

Entry limiting agreements for pharmaceuticals: pay-to-delay and authorized generic deals

Farasat A.S. Bokhari
Franco Mariuzzo
Centre for Competition Policy
School of Economics
University of East Anglia

Arnold Polanski
School of Economics
University of East Anglia

CCP Working Paper 15-5 v3
(updated April 2017)

Abstract

Pay-to-delay deals involve a payment from a branded drug manufacturer to a generic maker to delay entry. In return, the generic receives a payment and/or an authorized licensed entry at a later date but before the patent expiration. We examine why such deals are stable. We combine the first mover advantage for the first generic with the ability of the branded manufacturer to launch an authorized generic to show when pay-to-delay deals are an equilibrium outcome. Policy simulations show that removing the ability to launch authorized generics will deter such deals but removing exclusivity period for first generic will not.

Contact Details:

Farasat Bokhari

F.Bokhari@uea.ac.uk

The authors wish to acknowledge the contribution of the ESRC who supported their research through Centre for Competition Policy funding (ref: RES-578-28-0002).

Entry limiting agreements for pharmaceuticals: pay-for-delay and authorized generic deals[†]

Farasat A.S. Bokhari^{a,b,*} Franco Mariuzzo^{a,b} Arnold Polanski^a

^a*School of Economics, University of East Anglia, Norwich NR4 7TJ, UK*

^b*Centre for Competition Policy, University of East Anglia, Norwich NR4 7TJ, UK*

April, 2017

Abstract

Pay-for-delay deals involve a payment from a branded drug manufacturer to a generic maker to delay entry. In return, the generic receives a payment and/or an authorized licensed entry at a later date but before the patent expiration. We examine why such deals are stable. We combine the first mover advantage for the first generic with the ability of the branded manufacturer to launch an authorized generic to show when pay-for-delay deals are an equilibrium outcome. Policy simulations show that removing the ability to launch authorized generics will deter such deals but removing exclusivity period for first generic will not.

Key words: pharmaceuticals, pay-for-delay deals, reverse payments, authorized generics, Nash bargaining

JEL Classification: L41, K21, K41

[†]This draft has benefitted significantly from comments from Ashish Arora, Steve Davies, Emiliya Lazarova and Abhijit Ramalingam and Pierre Régibeau and Kai-Uwe Kühn. We are also very grateful for the useful feedback given by the seminar participants at the Centre for Competition Policy (University of East Anglia), the departmental seminar at the University of Malaga and the participants at the EARIE and IIOC conference.

*Corresponding Author: [Farasat Bokhari \(f.bokhari@uea.ac.uk\)](mailto:f.bokhari@uea.ac.uk)

1. INTRODUCTION

A pay-for-delay deal involves a ‘reverse payment’ from a patent holder to a generic manufacturer (the challenger) seeking entry for its generic equivalent. In return for the payment, the generic firm may abandon its challenge, but often also acquires a right from the patent holder to enter at a later date as an authorized licensed generic with an exclusive license, and before the patent expiration itself. Whereas previous studies have focused primarily on the welfare effects of patent expirations and out of court settlements with and without reverse payments, and under what conditions they may be anti or pro-competitive, we focus instead on incentives in making pay-for-delay deals [[Lakdawalla and Philipson, 2012](#), [Helland and Seabury, 2016](#), [Branstetter et al., 2016](#)].

