Joint Discussion Paper
Series in Economics

by the Universities of
Aachen · Gießen · Göttingen
Kassel · Marburg · Siegen

ISSN 1867-3678

No. 02-2017

Enrico Böhme, Jonas Severin Frank and Wolfgang Kerber

Optimal Incentives for Patent Challenges in the Pharmaceutical Industry

Coordination: Bernd Hayo • Philipps-University Marburg
School of Business and Economics • Universitätsstraße 24, D-35032 Marburg
Tel: +49-6421-2823091, Fax: +49-6421-2823088, e-mail: hayo@wiwi.uni-marburg.de

MACIE PAPER SERIES

Marburg Centre for
Institutional Economics

MACIE

Nr. 2017/01

Optimal Incentives for Patent Challenges in the Pharmaceutical Industry

Enrico Böhme, Jonas Severin Frank,
Wolfgang Kerber
MACIE, Philipps-Universität Marburg

Marburg Centre for Institutional Economics • Coordination: Prof. Dr. Elisabeth Schulte
c/o Research Group Institutional Economics • Barfuessertor 2 • D-35037 Marburg

Phone: +49 (0) 6421-28-23196 • Fax: +49 (0) 6421-28-24858 •
www.uni-marburg.de/fb02/MACIE • macie@wiwi.uni-marburg.de

Philipps Universität Marburg

Optimal Incentives for Patent Challenges in the Pharmaceutical Industry

Enrico Böhme*, Jonas Severin Frank**, Wolfgang Kerber ***

Abstract

Since the patent system relies on private litigation for challenging weak patents, and patent settlements might influence the incentives for challenging patents, the question arises whether the antitrust assessment of patent settlements should also consider their impact on the incentives to challenge potentially invalid patents. 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 therefore harming consumers. In this paper we present a model that analyzes the tradeoff between limiting the delay of generic entry through patent settlements and giving generic firms more incentives for challenging weak patents of the originator firms. We show that allowing patent settlements with a later market entry of generics than the expected market entry under patent litigation can increase consumer welfare under certain conditions. 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 generics' market entry decisions.

Keywords: patent settlements, collusion, patent challenges

JEL classification: L10, L40, O34

---

* Corresponding author: Post-Doc researcher at Philipps-University Marburg, School of Business & Economics, Research Group Institutional Economics, Barfüßertor 2, D-35037 Marburg, Germany. E-mail: boehmee@staff.uni-marburg.de.

** 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 Tim Friehe, Elisabeth Schulte, Thomas Gall, Bernd Hayo, Jan Voßwinkel, Moritz Suppliet, Laura Werner, Ramsi Woodcock, and Alexandra Zaby as well as the participants of EALE 2016, the 2016 Annual Meeting of the German Economic Association, the Fifth MaCCI Annual Conference 2016 and of the 45th Hohenheimer Oberseminar 2016 for valuable comments and suggestions on earlier versions of this paper. All remaining mistakes are our responsibility.

1. Introduction

It is a well-established empirical insight that patent offices grant many patents that are later found invalid when challenged in court (Lemley/Shapiro 2005, Lemley 2001, Allison/Lemley 1998). Hence, patents do not grant an ironclad “right to exclude”, but a weaker “right to try to exclude” (Shapiro 2003, p. 395). Therefore, economists have developed the notion of "probabilistic patents", which defines the strength of a patent as the probability that a patent can be upheld in patent litigation (Lemley/Shapiro 2005). Patents which are granted erroneously lead to unjustified monopolies that harm consumers through monopoly prices. In addition, they might also block further innovation. Since society has an interest in weeding out these unjustified patents (Ayres/Klemperer 1999, Shapiro 2003, Lemley/Shapiro 2005), patent systems usually have procedures for patent opposition and patent litigation, in which other firms can sue for invalidation of patents. Since, however, the patent systems rely on private litigation for patent challenges, the question emerges whether firms have socially optimal incentives for challenging potentially invalid ("weak") patents. One specific incentive problem for challenging patents is that the challenging firm cannot internalize all the benefits for invalidating a patent, because other firms (and through more competition also the consumers) can benefit from removing an unjustified monopoly. This has been described as a public good problem of challenging patents (or as the "multiple challenger" problem), which can lead to inefficiently small challenging incentives for individual firms (Farrell/Merges 2004, Farrrell/Shapiro 2008). Obvious solutions for inefficiently small challenging incentives are spending more money for better examinations in the patent offices or granting subsidies for firms that challenge weak patents. However, the additional costs would have to be borne by the taxpayer (Miller 2004).

Another well-known problem for solving the issue of weak patents is that patent holders try to defend their unjustified patents through patent settlements. Here, they pay firms for not challenging their weak patents with the consequence of delaying market entry and price competition (pay-for-delay agreements) and thus harming consumers. Patent settlements can therefore be an instrument for undermining patent litigation as an instrument for solving the weak patent problem. In this paper, we want to analyze whether the problem of inefficiently small challenging incentives should be considered in the antitrust policy dealing with such patent settlements. Particularly in the pharmaceutical industry, patent settlements have been scrutinized critically by competition authorities both in the U.S. and the EU. Since prices for pharmaceutical products are sharply decreasing after the entry of generics, any unjustified delay of the generic firms' entry can lead to high additional health costs for consumers and society. Especially patent settlements, in which the (patent-holding) originator firms pay large sums to generic firms ("reverse payments") and agree on future entry dates of generics have been the object of competition and antitrust law proceedings. Competition authorities in the U.S. and the EU have taken action against such patent settlements in a number of cases (as, e.g., the "Lundbeck" case in the EU (Case AT.39226 – Lundbeck)). Particularly important was the ruling of the US Supreme Court in the "Actavis" case, in which it decided that patent settlements with high unexplained reverse payments can be anticompetitive and violate antitrust law (570 U. S. ____ (2013) FTC v. Actavis).

The basic arguments about the potentially anticompetitive effects of patent settlements with high reverse payments are widely accepted in the economic and legal discussion. However, there is still much controversy about the specific assessment criteria for patent settlements in the pharmaceutical industry and to what extent reverse payments may also be justified (Shapiro 2003, Willig/Bigelow 2004, Elhauge/Krüger 2012, Woodcock 2016a). In economic models, it could be shown that under specific circumstances also patent settlements with reverse payments might not harm consumers or might even be necessary for achieving efficient settlement (see Section 2). One of the concerns about a too restrictive antitrust policy against patent settlements is that prohibiting patent settlements with reverse payments (beyond litigation costs) would decrease the generics' incentives for challenging weak patents.

