case f_g increases, because γ_{g2}^{set} decreases. Hence, an increase in \tilde{t} compensates for fewer patents being challenged as a consequence of an increasing f_g .

Apart from the incentive effects, we also have to take into account the entry delay effects for consumer welfare resulting from a marginal increase in \tilde{t} , i.e. from allowing for a marginally delayed market entry. In particular, we see that because of G_1 's delayed entry consumer welfare decreases by $CW^d - CW^m$ for all possible patent strengths $[0, \gamma_{g1}^{set}]$. In case of G_2 we have that consumer welfare marginally decreases by $(1-\lambda)(CW^t - CW^d)$ for all possible patent strengths $[0, \gamma_{g2}^{set}]$. As we have already seen, higher patent challenge costs, f_g , lead to lower values for the critical patent strength γ_{g1}^{set} and γ_{g2}^{set} . Hence, the entry delay effects

$\gamma_{g1}^{set}(CW^d - CW^m)$ for G_1 and $\gamma_{g2}^{set}(1-\lambda)(CW^t - CW^d)$ for G_2 , respectively, strictly decrease, because the range of patents, which is challenged anyway, ceteris paribus decreases if f_g increases. This outcome is intuitive: The entry delay effect for consumer welfare is based on a larger extent of collusion between $[0, \gamma_{g1}^{set}]$ and $[0, \gamma_{g2}^{set}]$. Since both patent strength ranges decrease in f_g , entry delay effects for G_1 and G_2 also decrease. Consequently, with respect to the marginal impact of \tilde{t} on CW^m a marginal increase in f_g leads to an increase in the incentive effects for G_1 and G_2 , whereas entry delay effects for G_1 and G_2 decrease. Hence, we obviously find that \tilde{t}_{opt} is strictly increasing in f_g .

Introducing a specific Competition Model for Parameter Analysis

The remaining exogenous parameters of our model can be divided into two different groups: While λ is an independent exogenous variable, we know that π^m , $\pi^d = \pi_o^d + \pi_g^d$, $\pi^t = \pi_o^t + \pi_g^t$ as well as CW^m , CW^d and CW^t are related as they all depend on the market structure and the mode of competition. Hence, the comparative static analysis of these “market-specific” parameters is very complex and requires many restrictive assumptions. For instance, an increase in π^m might stem from additionally exploited consumer surplus (which implies a decrease of CW^m) or from a higher market demand (which affects CW^m as well as welfare and profits under duopoly and triopoly). In order to cope with these issues, we introduce a specific competition model where i) firms compete in prices and ii) products are heterogeneous with a variable degree of substitutability. We expect assumptions i) and ii) to adequately reflect the characteristics of real world markets for pharmaceutical products. The details are summarized in Example 1.

Example 1: We consider a special case of the model presented in Häckner (2000) for the case of n firms where $n \in \{1, 2, 3\}$. In particular, we suppose the existence of a representative consumer with the utility function

$$U(\mathbf{q}, I) = \sum_{i=1}^n q_i - \frac{1}{2} \left(\sum_{i=1}^n (q_i)^2 + 2\delta \sum_{i=1, i \neq j}^n q_i q_j \right) + I$$

where q_i denotes firm i 's quantity, I represents the consumption of other goods, while δ reflects the degree of substitutability. We restrict the analysis to $\delta \in (0, 1]$, i.e. to the case where the products are substitutes. If $\delta = 0$, the products are independent (i.e. firms have monopoly power), whereas for $\delta = 1$ the products are perfect substitutes (i.e. we have perfect

competition). By solving the consumer's maximization problem subject to the budget constraint $\sum p_i q_i + p_I I = m$, where p_i denotes the price of product i , m denotes income and $p_I = 1$, we obtain firm k 's inverse demand function, which is

$$p_k = 1 - q_k - \delta \sum_{j=1, j \neq k}^n q_j.$$

We assume that firms compete in prices and find that firm k 's profit-maximizing prices and quantities in equilibrium are given by

$$p_k = \frac{\left[\delta^2 (n^2 - 5n + 5) + 3\delta(n-2) + 2 \right] - \delta \sum_{i=1}^n \left[\delta(n-2) + 1 \right]}{\left[\delta(n-3) + 2 \right] \left[\delta(2n-3) + 2 \right]}$$

$$q_k = \frac{\left[\left[\delta^2 (n^2 - 5n + 5) + 3\delta(n-2) + 2 \right] - \delta \sum_{i=1}^n \left[\delta(n-2) + 1 \right] \right] \left[\delta(n-2) + 1 \right]}{(1-\delta) \left[\delta(n-3) + 2 \right] \left[\delta(n-1) + 1 \right] \left[\delta(2n-3) + 2 \right]}.$$