Prior literature on reverse payments has relied on institutional details of the American legal system vis-à-vis the market authorization rules and provisions of the Hatch-Waxman Act of 1984, particularly section IV of the Act (called a ‘para IV challenge’) to provide an explanation of how pay-for-delay (P4D) deals come about in the pharmaceutical industry [[Bulow, 2004](#), [Frank, 2007](#), [Hemphill, 2009](#), [Scott Morton and Kyle, 2011](#), [Mulcahy, 2011](#), [Scott Morton, 2013](#), [Regibeau, 2013](#)]. As has been noted in this literature, they are typically in response to para IV challenges and that these are initiated after the principal composition-of-matter patent expires, i.e, the patent protecting the molecule itself expires, but while other patents associated with the drug, as registered by the US Food and Drug Administration (FDA) in the Orange Book, remain in force. For example, the first generic company to successfully file for market authorization under section IV of the Act is explicitly rewarded a six month exclusivity period during which no other generic firm can market its drug, which is in addition to any possible first mover advantages that the firm may acquire. A payment to a generic to stay out would require compensation for giving up these rewards. Conversely, a reverse payment settlement that keeps the challenger out of the market is also profitable for the branded firm [[Drake et al., 2015](#), [McGuire et al., 2016](#)]. Naturally, challenges are more frequent in larger and more profitable markets.

The American arrangement raises the entry cost via para IV challenges, but provides a clear gain for the challengers if successful. The 180-day award to the first filer also explains why P4D deals block entry by other generics and why the originator does not have to pay off all future challengers: if the first filer delays entry for say three years due to a P4D settlement, then it blocks entry for all other generics for three and half years. Based on this, many have called for the loop-hole to be closed, and indeed the Medicare Modernization Act of 2003 (MMA) made amendments to the Act which can trigger a forfeiture of the exclusivity under such cases and requires the settling parties to submit the terms to the Federal Trade Commission (FTC) for antitrust review if it relates to the generic application filed with the FDA.

Nonetheless, as [Hemphill \[2009\]](#) has pointed out, over the years settlements have evolved where the parties have devised new disguises for P4D payments involving ‘side payments’ for licenses or purchase of other products, or even the very existence of agreements. Similarly, [Hovenkamp and](#)

[Lemus \[2016\]](#) document several patent settlement cases that were resolved in the Patent Trial and Appeal Board (PTAB), as opposed to in patent infringement litigation in district courts, which begins if a firm files for generic entry with the FDA. Post settlement in the PTAB they do not find evidence of generic entry, indicating that these might be P4D deals, and may have escaped antitrust scrutiny by settling prior to filing with FDA. If a generic entry application was not already filed with the FDA, the firms would not need to disclose the terms of settlement to FTC. Thus, while the MMA solves one aspect of regulatory problems related to P4D deals, it leaves open the path for parties to settle prior to filing for generic entry with the FDA and avoid antitrust scrutiny.¹

This paper is about the alternative path to P4D deals. We focus on the incentives involved in reaching P4D deals before filing for generic entry i.e., *ex ante* P4D deals when no generic can use its first filer status to block entries by other generic challengers. A key question for us is then: if the originator can pay off the generic producer to refrain from challenging its patent, and to stay out of the market for some time, how much do they have to pay, and why do other generic challengers not grab the same opportunity to also get paid off? And if indeed this is possible, then how is the initial deal profitable for the originator? In the 5-3 majority opinion of the US Supreme Court on the ruling on P4D deals in the case of *Federal Trade Commission v. Actavis, Inc.*, the 180-day exclusivity of the Hatch-Waxman Act (along with a 30-month stay order at FDA in case of a challenge) is precisely why P4D deals are stable:

Would not a high reverse payment signal to other potential challengers that the patentee lacks confidence in its patent, thereby provoking additional challenges, perhaps too many for the patentee to ‘buy off?’ Two special features of Hatch-Waxman mean that the answer to this question is ‘not necessarily so’. [[US Supreme Court, 2013](#), p.16].

The majority opinion goes on to state that because the 180-day exclusivity is not available to latter challengers (even if the first challenger settles in a P4D case), the low potential reward prevents others from seeking entry.