In this paper, we are presenting a model that analyzes the tradeoff between, on the one hand, the negative price effects of longer collusion between originator firms and generics through an agreed later generic entry (including reverse payments), and, on the other hand, the positive effects on consumer welfare through incentivizing the challenging of more weak patents by generics. Such a solution would also have the advantage that the costs of challenging more weak patents would be borne by the consumers of pharmaceutical products who benefit from the removal of unjustified patents. Following this rationale, we introduce in our model a policy parameter that stipulates the additional collusion time that competition authorities can grant to the parties of patent settlements in order to maximize consumer welfare.

In the model, we can show that there exists an optimal policy parameter which would lead to a higher consumer welfare than under the litigation solution (without a settlement), which so far has been used as a normative benchmark for the antitrust assessment of patent settlements

(Shapiro (2003)). Therefore, the consideration of challenging incentive effects can lead to a different assessment of patent settlements, and challenging incentives can be another well-founded reason for the justification of patent settlements with later generic entry and reverse payments. However, our model also proves that it depends on parameter constellations, as, e.g. the size of challenging costs, whether the optimal policy parameter is indeed positive or whether it is negative. The latter case would imply that even patent settlements without reverse payments can lead to an inefficiently long collusion period, i.e. leading to inefficiently strong challenging incentives, and therefore rendering them anticompetitive. Therefore, our results show that the consideration of challenging incentives changes the criteria for the assessment of patent settlements. However, they do not necessarily lead to the recommendation of a longer period of collusion in patent settlements and to the justification of reverse payments. For the case of a positive optimal policy parameter, we show in comparative statics analysis that this policy parameter would increase with rising challenging costs, with a longer lag between the generics' entry decisions, and with a higher intensity of competition on the market for pharmaceuticals.

The paper is structured as follows: In Section 2, we discuss, how the problem of challenging incentives has so far been considered in the context of the legal and economic discussion about the antitrust assessment of patent settlements in the pharmaceutical industry. In addition, we explain the basic idea of our approach. Sections 3 - 5 present the basic model. After explaining the model framework in Section 3, we derive the optimal policy parameter in Section 4. Afterwards, we conduct a welfare analysis in Section 5 that shows that the optimal policy parameter maximizes consumer welfare, but also that it can be positive or negative. Section 6 provides our analyses of comparative statics. In Section 7, we briefly discuss the results and potential conclusions.

2. Patent Settlements and Incentives for Patent Challenges

In the U.S., patent settlements with reverse payments in the pharmaceutical industry were challenged by the Federal Trade Commission (FTC) since 1999, because they could be collusive behavior between originators and generics for delaying market entry of generics and therefore harming consumers (FTC (2002)). The main discussion in the U.S. focused on the question whether the existence of a large reverse payment (i.e., larger than litigation costs) should be sufficient for a presumption of the illegality of a patent settlement, or whether a broader rule-of-reason approach should be applied. The latter would require a more case-specific analysis of the efficiency-enhancing and anticompetitive effects of patent settlements and therefore allow for easier justifications for patent settlements with reverse payments. After contradictory decisions by U.S. courts, the U.S. Supreme Court decided in "Actavis" that large, unexplained reverse payments in patent settlements can be a signal for the weakness of patents and therefore for the anticompetitiveness and illegality of such patent settlements (570 U. S. ____ (2013) FTC v. Actavis). However, it also insisted on a rule-of-reason approach. In the EU, the European Commission also viewed patent settlements with restrictions of market entry and reverse value transfers as potentially anticompetitive agreements that require a close scrutiny by competition law. In the meantime, the Commission has decided in several patent settlement cases with reverse payments that they are violating Art. 101 TFEU. In the recent "Lundbeck" case, the European Court has confirmed the Commission (General Court Case T-467/13). Especially relevant for our research question is that the Commission emphasized in its competition guidelines about patent settlements that society also has an interest in removing wrongly granted patents to promote competition and innovation (Guidelines on the application of Article 101 of the Treaty on the Functioning of the European Union to technology transfer agreements, p. 44).

In the economic discussion, there is a broad consensus that patent settlements with reverse payments from originators to generics can be an effective instrument for protecting weak patents against their invalidation through patent challenges. In addition, the criterion of Shapiro (2003), which states that patent settlements should not lead to lower consumer welfare than under patent litigation, has been broadly accepted in the economic and legal discussion as a relevant normative criterion for the antitrust assessment of patent settlements. In the basic settlement model about the price effects of patent settlements on consumer welfare, it has been shown that if reverse payments are larger than litigation costs, patent settlements lead to an entry date of generics that is delayed beyond the expected entry date under litigation, i.e. compared to the benchmark of the litigation solution, consumers are harmed (Shapiro (2003), Elhauge/Krüger (2012), Frank/Kerber (2016)). This result strongly supports the concerns about reverse payments. However, subsequent economic analyses have also clarified that under less simple and more realistic assumptions about the negotiation situation (e.g., information asymmetries, risk aversion, multiple entrants etc.), the conclusions about reverse payments are less clear. On the one hand, there might be economically well-founded reasons for why patent settlements with large reverse payments might also be efficient and not harm consumers. On the other hand, patent settlements without reverse payments can also be anticompetitive. Therefore, the existence of reverse payments might not be such a clear criterion for identifying anticompetitive patent settlements, and hence in most cases a deeper investigation might be necessary (Willig/Bigelow (2004), Dickey et al. (2010)).

A serious problem of the current discussion is that it has so far almost exclusively focused on the effects of patent settlements on consumer welfare via price effects, i.e. on the question, whether patent settlements lead to an inefficiently late generic entry and therefore to an inefficiently late start of price competition through generics. However, patent settlements can also affect consumer welfare through two additional channels (Frank/Kerber (2016)). Since antitrust limits for patent settlements also influence the profits of the originators, the question emerges whether these limits to patent settlements might have negative effects on the innovation incentives for originators and therefore might harm consumers through the development of fewer pharmaceuticals. So far, only Elhauge/Krüger (2012) and Woodcock (2016a, 2016b) have taken this innovation incentive effect into account when analyzing the effects of patent settlements. Although we should be cautious about too far-reaching conclusions from their models, their results suggest that competition authorities don't have to worry too much about negative effects on innovation incentives. The second additional channel is the effect of antitrust limits of patent settlements on the challenging incentives for generics. For example, in the above-mentioned "Actavis" ruling of the U.S. Supreme Court, Chief Justice Roberts states in his dissenting opinion 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 addition, other scholars are also concerned that a too restrictive policy in regard to the possibilities of generics to make profits from patent settlements through reverse payments would decrease their incentives to challenge patents (e.g. Dickey et al. (2010, p. 399)).