If we focus on the case under consideration, i.e. on the case of $n \in \{1, 2, 3\}$, we finally obtain $\pi^m = \frac{1}{4}$, $\pi_o^d = \pi_g^d = (1-\delta) / \left[(2-\delta)^2 (1+\delta) \right]$, $\pi_o^t = \pi_g^t = (1-\delta^2) / \left[4(2\delta+1) \right]$ as well as $CW^m = \frac{1}{8}$, $CW^d = 1 / \left[(1+\delta)(-2+\delta)^2 \right]$, $CW^t = \left[3(1+\delta)(1+3\delta+2\delta^2) \right] / \left[8(1+2\delta)^2 \right]$. Note that in our model feasible solutions require $t_{g1}^{set}, t_{g2}^{set}, \tilde{t}_{opt}, \gamma_{g1}^{set}, \gamma_{g2}^{set} \in (0, 1)$. Under Example 1 this holds for $f_g \in [\underline{f}_g, \bar{f}_g]$, $\delta \in [\underline{\delta}, 1)$, and $\lambda \in [0, \bar{\lambda}]$, where $\underline{\delta} = 0,7478$ and

$$\begin{aligned} \bar{\lambda} = & (1792 + 2560\delta - 4736\delta^2 - 9344\delta^3 + 3808\delta^4 + 10816\delta^5 - 1704\delta^6 - 5832\delta^7 \\ & + 1155\delta^8 + 1554\delta^9 - 699\delta^{10} - 12\delta^{11} + 93\delta^{12} - 30\delta^{13} + 3\delta^{14}) / (-256 - 1536\delta - 1920\delta^2 \\ & - 1664\delta^3 + 5280\delta^4 + 5376\delta^5 - 2280\delta^6 - 6024\delta^7 + 1539\delta^8 + 1554\delta^9 - 699\delta^{10} \\ & - 12\delta^{11} + 93\delta^{12} - 30\delta^{13} + 3\delta^{14}). \end{aligned}$$

Example 1 generates feasible solutions for $\delta \in [\underline{\delta}, 1)$, i.e. for those cases where the firms' products are sufficiently close substitutes. Our model is therefore particularly relevant for the pharmaceutical industry, because consumers tend to perceive generics as being close (but imperfect) substitutes to the originator's product.

The Effects of a later Entry of the Second Generic

The entry decision of the second generic entrant is denoted by λ . We can also directly link this to the actual entry of the second generic since a later entry decision will lead to a later entry. Therefore the parameter λ reflects competitive pressure from the second generic on the

already present firms. We use Example 1 to analyze the impact of λ on \tilde{t}_{opt} , which requires studying the influence of λ on γ_{g1}^{set} and γ_{g2}^{set} at first. Our findings are summarized in Lemma 1:

Lemma 1: The settlement surplus in case of two generic entrants, s_1 , is strictly decreasing in λ , i.e. we have that $\partial s_1 / \partial \lambda < 0$. In addition, we find that $\partial \gamma_{g1}^{set} / \partial \lambda = 0$ as well as $\partial \gamma_{g2}^{set} / \partial \lambda < 0$, which implies that G_1 is not affected, while G_2 strictly challenges less patents if λ marginally increases.

Proof: See Appendix A.

The intuition behind Lemma 1 is based on two arguments: At first, the settlement surplus s_1 that firms can gain from a settlement compared to litigation is decreasing in λ , which might seem somewhat puzzling at first glance, because in general firms are able to reap a higher surplus from a settlement in case that market entry of potential entrants is delayed. Indeed, we find that an increase in λ ceteris paribus shifts the entry decision and hence the actual entry date $t_{g2}^{set} = \gamma + \tilde{t} + (1 - \gamma - \tilde{t})\lambda$ of the second generic to a later point in time. However, a marginal increase in λ also affects G_2 's expected entry date under litigation, which is $t_{g2}^{lit} = \gamma + (1 - \gamma)\lambda$. As we have that $\partial t_{g2}^{lit} / \partial \lambda > \partial t_{g2}^{set} / \partial \lambda$, we find that the marginal impact on the entry date is higher under litigation. Therefore, the surplus decreases in λ . Secondly, it is easy to see that the denominator in $\gamma_{g2}^{set} = 1 - \left(f_g - \frac{s_1}{3}\right) / ((1 - \lambda)\pi_g^t)$ is decreasing in λ , which reflects that triopoly profits from entering the market are realized for a shorter period of time. Thus, G_2 's critical level of patent strength, γ_{g2}^{set} , is decreasing in λ , because settlement surplus and triopoly profits decrease, making a patent challenge ceteris paribus less attractive. Based on Lemma 1's results, we can study the overall impact of λ on \tilde{t}_{opt} . Our findings are given in Proposition 3:

Proposition 3: Under Example 1 the optimal policy parameter \tilde{t}_{opt} is strictly increasing in λ , i.e. we have that $\partial \tilde{t}_{opt} / \partial \lambda > 0$.

Proof: See Appendix A.

To understand this result, we use Equation (7) to analyze the different effects of a marginal change in λ on $\partial CW^{set} / \partial \tilde{t}$, which determines \tilde{t}_{opt} at $\partial CW^{set} / \partial \tilde{t} = 0$. In general, we can conclude from Lemma 1 and from Equation (7) that an increasing λ influences $\partial CW^{set} / \partial \tilde{t}$ through the incentive- and through the entry delay effect of G_2 only. The corresponding

effects for G_1 remain unaffected. G_2 's entry delay effect, given in (7) by the expression $\gamma_{g2}^{set}(1-\lambda)(CW^t - CW^d)$, is decreasing in λ , because we know from Lemma 1 that $\partial\gamma_{g2}^{set}/\partial\lambda < 0$. Hence, we can conclude that the overall entry delay effect of the second generic is lower in case λ increases, which in any case positively affects \tilde{t}_{opt} . This result holds in general, i.e. it is independent from Example 1.