In this paper we investigate how P4D deals would arise if the 180-day exclusivity is available to the late filers, as in the *ex ante* settlements, or for instance if exclusivity was awarded to first successful challenger (hereinafter FSC system) rather than to the first filer (hereinafter FF system), as suggested in [Hemphill and Lemley \[2011\]](#). We show that compared to the FF system, P4D deals are less likely than under the FSC system. However, even with many challengers P4D deals are still possible *ex ante* or under FSC system, i.e., paying off one (or a few) challengers can prevent others from challenging if there is a strong first mover advantage. Thus making the 180-day exclusivity available to late challengers if they are successful would not be *per se* sufficient from preventing P4D deals (though would be an improvement over the FF system). Similarly, removing the 180-day

¹An example is of settlement between ViiV Healthcare and Apotex Corp involving the drug Triumeq reported in [Hovenkamp and Lemus \[2016\]](#). The parties settled in PTAB on August/3/2015. The drug in question had a ‘new chemical exclusivity’ designation attached to it, which was effective until August/12/2018. The exclusivity designation and the period implies that FDA cannot receive a generic drug application this early, and hence it appears that the settlement was prior to filing of generic entry.

exclusivity altogether, another policy option often proposed, would not necessarily prevent P4D deals either. Note that the first to market a generic enjoys no statutory exclusivity period in the EU, and yet entry limiting reverse payments take place on both sides of the Atlantic (see Table 1).²

TABLE 1. Pay to Delay Agreements in the US and Europe

EU: 45 in Jan/00-Jun/08	9 in Jul/08-Dec/09	3 in 2010	13 in 2011
US: 16 in FY08	19 in FY09	31 in FY10	28 in FY11

Source: EC [2012], FTC [2011a].

Based on recent literature we build on two key insights to provide an explanation for the stability of pay-for-delay deals. The first is a first mover advantage for a generic firm that is distinct from any exclusivity period or even an incumbency period and arises due to higher willingness to pay for the first generic relative to other generics, and the second is the ability of a branded manufacturer to launch a generic, known as the ‘pseudo’ generic or an authorized generic (AG), either itself or via a third party under a licensing agreement. We show that together these two factors can lead to P4D deals being stable even in the proposed FSC or in ex-ante settlements and may be relevant in explaining the European P4D deals as well.

To study the stability of P4D deals, we propose a simple model with one branded firm with a patent and many potential challengers. The branded firm can threaten the first challenger to launch its own generic (an in-house AG) and deprive the challenger of any first mover advantage, but would incur an entry cost associated with acquiring speciality to successfully market a generic. Alternatively, if the cost to launch its own AG is too high, it can pay off the first challenger to stay out of the market. If the deal is accepted, the branded firm can use the first challenger to ward off entry by any subsequent challengers. It can do so by threatening to launch a generic via the first paid-off challenger *prior* to the second challenger’s entry in case a patent litigation outcome is in favor of the second challenger. If at any stage the branded firm chooses to execute the threat (launch an AG), it takes away the challenger’s first mover advantage thereby reducing the latter’s incentive to contest entry. However, launching an AG either in-house or via the first challenger, also forces the branded firm to enter into a triopoly rather than engage in a competitive duopoly against the winning challenger, and hence the threat may not always be credible.

We show that if the first mover advantage is larger than a threshold, then under an endogenously determined licensing fee for an AG determined via a Nash Bargaining Solution (NBS), the branded firm is better off in a triopoly with an AG than in a competitive duopoly, thus making the threat credible. This is because if the first generic entrant can capture a significantly large share of the

²Even without any explicit 180-day market exclusivity period in the EU, there may be other barriers to entry, such as delays in market authorization by the medical agencies that create de facto duopoly periods for the first successful generic challenger, making it perhaps similar to the American case (these are discussed in more detail in the appendix). Also in the US, even after the 180-day exclusivity ends, remaining generics do not necessarily enter the market immediately. In fact, a report by FTC [pp. 98-99 2011b] shows that in markets with AGs, entry by later generics happens slowly over time, taking on average 36 months after initial generic entry to evolve to peak value of seven generics in such markets. Nonetheless, we explicitly allow for this duopoly period in our model, and show that unless the first mover advantage is large, the P4D deals may still not happen, i.e., the 180-days exclusivity for the first generic is not a sufficient condition for P4D deals.