These concerns about the generics' challenging incentives suggest that there might be a tradeoff between limiting the delay of generic entry through patent settlements for ensuring early price competition and offering generics more incentives for challenging additional weak patents of originators. In our model, we want to study this trade off and analyze whether consumer welfare increases, if competition authorities would grant an additional period of collusion (beyond the expected entry date of the litigation solution) in order to incentivize the challenging of additional patents. The economic intuition is, that the additional costs for consumers of pharmaceutical products from a later generic entry (in form of a later decrease in prices) might be smaller than their additional benefits from the challenging of more weak patents, which otherwise would allow their owners monopoly profits until their expiration.

The structure of our model is partly based upon the model of Gratz (2012), who was the first to offer an integrated analysis of price and challenging incentive effects in regard to 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 erroneously uphold anticompetitive patent settlements. This increases joint settlement profits and hence the challenging incentives for generics. Therefore, in her model judicial errors (due to information problems in the application of the rule-of-reason) lead to overall positive effects on consumer welfare through erroneously allowing later generic entry than in the benchmark litigation solution. In contrast to her approach, we introduce a policy parameter and analyze the optimal additional delay (compared to the litigation solution) that would maximize consumer welfare if challenging incentives are also taken into account.[1] As in the model of Gratz (2012), we also suppose the existence of two generic entrants that can challenge the patent and sequentially enter the market at two endogenously determined dates with an exogenously given lag between the entry decision of the first and the second generic. Therefore, we do not model the public good problem that emerges if several entrants could simultaneously challenge a patent (multiple challenger problem)[2] as well as we do not model strategic interaction between generics.

3. Model Framework

Our model follows the basic framework introduced by Gratz (2012). In this framework, an originator (O) holds a patent with remaining patent duration $1-t$, where $t \in [0,1]$, and patent strength γ, which is a random variable following a continuous uniform distribution over the unit interval, i.e. $\gamma \sim U[0,1]$. Hence, we can interpret γ as reflecting the probability that the patent is found valid in court. It is assumed that all patents have the same value and that the realizations of γ are common knowledge. Two generics (G_1 and G_2) are potential entrants in

[1] Compared to the policy parameter in our model, the errors which courts make in applying the rule-of-reason in the model of Gratz (2012) have a different effect, because this error always leads to higher challenging incentives for generics, while our policy parameter for additional collusion can be positive or negative.

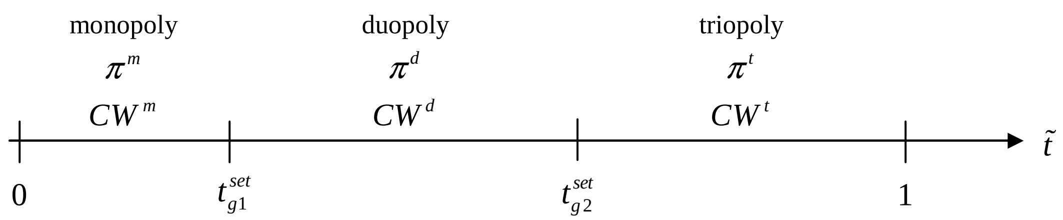
[2] In the U.S., the incentive problems resulting from the public good problem were addressed by the Hatch-Waxman Act, which grants the first challenger a 180 days marketing exclusivity of generic entry, before other entrants are allowed to enter the market (FTC (2011, p. 138), Hemphill (2006)). Therefore, sequential entry of generics has some similarities to the consequences of this U.S. regulation, although we do not intend to model the specific conditions of the U.S. Hatch-Waxman pharmaceutical sector regulation.

the market by challenging the patent at a specific time with challenge costs f_g. G_1 and G_2 make their entry decisions in a fixed sequence and, depending on the situation, it is possible that both enter, only G_1 enters, or none enters. The generics' challenging decisions are based on the rationale whether the generated net profits from such a patent challenge are greater than zero. It is assumed that firms maintain the option to litigate and hence to reap, in addition to the settlement surplus, their expected litigation profits. G_1 decides on challenging the patent at $t=0$, whereas G_2 decides to challenge at $t=\lambda$, where $\lambda \in (0,1)$. If the patent is declared invalid, entry by the first generic would occur at $t=0$ and entry of the second generic at $t=\lambda$. In case the patent is held valid, entry by both generics would occur at $t=1$. This implies that G_1's expected entry date under litigation is $t_{g1}^{lit} = \gamma \cdot 1 + (1-\gamma)\cdot 0 = \gamma$, while G_2's expected entry date is $t_{g2}^{lit} = \gamma \cdot 1 + (1-\gamma)\lambda > t_{g1}^{l}$. Hence, G_1 always enters prior to G_2.

If a generic challenges a patent, the originator and the generic can settle their patent dispute through a patent settlement with the agreement on a specific future entry date and a potential payment from the originator to the generic (reverse payment). We assume that the settling firms equally share the settlement surplus.

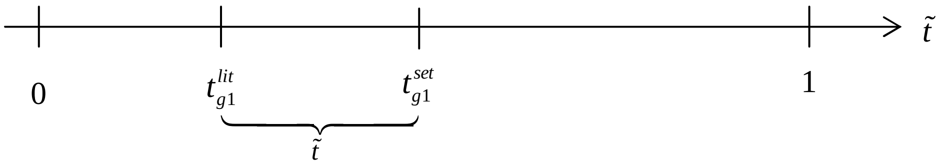
The parties' actual entry dates under a settlement, t_{g1}^{set} and t_{g2}^{set}, are endogenous with $t_{g1}^{set} \in [0,1]$ and $t_{g2}^{set} \in [\lambda,1]$. Since market entry occurs sequentially, the originator can reap monopoly profits π^m for the period $[0, t_{g1}^{set})$, whereas for $[t_{g1}^{set}, t_{g2}^{set})$, 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. *Fig. 1* summarizes this situation.