In addition, we have to analyze λ 's impact on G_2 's incentive effect, which is given by $(\partial\gamma_{g2}^{set}/\partial\tilde{t})(1-\gamma_{g2}^{set}-\tilde{t})(1-\lambda)(CW^t - CW^d)$ in Equation (7). Here, we find that the effect of a marginal increase in λ on this expression is ambiguous: Since we already know that γ_{g2}^{set} is decreasing in λ , we find that λ 's marginal impact on the incentive effect is ambiguous, because $(1-\lambda)$ decreases, whereas $(1-\gamma_{g2}^{set}-\tilde{t})$ increases. This ambiguity holds in any case, i.e. it is independent from the sign of $\partial\gamma_{g2}^{set}/\partial\tilde{t}\partial\lambda$. In order to understand this outcome, we analyze the second generic's actual market entry for γ_{g2}^{set} , which is given by $t_{g2}^{set}(\gamma_{g2}^{set}) = \gamma_{g2}^{set} + \tilde{t} + (1-\gamma_{g2}^{set}-\tilde{t})\lambda$. As we know from Lemma 1 that γ_{g2}^{set} is decreasing in λ , the overall effect on $t_{g2}^{set}(\gamma_{g2}^{set})$ is ambiguous. Hence, at γ_{g2}^{set} it is not clear whether the second generic actually enters earlier or later as a reaction of an increase in λ . This, however, seems to be in sharp contrast to how we previously argued in Lemma 1, where we found that a later entry decision of G_2 corresponds to later market entry and hence to a diminishing critical patent strength γ_{g2}^{set} . The difference results from the different nature of γ : When we described the influence of λ on the range of patents challenged by G_2 , t_{g2}^{set} was determined by a randomly drawn and hence exogenous γ , which does *not* depend on λ . On the other hand, if we analyze G_2 's incentive effect we have to take into account that t_{g2}^{set} depends on γ_{g2}^{set} , since the incentive effect is endogenously determined by specific patent strength values.

Overall, we can conclude in general that the entry delay effect for G_2 is strictly decreasing in λ , while the impact on the incentive effect is ambiguous. Hence, we cannot determine the overall impact on \tilde{t}_{opt} in general. However, under Example 1 we can show that the effect of an increase in λ on the incentive effect is not ambiguous anymore. Instead, we can show that $\partial[(\partial\gamma_{g2}^{set}/\partial\tilde{t})(1-\gamma_{g2}^{set}-\tilde{t})(1-\lambda)]/\partial\lambda > 0$,³ which allows us to conclude that G_2 's incentive effect strictly increases in λ . Since we know that the entry delay effect is always strictly decreasing in λ , it is easy to see that \tilde{t}_{opt} is in any case positively affected from a marginal increase in λ . This explains our findings in Proposition 3.

³ The proof has been established in Mathematica. The corresponding code is available from the authors upon request.

The Effects of an Increase in Intensity of Competition

The final step of our analysis addresses the influence of δ on the optimal policy. Since δ measures the degree of substitutability between the originator's and the generics' product(s), it may well be interpreted as reflecting the intensity of competition on the market. Again, we study the impact on the critical levels of patent strength, i.e. on $\gamma_{g1}^{set}, \gamma_{g2}^{set}$, at first. Our findings are given in Lemma 2:

Lemma 2: Under Example 1 the critical levels of patent strength are strictly decreasing in δ , i.e. we have that $\partial\gamma_{g1}^{set}/\partial\delta < 0$ as well as $\partial\gamma_{g2}^{set}/\partial\delta < 0$. In addition, it always holds that $\partial\gamma_{g1}^{set}/\partial\tilde{t}\partial\delta > 0$ and $\partial\gamma_{g2}^{set}/\partial\tilde{t}\partial\delta > 0$, i.e. the number of G_1 's and G_2 's additionally challenged patents resulting from an increase in \tilde{t} is strictly increasing in δ .

Proof: See Appendix A.

Analyzing the intuition for Lemma 2 reveals several insights about the impact of an increase in δ . If we recall that δ can be interpreted as reflecting the intensity of competition, it is easy to see that duopoly profits $\pi_o^d(\delta)$, $\pi_g^d(\delta)$ and triopoly profits $\pi_o^t(\delta)$, $\pi_g^t(\delta)$ for the firms decrease, while consumer welfare under duopoly and triopoly $CW^d(\delta)$, $CW^t(\delta)$ increases in δ . If we focus on G_1 at first, we can immediately conclude that due to the decreasing profit under duopoly, the joint settlement surplus $s_2(\delta)$ is strictly increasing. This effect ceteris paribus has a positive impact on $\gamma_{g1}^{set}(\delta) = 1 - \left(f_g - \frac{s_2(\delta)}{2}\right) / \pi_g^d(\delta)$. At the same time, $\gamma_{g1}^{set}(\delta)$ is negatively affected, because $\pi_g^d(\delta)$, i.e. G_1 's profit after market entry, is decreasing. Since we know from Lemma 2 that $\partial\gamma_{g1}^{set}/\partial\delta < 0$, we can conclude that the overall impact on $\gamma_{g1}^{set}(\delta)$ is negative. This is not surprising, because G_1 is directly affected from the decrease of $\pi_g^d(\delta)$, while the increasing settlement surplus is equally shared with the originator. The same argument holds for G_2 .