generic market, then both the branded firm and the first challenger can agree on a licensing fee that allows the launch. Similar reasoning applies to the case when the branded firm can launch its own in-house AG at zero (or low) entry cost, the only difference being that it fully captures any profits associated with the sales of the AG rather than a negotiated licensing fee. In the ensuing triopoly, the branded firm gets to recoup some of the losses relative to its favored monopoly position due the sales of the AG via the licensing fee (or all of AG's profit if it was self launched) and hence it is better off than being in a duopoly. Thus for a large enough first mover advantage, the threat to launch an AG either itself or via the first challenger is credible, and working backwards, second and subsequent potential challengers may optimally choose to stay out of the market if their expected profit is lower than the cost of litigation.

This result is also robust to any additional 'incumbency advantage' for the second (or later) successful challenger relative to other firms if it enters during the exclusivity period. For instance, suppose the brand settles with the first challenger (which becomes its AG) and a second challenger sues and wins the case. Then if both the AG and the winning generic are launched during the exclusivity period, the profits of the winning challenger may be larger than all later entrants in the post-exclusivity period due to an incumbency advantage. Nonetheless, if first mover advantage is large, we show that P4D deals where only the first generic is paid off and second onwards choose not to challenge is still the outcome.

Finally, when the first mover advantage is not large, subsequent generic firms may choose to challenge entry, in which case it is necessary for the branded firm to make smaller payments to all subsequent challengers to maintain its monopoly position (these payments are slightly larger if there is an additional incumbency advantage). With just a few potential challengers, the branded firm can pay off all the challengers and still be better off than facing litigation. However, if the number of challengers is large, the net surplus from paying off multiple challengers eventually becomes negative and hence no P4D deals are possible when first mover advantage is small.

To demonstrate all this, we first model equilibrium profits and payments with just three firms (brand and two challengers under the FSC system) and show how P4D deals come about when the branded manufacturer can pay off one or both challengers and/or launch its own AG. We then extend it to the case with many challengers and show that P4D deals are still possible if the first mover advantage is larger than a threshold. We compare these to the case when exclusivity is restricted to either just the first filer (the current FF system) or when exclusivity period is removed altogether. Both of these cases give similar outcomes and show that P4D deals are possible over a much larger range of model parameters and that the first mover advantage does not matter in these cases. Finally, we also consider the case where we limit the ability of the branded firm to launch an AG if a subsequent challenger wins patent litigation and show that P4D deals are then not possible.

The next section discusses relevant literature. Section three provides a stylized game between a branded firm and several challengers seeking entry. Section four compares (equilibrium) profits and payoffs for firms under different market structures. This section characterizes the range of parameter values over which P4D deals, litigation or unchallenged monopoly is the outcome. This is followed by a section on policy options. The last section concludes. The paper comes with four appendices. Institutional details of patent litigation and market entry in the US and EU are given in appendix A and our model is based on these details. Appendix B provides proofs of various propositions in the paper. Appendix C extends the game tree to more than two challengers and lists payoffs under different policy options. Appendix D is a sub-model of demand for differentiated products, where consumers have different willingness to pay for branded vs first and later generic entrants, and generates equilibrium profits that are used in the main paper.

2. RELATED LITERATURE

Sequential P4D deals with potential challengers are similar to the logic developed by [Bernheim \[1984\]](#), but with deterrence investment substituted with P4D deals and licensing an AG. In fact, the strategy of launching an AG via a P4D deal with a challenger, as discussed in this paper, is similar to earlier studies that focus on licensing as a strategy to maintain market leadership and/or deter entry. For instance, [Gallini \[1984\]](#) shows the conditions where the incumbent licenses its production technology to a potential entrant in exchange for terminating research into competing or better technology, while [Rockett \[1990\]](#) and [Eswaran \[1994\]](#) provide models where the incumbent licenses either the weaker competitor or a competitor from outside of the industry, so as to crowd the market and discourage stronger competitors from entering. Yet, despite these similarities, important differences exist between our paper and previous studies on licensing. In our paper the generic with the AG licence is the *de facto* strongest competitor to the brand as it enters before other generics and grabs the first mover advantage. Additionally, instead of a license being introduced prior to the potential competitor incurring entry costs, in our paper the license is issued and AG launched only if the next potential entrant has incurred an entry cost, (i.e., a litigation cost) and is successful.