Figure 1: Market structure, profits, and consumer welfare in the settlement solutions



According to the criterion of Shapiro, competition authorities would prohibit all patent settlements through which consumers are harmed by generic entries that are later than expected under litigation. This implies that the firms are not allowed to specify entry dates beyond $t_{g1}^{lit} = \gamma$ and $t_{g2}^{lit} = \gamma + (1-\gamma)\lambda$. However, deviating from Gratz (2012), we introduce a policy parameter $\tilde{t}$, which explicitly allows the competition authorities 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 (certain) entry date of G_1 under this policy parameter. Since the original remaining patent duration, i.e. the remaining duration for $t = 0$, is equal to one, $\tilde{t}$ can be interpreted as a percentage share of the original remaining patent duration that is *additionally* granted for collusion. We illustrate this situation in *Fig. 2*.

Figure 2: Policy parameter $\tilde{t}$ as a share of the original remaining patent duration



4. Equilibrium Analysis: The Optimal Policy Parameter

Since competition authorities will not challenge settlements with $t_{g1}^{set} \le t_l + \tilde{t}$, O and G_1 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 $t_{g2}^{set} = t_{g1}^{set} + \left(1 - t_{g1}^{set}\right)\lambda$. Hence, the joint settlement profits under $\tilde{t}$ are given by

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

where γ_{g1}^{set} and γ_{g2}^{set} denote the critical levels of patent strength, for which generic companies are indifferent between challenging a patent or not. Therefore, the generated surplus compared to litigation is

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

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

which allows us to determine the critical levels of patent strength. Respecting that the relevant expected litigation profits for G_1 and G_2 are given by $\pi_{g1}^{lit}(\gamma) = \left(1 - t_{g1}^{lit}\right)\pi_g^d = (1-\gamma)\pi_g^d$ and $\pi_{g2}^{lit}(\gamma) = \left(1 - t_{g2}^{lit}\right)\pi_g^t = (1-\gamma)(1-\lambda)\pi_g^t$, we find that γ_{g1}^{set} and γ_{g2}^{set} are determined by

(4) $$\pi_{g1}^{lit}(\gamma)+\frac{s_2}{2}-f_g=0 \quad \Leftrightarrow \quad \gamma_{g1}^{set}=1-\frac{f_g-\frac{s_2}{2}}{\pi_g^d},$$

(5) $$\pi_{g2}^{lit}(\gamma)+\frac{s_1}{3}-f_g=0 \quad \Leftrightarrow \quad \gamma_{g2}^{set}=1-\frac{f_g-\frac{s_1}{3}}{(1-\lambda)\pi_g^t}<\gamma_{g1}^{set}.$$

Given Equations (4) and (5) we can conclude that G_1 challenges patents for $\gamma\in\left[0,\gamma_{g1}^{set}\right]$, while G_2 challenges for $\gamma\in\left[0,\gamma_{g2}^{set}\right]$, which leads to the following market structures: If neither G_1 nor G_2 challenges any patents, i.e. for $\gamma\in\left(\gamma_{g1}^{set},1\right]$, the originator's monopoly covers the entire remaining patent duration. For $\gamma\in\left(\gamma_{g2}^{set},\gamma_{g1}^{set}\right]$ 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 $1-t_{g1}^{set}=1-\gamma-\tilde{t}$. However, for $\gamma\in\left[0,\gamma_{g2}^{set}\right]$ we find that G_2 additionally enters the market. Therefore, monopoly lasts for $t_{g1}^{set}-0=\gamma+\tilde{t}$, duopoly for $t_{g2}^{set}-t_{g1}^{set}=\left(1-\gamma-\tilde{t}\right)\lambda$, and triopoly for $1-t_{g2}^{set}=\left(1-\gamma-\tilde{t}\right)(1-\lambda)$. Hence, respecting CW^m, CW^d, and CW^t, we know that consumer welfare under $\tilde{t}$ is described by

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

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

(7) $$\begin{aligned}\frac{\partial CW^{set}}{\partial\tilde{t}}=&\underbrace{\frac{\partial\gamma_{g1}^{set}}{\partial\tilde{t}}\left(1-\gamma_{g1}^{set}-\tilde{t}\right)\left(CW^d-CW^m\right)}_{\text{Incentive Effect - G}_1}-\underbrace{\gamma_{g1}^{set}\left(CW^d-CW^m\right)}_{\text{Entry Delay Effect - G}_1}\\&+\underbrace{\frac{\partial\gamma_{g2}^{set}}{\partial\tilde{t}}\left(1-\gamma_{g2}^{set}-\tilde{t}\right)(1-\lambda)\left(CW^t-CW^d\right)}_{\text{Incentive Effect - }G_2}-\underbrace{\gamma_{g2}^{set}(1-\lambda)\left(CW^t-CW^d\right)}_{\text{Entry Delay Effect - G}_2}.\end{aligned}$$

The interpretation of (7) is straightforward: If competition authorities are 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 $\left(CW^d - CW^m\right)$ for the period after market entry under the marginal settlement, i.e. for $1-t_{g1}^{set} = 1-\gamma_{g1}^{set} - \tilde{t}$. However, due to G_1's delayed market entry, consumer welfare decreases by $\left(CW^d - CW^m\right)$ 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

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

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

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

where $\varphi_1 = \pi^m - \lambda\left(\pi_o^d + \pi_g^d\right) - (1-\lambda)\left(\pi_o^t + 2\pi_g^t\right)$ and $\varphi_2 = \pi^m - \left(\pi_o^d + \pi_g^d\right)$. 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

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

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

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 of the litigation solution. Using Equation (6) we can easily compute consumer surplus under litigation by evaluating $CW^{set}\left(\tilde{t}\right)$ at $\tilde{t}=0$. Hence, consumer welfare under litigation is given by $CW^{set}\left(0\right)$, whereas consumer welfare under the optimal policy $\tilde{t}_{opt}$ is determined by $CW^{set}\left(\tilde{t}_{opt}\right)$. Our results are summarized in Proposition 1:

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

Proof: See Appendix A.

Proposition 1 shows that the implementation of $\tilde{t}_{opt}$ is indeed welfare increasing for consumers. 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 $\tilde{t}_{opt}<0$. In Appendix B, we show that Proposition 1 also holds for the case of one generic entrant. Therefore, our results do not depend on the specific assumption of two generic entrants. 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.