In addition, we know from Lemma 2 that $\partial\gamma_{g1}^{set}/\partial\tilde{t}\partial\delta > 0$, which also results from the decreasing $\pi_g^d(\delta)$. In general, a marginal increase in \tilde{t} affects G_1 in two ways: On the one hand, if \tilde{t} marginally increases, G_1 benefits from a longer period of collusion and hence from higher settlement profits. On the other hand, G_1 's market entry is marginally delayed, which reduces the individual profit from entering the market. However, if δ increases, the overall effect is strictly positive: Since $\pi_g^d(\delta)$ decreases, the settlement surplus increases, i.e. collusion is more valuable under a more competitive market. At the same time, the decline of $\pi_g^d(\delta)$ makes an additional delay of market entry less costly for G_1 . Hence, G_1 's number of

additionally challenged patents resulting from a marginal increase of \tilde{t} is strictly increasing in δ . Again, the same logic holds for G_2 .

Given the results from Lemma 2, we can finally analyze the overall impact of δ on the optimal policy parameter. The result is summarized in Proposition 4:

Proposition 4: Under Example 1 the optimal policy parameter \tilde{t}_{opt} is strictly increasing in δ , i.e. we have that $\partial \tilde{t}_{opt} / \partial \delta > 0$.

Proof: See Appendix A.

In order to explain our result in Proposition 4, we again use Equation (7). Once more, we distinguish entry delay effect and incentive effect, which determine the impact of our policy parameter on consumer welfare. If we focus on G_1 at first, it is easy to see that the difference $CW^d(\delta) - CW^m$ is increasing in δ . Since this expression enters both the incentive- and the entry delay effect of G_1 , we ceteris paribus find a countervailing combined effect. Hence, we have to examine the influence of $\gamma_{g1}^{set}(\delta)$ on the incentive- and the entry delay effect. We already know from Lemma 2 that $\partial \gamma_{g1}^{set} / \partial \delta < 0$, which has two effects: At first, G_1 's entry delay effect decreases, since the costs of collusion, i.e. $CW^d(\delta) - CW^m$, apply to less already challenged patents. In addition, the incentive effect increases, because the reward of higher consumer welfare through additional patent challenges is realized for a longer time period, since G_1 's market entry takes place earlier, i.e. $t_{g1}^{set}(\gamma_{g1}^{set}) = \gamma_{g1}^{set} + \tilde{t}$ decreases. Both effects positively influence \tilde{t}_{opt} . Moreover, we have found in Lemma 2 that $\partial \gamma_{g1}^{set} / \partial \tilde{t} \partial \delta > 0$, which has an additional positive impact on the incentive effect. Hence, we have that \tilde{t}_{opt} is strictly increasing in δ . For the second generic the same logic for incentive- and entry delay effect in (7) applies. Our result in Proposition 4 implies that for $\tilde{t}_{opt} > 0$ the competition authorities should grant more collusion on markets that are more competitive, which is not what we would normally expect since collusion is more harmful on competitive markets. However, in terms of our model, this counterintuitive outcome can be explained by using the above rational, since ceteris paribus the benefits from additionally challenged patents (the incentive effects) outweigh the costs of more collusion.

The results of our comparative static analysis are summarized in Table 1:

Table 1: Summary of comparative statics results

Marginal Impact on \tilde{t}_{opt}	Parameters		
	f_g	λ	δ
Incentive Effect	+	+	+
Entry Delay Effect	-	-	-
Overall Effect	+	+	+

“+” and “-” indicate that the absolute values of the effects increase/decrease.

“*” indicates that the effects can be shown under Example 1.

7. Conclusions

In our paper we model the tradeoff between granting incentives for challenging weak patents and allowing collusion between originator and generic firms. We introduce a policy parameter to specify the time for which companies are allowed to share monopoly profits additional to the expected entry under litigation. We find that there exists an optimal specification for this policy parameter to maximize consumer welfare. Our key contribution is that consumer welfare under this optimal policy parameter is higher than under the benchmark litigation solution. This result holds under very general conditions. We examine the marginal effects of an increase in challenging costs, the later entry decision of the second generic challenger and the intensity of competition on our policy parameter and find that it should increase.

Our results have substantial implications for the competition policy assessment of patent settlements. Many scholars, competition authorities and courts suggest that a high reverse payment is a sign to determine the anticompetitiveness of patent settlements. The Supreme Court in the US in line with the US Federal Trade Commission (FTC) argued that a high and otherwise unexplained reverse payment from the originator to the generic is a sign for patent weakness and an antitrust violation if entry is delayed (570 U. S. ____ (2013) FTC v. Actavis). Also the European Commission states in its guidelines on the assessment of patent settlements that large reverse payments are critical (Guidelines on the application of Article 101 of the Treaty on the Functioning of the European Union to technology transfer agreements, p. 44). Our model shows that this is not necessarily the case since we add a new

perspective on the assessment of patent settlements which is to maintain optimal challenging incentives for weak patents.

We need to distinguish two cases:

First, depending on parameter constellations, the policy parameter can be negative. If challenging costs are particularly low in relation to the exogenous parameters, i.e. below \underline{f}_g , the policy parameter becomes negative which means that challenging incentives are too high and market entry too late. Under these conditions every patent settlement which is established by the parties decreases consumer welfare. Also, if the parties could (be forced to) engage into a specific agreement with earlier market entry than under litigation, consumer welfare could be increased. This approach would be clearly different from the Shapiro criterion since it would declare all settlements to violate antitrust rules where a better- or less restrictive alternative for consumers could be found (Shapiro 2003, pp. 396). Such settlements are not realistic since agreements which are worse than the litigation option for the parties are not feasible. A state subsidy might incentivize parties to agree to an earlier market entry than under litigation, but this would also affect consumer welfare and change our model since the subsidy needs to be financed.