Several studies have documented the impact branded manufacturers have when they launch their own generic or an authorized generic (AG) via a third party on independent generic entry. [Hollis \[2003\]](#) argues that authorized generics deter independent generic entry in intermediate sized markets (and “probably” in other markets as well) while [Reiffen and Ward \[2007\]](#) show that authorized generic entry may deter independent generic entry in small and intermediate sized markets only and raise the long run prices by 1-2%. [Berndt et al. \[2007\]](#) argue that the effect of authorized entry on independent generic entry and ultimately on consumer welfare is likely to be small but still positive. However, [Appelt \[2015\]](#) reports that early authorized entry has no impact on the likelihood of generic entry. As documented in a report by the Federal Trade Commission [[FTC, 2011b](#), pp.17-18], authorized generics can be launched by the branded firm itself (in-house) or via third parties but require expertise in generic marketing. This is because whereas brand name drugs

are typically marketed to physicians and consumers emphasizing the product differentiation and securing placement on formularies, generic drugs are marketed to wholesalers and pharmacies on the basis of price, consistency of supply, and ability to offer a large portfolio of drugs, which is a different expertise.

The first mover advantage for the first generic is in part due to the fact that it enters and serves the market for a longer period of time compared to other generics, but also because it captures and sustains a much larger share of the generic market over a period of several years [Caves et al., 1991, Grabowski and Vernon, 1992, Hollis, 2002, Yu and Gupta, 2014, Shajarizadeh et al., 2015]. For instance, as noted in Hollis [2002], in the Canadian market, the first generic advantage arises due to patients' unwillingness to switch between generic medications, the search and persuasion costs on the part of doctors, and the additional administrative costs of pharmacies when stocking several identical generic drugs with no real monetary incentives due to reference pricing. As our point of entry, we take the first mover advantage as given, and model its impact via differences in maximum willingness to pay for a product, leading to asymmetries in demand curves for differentiated products. Thus, the 'prize' of being the first generic is not just a legislative market exclusivity period where the first generic entrant can operate as a duopolist, but also the relative order of entry – the rewards for which (due to the first mover advantage) are recouped by the entrant in the current period, as well as in the post-patent period when there may be several generic firms.

Building on the theory of harm and the probabilistic nature of patents, the economic and legal literature has focused on the role of antitrust laws on out of court settlements that involve payments from the originator and an agreement on the date of entry by the generic [Shapiro, 2003a, Lemley and Shapiro, 2005, Gratz, 2012]. Under Shapiro's antitrust welfare criteria – that a settlement should leave the consumers at least as well off as the ongoing patent litigation – a payment that exceeds the expected litigation costs of the licensor is sufficient to establish that consumers lose from the settlement [Shapiro, 2003b]. In line with this reasoning, several authors have argued that pay-for-delay settlements should carry a presumption of *per se* anticompetitive behavior (see for instance, Hovenkamp et al. [2003], Bulow [2004], Leffler and Leffler [2004], Hemphill [2009]). In the ruling on P4D deals in the *FTC v. Actavis, Inc.*, the US Supreme Court did not find such settlements to be presumptively anticompetitive. In fact the Court favored the "rule of reason" approach, and reversed the earlier decision by the Eleventh Circuit, which had upheld the P4D agreement as legal and restricted the application of antitrust law under the "scope of the patent test" [US Supreme Court, 2013]. Others have pointed out that while the theory of harm is useful, it has limitations and cannot be applied directly to the more complex agreements between the parties, or that P4D deals can in fact be pro-competitive in some situations, and hence such deals should not be *per se* illegal [Crane, 2002, Willig and Bigelow, 2004, Dickey et al., 2010, Regibeau, 2013]. Similarly, concerns regarding dynamic efficiency and innovation have also prompted researchers to investigate the impact of "prospecting" by generics on the market exclusivity period of the originators, and reported mixed findings [Grabowski and Kyle, 2007, Hemphill and Sampat, 2012].