Given the paper's focus, it is also a very important result that, depending on parameter constellations, the optimal policy parameter $\tilde{t}_{opt}$ can be positive or negative. However, firms would never accept any settlement profits that are lower than profits under litigation. Therefore, a requirement of competition authorities that the entry date in a settlement should be prior to the expected entry date under litigation would make settlements impossible and lead to the litigation solution. In addition, market entry cannot be delayed beyond the original patent duration, i.e., $\tilde{t}_{opt}\leq 1$. Hence, in our model framework feasible solutions for patent settlements require $\tilde{t}_{opt}\in\left[0,1\right]$, which holds for $f_g\in\left[\underline{f}_g,\overline{f}_g\right]$, where

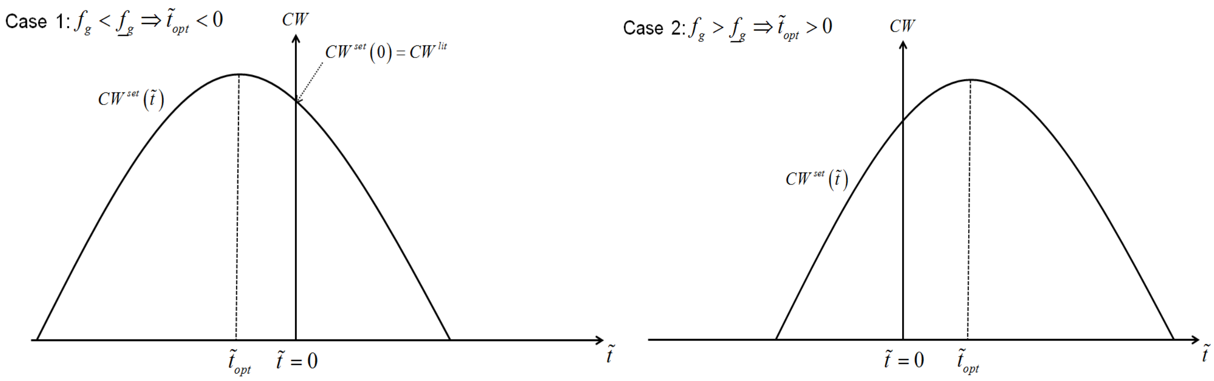
$$\underline{f}_g=\left[-6(1-\lambda)\left(CW^m-CW^t(1-\lambda)-CW^d\lambda\right)\left(\pi_g^d\right)^2\left(\pi_g^t\right)^2\right]\Big/\Big[2\varphi_1\left(CW^t-CW^d\right)\left(\pi_g^d\right)^2$$
$$-3(1-\lambda)\pi_g^t\left(2\left(CW^d-CW^t\right)\left(\pi_g^d\right)^2-\left(CW^d-CW^m\right)\left(\varphi_2+2\pi_g^d\right)\pi_g^t\right)\Big],$$

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

Note that we can show that $0 < \underline{f}_g < \overline{f}_g$, 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 $\tilde{t}_{opt} < 0$, i.e., firms would strictly prefer to go for litigation. If challenge costs are above the upper threshold, we would have $\tilde{t}_{opt} > 1$. Hence, we end up with the corner solution where $\tilde{t}_{opt} = 1$.

Our finding that $\tilde{t}_{opt}$ can be positive or negative implies that we have to distinguish two cases (see *Fig. 3*). In Case 1 ($f_g < \underline{f}_g$), consumer welfare would be maximized for $\tilde{t}_{opt} < 0$. Economically, this would imply that the optimal date of generic entry is even prior to the expected entry date under litigation. Here, the challenging incentives provided in the litigation solution are inefficiently high. Hence, reducing challenging incentives would be welfare-increasing for consumers. In addition, we can conclude that all settlements that specify an entry date beyond the one under litigation are clearly anticompetitive. However, as we have seen above, as long as the law does not limit the parties' rights to litigate, competition authorities cannot enforce earlier market entry than under litigation. In Case 2 ($f_g > \underline{f}_g$), the situation is entirely different. If $\tilde{t}_{opt} > 0$, challenging incentives under litigation are inefficiently small. Therefore, allowing for a longer period of collusion benefits consumers. Since generic entry delay is only possible if the generics get a share of the settlement surplus, competition authorities have to allow reverse payments to a certain extent. Hence, our analysis shows that allowing for more collusion in patent settlements (with reverse payments) can be procompetitive. However, this only holds for Case 2, i.e. for $\tilde{t}_{opt} > 0$.

Figure 3: Impact of $\tilde{t}_{opt}$ on consumer welfare



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 $t_{g2}^{set} = t_{g1}^{set} + \left(1 - t_{g1}^{set}\right)\lambda$ 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, the critical level of patent strength under such a settlement is larger than under litigation, i.e. $\gamma_{g1}^{set} > \gamma_{g1}^{lit}$ and $\gamma_{g2}^{set} > \gamma_{g2}^{lit}$ (see *Fig. 4*).