Secondly, there exist parameter constellations which allow a non-negative policy parameter. This means that periods of collusion deviating from the expected entry under litigation with substantial reverse payments can be welfare increasing, especially under relatively high challenge costs, later entry decisions of the second generic or a market with relatively intense competition:

In case the policy parameter is non-negative we find that:

- Higher challenge costs reduce the range of patent strengths which are profitable to challenge. Thus to allow more collusion offsets this effect with the cost that for already challenged patents later market entries are created – consumer welfare increases overall.
- More time between the market entry of the first and second generic should allow for more collusion to increase welfare. Although the later entry decision of a second generic is a rather theoretical parameter which is hard to interpret and to assess in actual patent law cases, it is ceteris paribus directly linked in the model to the actual later entry of the second generic which, as a result, could be interpreted as competitive pressure by the second generic. To allow more collusion in situations where the entry decision of the

second generic occurs later is welfare increasing. If the second generic enters later, it can reap lesser triopoly profits and challenge incentives decrease. To maintain challenging incentives it is more valuable for the second generic to get a share of collusive profits instead of entering earlier. The benefit of additionally challenged patents by the second generic offsets the costs of later entry for already challenged patents by both generics.

- More intense competition between firms in the market triggers reduced incentives to challenge patents and thus should allow for more collusion to increase welfare. This on first glance counterintuitive result is explained by our model since benefits of additionally challenged patents in the future outweigh the loss occurring from later entry for already challenged patents.

It should be noted that the policy parameter theoretically could be such as to allow for collusion larger than the patent duration. In our model we do not include such a case. However since it is not clear whether the actual patent duration granted by patent offices is an optimal one, the question of different periods of granted collusion can be raised. It might be that there exist parameter constellations that challenging incentives are optimal in case firms collude even after patent expiry. Patents inherently aim at solving the problem of optimal innovation incentives in the tradeoff with resulting price effects from monopoly. We do not include in our analysis the perspective of innovation effects. Thus we do not model that we might eventually extend the actual patent duration (to incentivize more patent challenges) since the question of the optimal duration of patents relies on innovation effects as well. We see in our model the patent duration as a given limitation of collusion.

Challenge costs or the competitiveness of the market environment are parameters which likely play a significant role in actual patent settlement cases. If it is observed that, under a non-negative policy parameter, e.g. it is relatively costly to challenge the patent and enter the market or competition is intense, competition authorities might consider allowing more collusion to increase welfare. This can be an important insight for looking at individual patent settlement cases. If it is the competition authorities' or court's rational to only prevent collusion in the individual case, they might prohibit a certain settlement in case they detect later market entry than expected under litigation. However, if they take into account that to encourage future patent challenges it might actually be justified to grant more collusion in deviation of the expected market entry under litigation, they could induce a welfare enhancing result.

With our model we firstly add an explicit argument for allowing reverse payments in patent settlements next to the already established arguments that e.g. high litigation costs, risk aversion or asymmetric information could be a justification for these payments (e.g. Willig/Bigelow 2004, Dickey/Orszag/Tyson 2010). Secondly we deliver a new perspective on the existing welfare benchmark, to only allow settlements which are not worse for consumers than the expected outcome under litigation (Shapiro criterion): To allow settlements in line with a consumer welfare standard could mean to accept a certain amount of collusion specified by our policy parameter, in clear deviation from the expected outcome under litigation, with the goal to make sure that in future more patents get challenged. *Also* in line with a consumer welfare standard, settlements perfectly in line with the Shapiro criterion could be prohibited and instead earlier generic market entry demanded.

Very important is, that we do not model the overall optimal way to incentivize the challenging of patents. For instance we could think about finding the optimal level of patent challenges by including price effects but also innovation effects (Frank/Kerber 2015). If more patents are challenged this might deteriorate innovation incentives of the originator which could yield entirely different results for our policy parameter, i.e. the policy parameter could become smaller or even negative than without taking into account innovation effects. This is because deteriorating innovation incentives for originators is a supplementary negative effect for consumer welfare besides the mentioned entry delay effect. Also taking into account patent value, as a different dimension besides patent strength, could entirely change our results, since the challenging of patents even with high challenging costs could become feasible in case high value patents are concerned. This would also raise questions which patents exactly should be challenged in terms of patent strength and patent value to maximize consumer welfare (e.g. should low value patents be challenged at all?). We assume an equal value of patents in our model.

Another important problem is whether our results for an optimal policy parameter should be regarded only on an individual case basis or can be seen as relevant for case groups, i.e. a population of patents with different patent strengths. E.g. it could be thought of using average values for our exogenous parameters based on a case group to calculate the optimal policy parameter. Although practically there occur difficult problems in terms of the specification of these parameters (e.g. average parameter values could be difficult in terms of market delineation), the idea that the policy parameter should incentivize future challenges to *unspecified patent populations* rather points at a relevance of our results for the average- or

case group perspective. Also we note that our results entail difficult competition policy issues regarding the balancing of benefits (or losses) for different groups of consumers (e.g. present- and future consumers). Difficult tradeoffs of this kind are something competition policy and also the law have tried to address before e.g. in including innovation effects in competition law.

In our paper we put forward a different perspective on the assessment of patent settlements and show that optimal incentives to challenge weak patents can be beneficial for consumers. Since the actual legislation and debate in this field, currently seems to have a different focus by e.g. being very skeptical towards large reverse payments, we believe that this result adds an important insight for policy and the law in this field.