3. MODEL SETUP

We begin by describing a P4D deal from the US which serves as a motivating example of the stylized model described below. Shire Pharmaceuticals introduced an extended release version of its ADHD drug called Adderall XR in 2001. Under the Hatch-Waxman terms it had exclusivity until April 2005 (initial exclusivity was until October 2004, but then had received pediatric extensions). The underlying patents for the extended release version, unless invalidated, were effective until 2018. In November 2002, Barr laboratories filed an abbreviated new drug application (ANDA) which was followed by a second filing by IMPAX in November 2003. Patent litigation ensued, but Shire settled with both parties before any court outcome. Shire settled with IMPAX (the second filer) to enter the market no later than December 2010, but with a non-exclusive license. It also settled with Barr laboratories (the first filer), which acknowledged that Shire's patents were valid and to agreed to delay entry until April 1, 2009. At that point, Barr would enter with a 180-day exclusive licence from Shire and pay royalties as a proportion of its profits from the sales of generic Adderall XR over the exclusivity period [[Barr Laboratories, Inc., 2006](#)]. Per the terms of the agreement, Barr would also be allowed to enter earlier if another party were to launch a generic version of the drug. Similarly, Teva (which had acquired Barr laboratories in the meantime) started marketing generic version of Adderall XR in the US on April 2, 2009, and six months later IMPAX also entered the market. For a discussion on side payments and additional examples, see [Hemphill \[2007\]](#). Further details of patent litigation and market entry rules in the US and EU are given in [Appendix A](#) and our model is based on these institutional details.

We propose a dynamic game Γ with $J+1$ players that illustrates the essential elements of interactions between a brand name firm B (player 0), which is protected by a patent, and $J \geq 1$ potential generic challengers (G_1, \dots, G_J). As in [Friedman and Wittman \[2007\]](#), our game unfolds in the shadow of a trial. Our stylized game is designed to capture the market authorization rules and main features of P4D cases described earlier and stylized below.

- (1) There are two periods, period 1 which is pre-patent expiration, and period 2, which is post-patent expiration period.
- (2) In period one, the J potential entrants can sequentially contest entry. The branded firm can offer a payment to a challenger to stay out of the market during period one (a P4D deal), and guarantee the order of entry in the post-patent period, as long as the patent is not invalidated by another challenger (order of entry is not guaranteed if the patent is invalidated).³
- (3) If at any stage a challenger (say the j th) does not accept a P4D deal and wins the court case (patent is invalidated), that challenger enters immediately in period one. However, the remaining $J - j$ entrants can only enter in the next period. In a later section we relax

³For instance, the branded firm can allow a generic to use its own production facilities to achieve all regulatory market approval requirements and enter just before other generic firms enter.

the assumption of exclusivity for first successful challenger to no exclusivity for anyone, or exclusivity restricted to just the first filer.