Figure 4: Entry delay and incentive effects under $\tilde{t}_{opt} > 0$

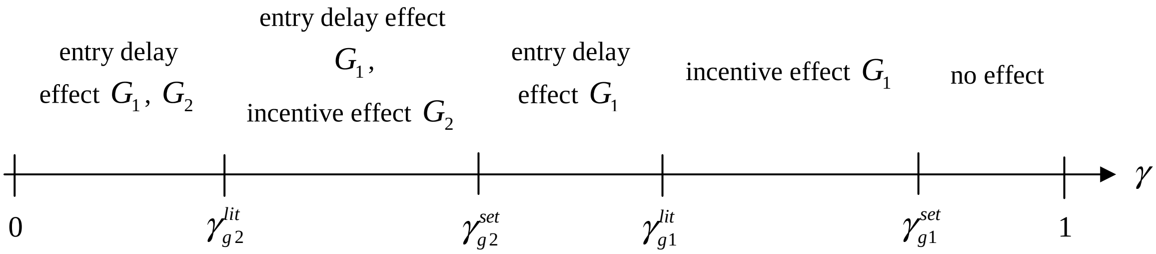


Fig. 4 refers to the case of $\tilde{t}_{opt} > 0$ and depicts entry delay and incentive effects for different intervals of patent strength. These intervals are bound by the critical levels of patent strength under litigation as well as under $\tilde{t}_{opt}$. Thus, we can derive for which patent strengths incentive and entry delay effects are relevant and to which extent they affect consumer welfare. As we can see from *Fig. 4*, G_2 would challenge γ_{g2}^{lit} patents under litigation, while G_1 challenges $\gamma_{g1}^{lit} > \gamma_{g2}^{lit}$ patents. Hence, G_1 strictly challenges more patents than G_2. This result is a consequence from the assumption of sequential entry decisions of the parties: Since G_1 always decides first about entry, it can reap more profits and 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 *Fig. 4* that for $\gamma \in \left[0, \gamma_{g2}^{lit}\right]$, both generics would challenge these patents under litigation anyway. Hence, in this interval the implementation of $\tilde{t}_{opt}$ induces an entry delay effect of G_1 and G_2, which negatively affects consumer welfare. For $\gamma \in \left(\gamma_{g2}^{lit}, \gamma_{g2}^{set}\right]$, we find that G_2 additionally challenges patents that would not have been challenged under litigation, which describes G_2's incentive effect. However, this patent range would still have been challenged by G_1 under litigation, i.e. we have an additional entry delay effect for G_1. Hence, for $\gamma \in \left(\gamma_{g2}^{lit}, \gamma_{g2}^{set}\right]$, we find two effects, which affect consumer welfare in opposite directions. For $\gamma \in \left(\gamma_{g2}^{set}, \gamma_{g1}^{lit}\right]$, G_1's entry delay effect is still present, while there is no effect resulting from G_2, since she does not challenge these (or stronger) patents. For $\gamma \in \left(\gamma_{g1}^{lit}, \gamma_{g1}^{set}\right]$, we have that G_1 additionally challenges patents 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 $\gamma \in \left(\gamma_{g1}^{set}, 1\right]$, we do not observe any effects, because these patents are neither challenged under litigation nor under $\tilde{t}_{opt}$.

6. Comparative Statics

In this chapter, we analyze the comparative statics of the optimal policy parameter $\tilde{t}_{opt}$. Since we have found that $\tilde{t}_{opt} < 0$ strictly results in litigation, we restrict our analysis to Case 2, i.e. to the case of $\tilde{t}_{opt} > 0$. We will analyze how $\tilde{t}_{opt}$ is changing if i) the challenging costs f_g increase, ii) the lag between the generics' entry decisions increases, and iii) the intensity of competition on the pharmaceutical market increases. For the analyses ii) and iii), we introduce a specific competition model.

The Effects of an Increase in Challenging Costs

As we already know that challenge costs have a crucial impact on the optimal policy, we examine at first, how $\tilde{t}_{opt}$ changes in f_g. The challenge costs include costs for preparing the challenge as well as firm investments for entering the market, i.e. technology and marketing investments to overcome market entry barriers. 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 find by using (8) and (9) that a marginal increase in f_g does not 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 patent challenges more costly. Hence, ceteris paribus firms would 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 $\left(\partial\gamma_{g1}^{set} / \partial\tilde{t}\right)\left(1-\gamma_{g1}^{set}-\tilde{t}\right)\left(CW^d - CW^m\right)$, which describes the incentive effect of G_1. If f_g increases, this incentive effect is stronger, because γ_{g1}^{set} decreases. The interpretation is intuitive: Since γ_{g1}^{set} decreases, $t_{g1}^{set}\left(\gamma_{g1}^{set}\right) = \gamma_{g1}^{set} + \tilde{t}$ decreases as well, i.e. firms would ceteris paribus agree on earlier market entry in the marginal 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 stronger impact on CW^{set} if f_g marginally increases. The same logic applies to G_2's incentive effect, where a marginal increase in $\tilde{t}$ positively affects CW^{set} by $\left(\partial\gamma_{g2}^{set}/\partial\tilde{t}\right)\left(1-\gamma_{g2}^{set}-\tilde{t}\right)(1-\lambda)\left(CW^t-CW^d\right)$. This incentive effect for G_2 is also stronger if 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 find that because of G_1's delayed entry, consumer welfare decreases by CW^d-CW^m for all $\gamma\in\left[0,\gamma_{g1}^{set}\right]$. In case of G_2, we have that consumer welfare marginally decreases by $(1-\lambda)\left(CW^t-CW^d\right)$ for all $\gamma\in\left[0,\gamma_{g2}^{set}\right]$. As we have already seen, a marginal increase in f_g leads to lower values for the critical patent strengths γ_{g1}^{set} and γ_{g2}^{set}. Hence, the entry delay effects $\gamma_{g1}^{set}\left(CW^d-CW^m\right)$ for G_1 and $\gamma_{g2}^{set}(1-\lambda)\left(CW^t-CW^d\right)$ for G_2, respectively, strictly decrease, because the range of patents, which is challenged anyway, ceteris paribus decreases. This outcome is intuitive: The entry delay effect for consumer welfare is based on a larger extent of collusion between $\left[0,\gamma_{g1}^{set}\right]$ and $\left[0,\gamma_{g2}^{set}\right]$. 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^{set}, a marginal increase in f_g leads to an increase in the generics' incentive effects, whereas entry delay effects for both generics 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

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

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

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

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

If we focus on the case under consideration, i.e. on the case of $n \in \{1,2,3\}$, we finally obtain $\pi^m = \frac{1}{4}$, $\pi_o^d = \pi_g^d = (1-\delta) \Big/ \left[(2-\delta)^2 (1+\delta)\right]$, $\pi_o^t = \pi_g^t = \left(1-\delta^2\right) \Big/ \left[4(2\delta+1)\right]$ as well as $CW^m = \frac{1}{8}$, $CW^d = 1 \Big/ \left[(1+\delta)(-2+\delta)^2\right]$, $CW^t = \left[3(1+\delta)\left(1+3\delta+2\delta^2\right)\right] \Big/ \left[8(1+2\delta)^2\right]$. 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 $f_g \in \left[\underline{f}_g, \overline{f}_g\right]$, $\delta \in [\underline{\delta}, 1)$, and $\lambda \in \left[0, \overline{\lambda}\right]$, where $\underline{\delta} \approx 0{,}7478$ and

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

$$-12\delta^{11}+93\delta^{12}-30\delta^{13}+3\delta^{14}\Big).$$

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 an Increase in G_2's Entry Decision Date

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 can be interpreted as reflecting 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: First, the settlement surplus s_1 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 G_2's actual entry date, $t_{g2}^{set}=\gamma+\tilde{t}+(1-\gamma-\tilde{t})\lambda$, 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 stronger under litigation. Therefore, the surplus decreases in λ. Secondly, it is easy to see that the denominator in $\gamma_{g2}^{set}=1-\left(f_g-(s_1/3)\right)/(1-\lambda)\pi_g^t$ 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 $\gamma_{g2}^{set}(1-\lambda)\left(CW^{t}-CW^{d}\right)$, 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 $\left(\partial \gamma_{g2}^{set} / \partial \tilde{t}\right)\left(1-\gamma_{g2}^{set}-\tilde{t}\right)(1-\lambda)\left(CW^{t}-CW^{d}\right)$ 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 $\left(1-\gamma_{g2}^{set}-\tilde{t}\right)$ 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 date for γ_{g2}^{set}, which is given by $t_{g2}^{set}\left(\gamma_{g2}^{set}\right) = \gamma_{g2}^{set} + \tilde{t} + \left(1-\gamma_{g2}^{set}-\tilde{t}\right)\lambda$. As we know from Lemma 1 that γ_{g2}^{set} is decreasing in λ, the overall effect on $t_{g2}^{set}\left(\gamma_{g2}^{set}\right)$ 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 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 $\partial\left[\left(\partial\gamma_{g2}^{set}/\partial\tilde{t}\right)\left(1-\gamma_{g2}^{set}-\tilde{t}\right)\left(1-\lambda\right)\right]/\partial\lambda>0$,[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.