References

- Allison, John R./Lemley, Mark A. (1998): Empirical Evidence On The Validity Of Litigated Patents, in: American Intellectual Property Law Association Quarterly Journal, Vol. 26, No. 3, pp. 185-277.
- Ayres, Ian/Klemperer, Paul (1999): Limiting Patentees' Market Power Without Reducing Innovation Incentives: The Perverse Benefits of Uncertainty and Non-Injunctive Remedies, in: Michigan Law Review, Vol. 97, pp. 985-1033.
- Court of Appeals for the Eleventh Circuit *FTC v Watson Pharmaceuticals, Inc., Solvay Pharmaceuticals, Inc.*, No. 10-12729.
- Dickey, Bret/ Orszag, Jonathan/Tyson, Laura (2010): An Economic Assessment of Patent Settlements in the Pharmaceutical Industry, in: Annals of Health Law, Volume 19, Issue 2 Winter 2010, Article 5, pp. 367-400.
- European Commission Competition DG: Case At.39226 – Lundbeck
- European Commission Competition DG: 5th Report on the Monitoring of Patent Settlements, period: January-December 2013, Published on 5 December 2014
- European Commission Competition DG: Pharmaceutical Sector Inquiry Final Report 8 July 2009.
- Federal Trade Commission (2011): FTC Report: Authorized Generic Drugs: Short-Term Effects and Long-Term Impact, August 2011.
- Frank, Severin/Kerber, Wolfgang (2015): Patent Settlements in the Pharmaceutical Industry: What Can We Learn from Economic Analysis?, in: MAGKS Joint Discussion Paper Series in Economics No. 01-2016.
- Fischmann, Filipe (2014): „Reverse Payments“ als Mittel zur Beilegung von Patentstreitigkeiten – Ein Verstoß gegen das Kartellrecht?, Inaugural-Dissertation zur Erlangung der Doktorwürde einer Hohen Juristischen Fakultät der Ludwig-Maximilians-Universität zu München.
- Grabowski, Henry G./Kyle, Margaret (2007): Generic Competition and Market Exclusivity Periods in Pharmaceuticals, in: Managerial Decision Economics, Vol. 28, pp. 491-502.
- Gratz, Linda (2012): Economic Analysis of Pay-for-delay Settlements and Their Legal Ruling, Munich Discussion Paper No. 2012-6.
- Guidelines on the application of Article 101 of the Treaty on the Functioning of the European Union to technology transfer agreements (2014/C 89/03), 28.3.2014 Official Journal of the European Union C 89/3.

- Häckner, Jonas (2000): A Note on Price and Quantity Competition in Differentiated Oligopolies, in: *Journal of Economic Theory*, Vol. 93, Issue 2, August 2000, Pages 233–239.
- Hemphill, C. Scott (2006): Paying For Delay: Pharmaceutical Patent Settlement As A Regulatory Design Problem, in: *New York University Law Review*, Vol. 81, pp. 1553-1623.
- Hovenkamp, H. (2004) Sensible Antitrust Rules For Pharmaceutical Competition, in: *University Of San Francisco Law Review*, Vol. 39, Rev. 11, 25, Fall 2004, pp. 11-32.
- Lemley, Mark A /Shapiro, Carl (2005): Probabilistic Patents, in: *Journal of Economic Perspectives*, Vol. 19, No. 2, Spring 2005, pp. 75-98.
- Lemley, Mark A. (2001): Rational Ignorance At The Patent Office, in: *Northwestern University Law Review*, Vol. 95, No. 4, pp. 1497-1532.
- Miller, Joseph Scott (2004): Building a Better Bounty: Litigation-Stage Rewards for Defeating Patents, in: *Berkeley Technology Law Journal*, Vol. 19, Issue 2, Article 3.
- Shapiro, Carl (2003): Antitrust limits to patent settlements, in: *RAND Journal of Economics*, Vol. 34, No. 2, Summer 2003, pp. 391–411.
- United States Code U.S.C. Title 21 § 355 (j).
- Willig, Robert D./Bigelow, John P. (2004): Antitrust policy toward agreements that settle patent litigation, in: 49 *Antitrust Bull.* 655 2004.
- 570 U. S. ____ (2013) *FTC v. Actavis, Inc.*

Appendix A: Proofs

Proof of Proposition 1:

Given Equation (6) we find that

$$(11) \quad \begin{aligned} & CW^{set}(\tilde{t}_{opt}) - CW^{set}(0) = \\ & \left[(CW^d - CW^t) f_g(\pi_g^d)^2 - 3(1-\lambda) \pi_g^t \left(2(CW^t - CW^d) f_g(\pi_g^d)^2 + (\varphi_2(CW^d - CW^m) f_g \right. \right. \\ & \quad \left. \left. + 2\pi_g^d ((CW^d - CW^m) f_g + (CW^m - CW^t(1-\lambda) - CW^d \lambda) \pi_g^d)) \pi_g^t \right) \right]^2 / \\ & \left[2(1-\lambda) (\pi_g^d)^2 (\pi_g^t)^2 \left(4(\varphi_1)^2 (CW^t - CW^d) (\pi_g^d)^2 + 24\varphi_1 (CW^t - CW^d) (1-\lambda) (\pi_g^d)^2 \pi_g^t \right. \right. \\ & \quad \left. \left. + 9\varphi_2 (CW^d - CW^m) (1-\lambda) (\varphi_2 + 4\pi_g^d) (\pi_g^t)^2 \right) \right]. \end{aligned}$$

Since we have that $CW^t > CW^d > CW^m$, $\lambda \in [0,1)$ as well as $\varphi_1, \varphi_2, \pi_g^t > 0$, we can show that $CW^{set}(\tilde{t}_{opt}) - CW^{set}(0) \geq 0$, because the single-rooted⁴ numerator in (11) is nonnegative, while the denominator is strictly positive. Since for $\tilde{t}_{opt} = 0$ it obviously holds that $CW^{set}(\tilde{t}_{opt}) = CW^{set}(0)$, we can conclude that $CW^{set}(\tilde{t}_{opt}) - CW^{set}(0) > 0$ for $\tilde{t}_{opt} \in (0,1]$.