The Effects of an Increase in the 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' products, it can 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 γ_{g1}^{set} and γ_{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. Since 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 $\pi_o^t(\delta)$, $\pi_g^t(\delta)$ decrease, while consumer welfare under duopoly and triopoly, i.e. $CW^d(\delta)$ and $CW^t(\delta)$, 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 $\gamma_{g1}^{set}(\delta)=1-\left(f_g-\left(s_2(\delta)/2\right)\right)/\pi_g^d(\delta)$. 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

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

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 the result of 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 consider 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 benefit of higher consumer welfare through additional patent challenges is realized for a longer period of time, since G_1's market entry under the marginal settlement 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 our model, 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. Discussion and Conclusion

In the model, we have analyzed the tradeoff between entry delay- and incentive effects in patent settlements between originators and generics. By introducing a policy parameter, we explicitly allow in antitrust law for a longer period of collusion, which negatively affects consumer welfare. However, the provision of more challenging incentives for potentially unjustified (weak) patents has a positive impact on consumer welfare. We show under very general conditions that there exists an optimal specification of the policy parameter, and our key result is that consumer welfare under this optimal policy parameter is higher than under the benchmark case of litigation.

However, it is also a crucial result of our model that, depending on parameter constellations (as, e.g., challenging costs), this optimal policy parameter can be positive or negative. If the optimal policy parameter is positive, limiting the collusion of originators and generics through antitrust law to the expected entry date of the litigation solution would lead to inefficiently small challenging incentives for generics. Hence, a longer collusion period would be beneficial from a consumer welfare perspective. Therefore, another implication of our model is that the consideration of challenging incentives for generics can provide an additional reason why patent settlements with reverse payments might not harm consumers and are therefore not anticompetitive. In addition, we have shown for this case of $\tilde{t}_{opt} > 0$ that (under certain assumptions) the optimal policy parameter increases with the size of challenging costs, the lag of the second generic's market entry decision, and with the intensity of competition between originators and generics after generic entry.

If, however, the optimal policy parameter $\tilde{t}_{opt}$ is negative (i.e. challenging costs are sufficiently low, $f_g < \underline{f}_g$, see Case 1 in *Fig. 3*), then the expected entry date under litigation would be inefficiently late, i.e. the marginal challenging incentives provided in the litigation benchmark solution would be too large. However, claiming from an antitrust perspective that the rights of the parties to litigate should be limited, might lead to a serious conflict between antitrust and patent law. This possible case of a negative policy parameter shows nonetheless that taking into account challenging incentives can lead to the conclusion that patent settlements without reverse payments can also harm consumers and be anticompetitive.

What conclusions can be drawn from our results? In a comprehensive analysis of the economic literature about patent settlements, Frank/Kerber (2016) have shown that a correct antitrust analysis of the effects of patent settlements requires a deep understanding of the interaction between patent and antitrust law, and therefore an integrated economic analysis of the effects of patent settlements on both innovation and competition. Therefore, the current main focus on price effects is not sufficient and, in addition, the effects on innovation incentives and challenging incentives have to be analyzed. Our model provides an integrated analysis of price effects and challenging incentives and shows that a tradeoff between both can exist, which is relevant for the antitrust assessment of patent settlements. Regarding the analysis of price effects and innovation incentive effects, results of the research by Elhauge/Krüger (2012) and Woodcock (2016a/2016b) seem to suggest that in this case there might be no tradeoff, i.e. an antitrust policy that uses the expected entry under litigation as a benchmark would not lead to inefficiently small innovation incentives.[4] Therefore, innovation incentive arguments might not support claims for allowing longer collusion periods (and therefore reverse payments). Models that try to simultaneously analyze all three effects of antitrust limits for patent settlements on consumer welfare are still missing.

[4] See also Frank/Kerber (2016, pp.12) where a critical analysis of the model of Elhauge/Krüger (2012) can be found, and, from a much broader perspective, the discussion of the proportionality principle in regard to probabilistic patents for innovation incentives in Farrell/Shapiro (2008) and Encaoua/Lefouili (2009).

Another question is whether the results of our model do also hold under less restrictive assumptions. For example, we assumed that i) all patents have the same value, ii) there is a uniform distribution of patent strength in the continuum of all patents, and iii) the challenging costs are the same for all patents. Under more relaxed assumptions, our specific results and the optimal policy parameter would certainly change, but it is not clear why this should change our main results about the existence of an optimal policy parameter and the possibility of a tradeoff with its conclusion that a longer collusion might be justified. Another assumption in our model is that originators and generics know the true patent strength. In previous research, the consequences of wrong and/or asymmetric beliefs about the strength of patents have been analyzed. Allowing for such more realistic assumptions can change the settlement ranges in different ways, and can lead to a number of problems with respect to achieving efficient settlement solutions. Although this makes the antitrust assessment considerably more difficult and would presumably influence the size of the optimal policy parameter, we do not expect our results to change qualitatively. Another specific assumption is the existence of two generic entrants that make sequential entry decisions with an exogenously assumed lag between first and second generic. Since we have shown that the results also hold for one entrant (Appendix B), they do not depend on the specific assumption of two entrants. In addition, they presumably also hold for more than two entrants, as long as a similar structure of sequential entry decisions is assumed. However, we do not model the multiple challenger problem, i.e. that there might be several generic entrants that simultaneously decide on challenging patents with the ensuing public good problem. So far, we are not aware of a model that includes this public good problem in the economic analysis of the antitrust treatment of patent settlements.[5] It is clear that a more explicit analysis of the effects of the interaction and competition between potential generic entrants is one of the important gaps in previous research. Therefore, there are still a large number of important questions for future economic research.