Proof of Proposition 2:

Using Equation (10) we can show that for $CW^t > CW^d > CW^m$, $\lambda \in [0,1)$ as well as $\varphi_1, \varphi_2, \pi_g^d, \pi_g^t > 0$ it holds that

$$\frac{\partial \tilde{t}_{opt}}{\partial f_g} = \frac{-\frac{(CW^d - CW^m)(\varphi_2 + 2\pi_g^d)}{2(\pi_g^d)^2} - \frac{\varphi_1(CW^t - CW^d)}{3(1-\lambda)(\pi_g^t)^2} - \frac{CW^t - CW^d}{\pi_g^t}}{-\frac{\varphi_2(CW^d - CW^m)(\varphi_2 + 4\pi_g^d)}{4(\pi_g^d)^2} - \frac{(\varphi_1)^2(CW^t - CW^d)}{9(1-\lambda)(\pi_g^t)^2} - \frac{2\varphi_1(CW^t - CW^d)}{3\pi_g^t}} > 0.$$

Proof of Lemma 1:

Since $\pi_o^d + \pi_g^d > \pi_o^t + 2\pi_g^t$ and $\tilde{t} > 0$, it is easy to show that

$$\partial s_1 / \partial \lambda = \tilde{t} \left(-(\pi_o^d + \pi_g^d) + (\pi_o^t + 2\pi_g^t) \right) < 0.$$

⁴ Ibid. fn.4.

In addition, we see that λ does not enter Equations (3) and (4), so we have that $\partial \gamma_{g1}^{set} / \partial \lambda = 0$. Moreover, we know that $\partial \gamma_{g2}^{set}$ is given by Equation (5). Taking the derivative with respect to λ yields

$$\frac{\partial \gamma_{g2}^{set}}{\partial \lambda} = - \frac{\left[-\frac{1}{3} \frac{\partial s_1}{\partial \lambda} (1-\lambda) \pi_g^t - \left(f_g - \frac{s_1}{3} \right) (-\pi_g^t) \right]}{(1-\lambda)^2 (\pi_g^t)^2},$$

which is strictly smaller than zero, because $\pi_g^t > 0$, $\lambda \in [0,1)$, $f_g - \frac{s_1}{3} > 0$ and $\frac{\partial s_1}{\partial \lambda} < 0$.

Proof of Proposition 3:

By plugging Example 1's expressions for profits and welfare into Equation (10) we obtain $\tilde{t}_{opt}(\delta)$, which is given by

$$(12) \quad \begin{aligned} \tilde{t}_{opt}(\delta) = & [24(-2+\delta)^2(1+\delta)(-24(-1+\delta)^2(2+\delta-\delta^2)^4(2+\delta(4+3\delta)) \\ & -(-2+\delta)^2(1+\delta)(1+2\delta)(-320+\delta(-896+\delta(-656+3\delta(176+\delta(404 \\ & +\delta(64+\delta(-99+\delta(-43+\delta(22+(-5+\delta)(-2+\delta)\delta))))))))))f_g \\ & +(24(-2+\delta)^2(-1+\delta)^2(1+\delta)^3(12+\delta(24+\delta(9+\delta(-25-8\delta+6\delta^2))))+(1+2\delta) \\ & (-1792+\delta(-4608+\delta(-384+\delta(9088+3\delta(2176+\delta(-2688+\delta(-984 \\ & +\delta(296+\delta(1+\delta)(225+\delta(-91+\delta(-46+\delta(42+(-11+\delta)\delta))))))))))f_g)\lambda \\ & -24(-2+\delta)^2(-1+\delta^2)^3(-4+3\delta(-4-5\delta+\delta^3))\lambda^2)]/ \\ & [(-2+\delta)^4(1+\delta)^2(3776+\delta(5376+\delta(-11472+\delta(-25616+3\delta(-4+\delta(9368 \\ & +\delta(3159+\delta(-3667+3\delta(-384+\delta(246+\delta(13+\delta(-13+2\delta)))))))))))+ \\ & (-2+\delta)^2(1+\delta)(21248+\delta(25600+\delta(-91008+\delta \\ & (-172160+\delta(102080+3\delta(107136+\delta(-11720+\delta(-83224+\delta(-1709+3\delta(10824 \\ & +\delta(-701+\delta(-1942+\delta(407+\delta(52+\delta(-19+2\delta))))))))))\lambda-64(-1+\delta)^3(-4 \\ & +3\delta(-4-5\delta+\delta^3))(4+\delta(-4+3(-3+\delta)\delta(1+\delta)))(20+\delta(28+3(-3+\delta)\delta(1+\delta)))\lambda^2]. \end{aligned}$$

Based on $\tilde{t}_{opt}(\delta)$ we compute $\partial \tilde{t}_{opt} / \partial \lambda$. We omit the details, because the corresponding expression is very complex. Then we can show that for $t_{g1}^{set}, t_{g2}^{set}, \tilde{t}_{opt}, \gamma_{g1}^{set}, \gamma_{g2}^{set} \in (0,1)$ as well as for $f_g \in [\underline{f}_g, \bar{f}_g]$, $\delta \in [\underline{\delta}, 1)$, and $\lambda \in [0, \bar{\lambda}]$ we have that $\partial \tilde{t}_{opt} / \partial \lambda > 0$.⁵

⁵ Ibid. fn.4.

Proof of Lemma 2:

Using Example 1's expressions for firms' profits in Equations (4) and (5) we know that

$$\frac{\partial \gamma_{g1}^{set}}{\partial \delta} = \frac{(-2+\delta)(1+(-1+\delta)\delta)(8f_g - \tilde{t})}{4(-1+\delta)^2},$$

$$\frac{\partial \gamma_{g2}^{set}}{\partial \delta} = \frac{24\varphi_3 f_g - 2(\varphi_3 + 8(-1+\delta)^2(1+\delta+3\delta^2))\lambda}{3(-2+\delta)^3(-1+\delta)^2(1+\delta)^3(-1+\delta)},$$

where $\varphi_3 = (-2+\delta)^3(1+\delta)(1+\delta+\delta^2)$. Then we find that for $t_{g1}^{set}, t_{g2}^{set}, \tilde{t}, \gamma_{g1}^{set}, \gamma_{g2}^{set} \in (0,1)$ as well as for $f_g \in [\underline{f}_g, \bar{f}_g]$, $\delta \in [\underline{\delta}, 1)$, and $\lambda \in [0, \bar{\lambda}]$ we have that $\partial \gamma_{g1}^{set} / \partial \tilde{t} \partial \delta > 0$ and $\partial \gamma_{g2}^{set} / \partial \tilde{t} \partial \delta > 0$.

Moreover, we know from (8) and (9) that under Example 1 we have that

$$\frac{\partial \gamma_{g1}^{set}}{\partial \tilde{t} \partial \delta} = \frac{1}{4} \left(1 - \delta + \frac{1}{(-1+\delta)^2} \right),$$

$$\frac{\partial \gamma_{g2}^{set}}{\partial \tilde{t} \partial \delta} = - \frac{2\varphi_3 + 16(-1+\delta)^2(1+\delta+3\delta^2)\lambda}{3(-2+\delta)^3(-1+\delta)^2(1+\delta)^3(-1+\lambda)}.$$

We can show that for $\delta \in [0,1)$ and $\lambda \in [0,1)$ it always holds that $\partial \gamma_{g1}^{set} / \partial \tilde{t} \partial \delta > 0$ and $\partial \gamma_{g2}^{set} / \partial \tilde{t} \partial \delta > 0$.

Proof of Proposition 4:

We already know that under Example 1 the optimal policy parameter $\tilde{t}_{opt}(\delta)$ is given by Equation (12). Based on (12) we compute $\partial \tilde{t}_{opt} / \partial \delta$. Again, the details are omitted due the output's complexity. We can show that for $t_{g1}^{set}, t_{g2}^{set}, \tilde{t}_{opt}, \gamma_{g1}^{set}, \gamma_{g2}^{set} \in (0,1)$ as well as for $f_g \in [\underline{f}_g, \bar{f}_g]$, $\delta \in [\underline{\delta}, 1)$, and $\lambda \in [0, \bar{\lambda}]$ it holds that $\partial \tilde{t}_{opt} / \partial \lambda > 0$.

Appendix B: Optimal policy in case of one generic entrant

The case of one generic entrant is equivalent to the special case of our model where $\gamma_{g2}^{set} = 0$. Hence, we can immediately conclude from Equation (6) that consumer welfare is given by

$$CW^{set} = \left[(1 - \tilde{t}) \gamma_{g1}^{set}(\tilde{t}) - \frac{\gamma_{g1}^{set}(\tilde{t})^2}{2} \right] (CW^d - CW^m) + CW^m.$$

Maximizing CW^{set} with respect to \tilde{t} yields

$$\frac{\partial CW^{set}}{\partial \tilde{t}} = \underbrace{\frac{\partial \gamma_{g1}^{set}}{\partial \tilde{t}} (1 - \gamma_{g1}^{set} - \tilde{t}) (CW^d - CW^m)}_{\text{Incentive Effect - } G_1} - \underbrace{\gamma_{g1}^{set} (CW^d - CW^m)}_{\text{Entry Delay Effect - } G_1},$$

so that by respecting (8) we find that

$$\tilde{t}_{opt} = \frac{2\varphi_2 f_g + 4f_g \pi_g^d - 4(\pi_g^d)^2}{(\varphi_2)^2 + 4\varphi_2 \pi_g^d}.$$

Comparing consumer welfare under litigation and consumer welfare under the optimal policy yields

$$CW^{set}(\tilde{t}_{opt}) - CW^{set}(0) = \frac{(CW^d - CW^m)(\varphi_2 f_g + 2(f_g - \pi_g^d)\pi_g^d)^2}{2\varphi_2 (\pi_g^d)^2 (\varphi_2 + 4\pi_g^d)},$$

for which we can show that $CW^{set}(\tilde{t}_{opt}) - CW^{set}(0) \geq 0$, because $CW^d > CW^m$ and $\varphi_2, \pi_g^d > 0$.