What conclusions can be drawn for the initial question of how to deal with the weak patent problem? Taking challenging incentives into account in the antitrust assessment does not change the necessity of a very critical antitrust scrutiny of patent settlements with reverse payments for eliminating an easy way for originators to protect potentially unjustified and weak patents that can harm consumers. The consideration of challenging incentives can help

[5] Some authors have discussed scenarios of multiple generic entries in specific frameworks. However, they have not analyzed the public good problem (Edlin et al. (2015), Kobayashi et al. (2015)).

to optimize the antitrust assessment of patent settlements, but it is also clear that this can only contribute to a limited extent to the solution of the weak patent problem. Therefore, it is still necessary to search for other solutions as part of the general problem of the optimal patent system's design. Important proposals about improving patent examination procedures in patent offices, facilitating and strengthening of patent opposition and patent litigation procedures, e.g. by joint challenges, and promoting subsidization of patent challenges can be found in Miller (2004), Farrell/Shapiro (2008), and Encaoua/Lefouili (2009).

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 Benefit s of Uncertainty and Non-Injunctive Remedies, in: Michigan Law Review, Vol. 97, pp. 985-1033.

Case AT.39226 – Lundbeck.

Consolidated Version of The Treaty On The Functioning Of The European Union - Article 101 (Ex Article 81 Tec), Official Journal 115 , 09/05/2008 P. 0088 – 0089.

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.

Edlin, Aaron/Hemphill, Scott/ Hovenkamp, Herbert, Shapiro, Carl (2015): The Actavis Inference: Theory And Practice, in: Rutgers University Law Review, Vol. 67, No.3, pp. 1-51.

Elhauge, Einer/Krüger, Alex (2012): Solving the Patent Settlement Puzzle, in: Texas Law Review, Vol. 91, pp. 283-330.

Encaoua, David/Lefouili, Yassine (2009): Licensing weak patents, in: Journal of Industrial Economics, Wiley-Blackwell, 2009, Vol. 57, No. 3, pp. 492-525.

Farrell, Joseph/Shapiro, Carl (2008): How Strong Are Weak Patents?, in: American Economic Review 2008, Vol. 98, No. 4, pp. 1347–1369.

Farrell, Joseph/Merges, Robert P. (2004): Incentives To Challenge And Defend Patents: Why Litigation Won't Reliably Fix Patent Office Errors And Why Administrative Patent Review Might Help, in: Berkeley Technology Law Journal, Vol. 19, pp. 943-970.

Federal Trade Commission (2011): FTC Report: Authorized Generic Drugs: Short-Term Effects and Long-Term Impact, August 2011.

Federal Trade Commission (2002): Generic Drug Entry Prior to Patent Expiration: An FTC Study, July 2002.

Frank, Severin/Kerber, Wolfgang (2016): Patent Settlements in the Pharmaceutical Industry: What Can We Learn from Economic Analysis?, in: MAGKS Joint Discussion Paper Series in Economics No. 01-2016.

General Court Case T-467/13.

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.

Kobayashi, Bruce H./Wright, Joshua D./Ginsburg, Douglas H./Tsai, Joanna (2015): Actavis And Multiple Anda Entrants: Beyond The Temporary Duopoly, in: Antitrust, Vol. 29, No. 2, pp. 89-97.

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.

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.

Willig, Robert D./Bigelow, John P. (2004): Antitrust policy toward agreements that settle patent litigation, in: The Antitrust Bulletin, Vol. 49, Fall 2004, pp. 655-698.

Woodcock, Ramsi A. (2016a): Innovation, Litigation, and New Drugs retrieved from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2702474 (December 11th 2016).

Woodcock, Ramsi A. (2016b): Innovation and Reverse Payments, retrieved from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2818470&download=yes## (November 29th 2016).

Appendix A: Proofs

Proof of Proposition 1:

Given Equation (6) we find that

(11) $$CW^{set}\left(\tilde{t}_{opt}\right)-CW^{set}\left(0\right)=$$

$$\Big[\left(CW^{d}-CW^{t}\right)f_g\left(\pi_g^d\right)^2-3\left(1-\lambda\right)\pi_g^t\Big(2\left(CW^{t}-CW^{d}\right)f_g\left(\pi_g^d\right)^2+\Big(\varphi_2\left(CW^{d}-CW^{m}\right)f_g$$
$$+2\pi_g^d\left(\left(CW^{d}-CW^{m}\right)f_g+\left(CW^{m}-CW^{t}\left(1-\lambda\right)-CW^{d}\lambda\right)\pi_g^d\right)\Big)\pi_g^t\Big)\Big]^2\Big/$$
$$\Big[2\left(1-\lambda\right)\left(\pi_g^d\right)^2\left(\pi_g^t\right)^2\Big(4\left(\varphi_1\right)^2\left(CW^{t}-CW^{d}\right)\left(\pi_g^d\right)^2+24\varphi_1\left(CW^{t}-CW^{d}\right)\left(1-\lambda\right)\left(\pi_g^d\right)^2\pi_g^t$$
$$+9\varphi_2\left(CW^{d}-CW^{m}\right)\left(1-\lambda\right)\left(\varphi_2+4\pi_g^d\right)\left(\pi_g^t\right)^2\Big)\Big].$$

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 $CW^{set}\left(\tilde{t}_{opt}\right)-CW^{set}\left(0\right)\geq 0$, because the single-rooted[6] numerator in (11) is nonnegative, while the denominator is strictly positive. Since for $\tilde{t}_{opt}=0$ it obviously holds that $CW^{set}\left(\tilde{t}_{opt}\right)=CW^{set}\left(0\right)$, we can conclude that $CW^{set}\left(\tilde{t}_{opt}\right)-CW^{set}\left(0\right)>0$ for $\tilde{t}_{opt}\neq 0$.

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

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

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

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

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

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

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

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

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

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 $f_g \in \left[\underline{f}_g, \overline{f}_g\right]$, $\delta \in [\underline{\delta}, 1)$, and $\lambda \in \left[0, \overline{\lambda}\right]$ we have that $\partial\tilde{t}_{opt}/\partial\lambda > 0$.[7]

Proof of Lemma 2:

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

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

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

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

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 $f_g \in \left[\underline{f}_g, \overline{f}_g\right]$, $\delta \in [\underline{\delta},1)$, and $\lambda \in (0,\overline{\lambda})$ 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 $f_g \in \left[\underline{f}_g, \overline{f}_g\right]$, $\delta \in [\underline{\delta},1)$, and $\lambda \in (0,\overline{\lambda})$ 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

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

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

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

so that by respecting (8) we find that

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

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

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

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