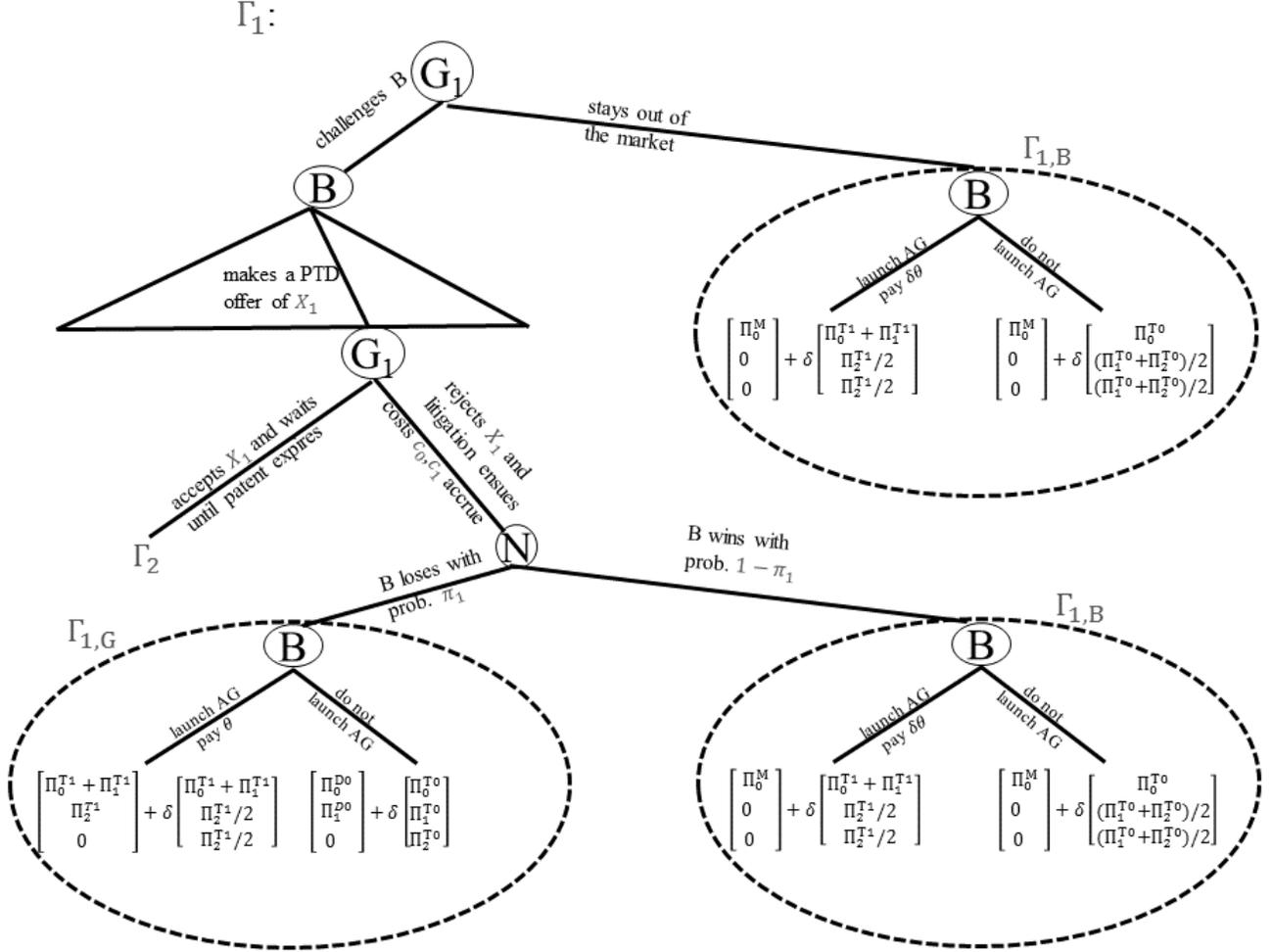
- (4) Additionally, if the j th firm wins the court case, the brand can opt to launch an authorized generic (AG), either itself at an additional cost θ and earn two profits from the brand and its generic product, or via any of the previously paid-off firms, in which case it earns profits from the brand plus a licensing fee L . If the brand launches an AG, period one consists of a triopoly. In what follows we also assume that if the brand launches an AG externally, it is only via the first generic challenger.⁴
- (5) Payoffs from the second period are discounted by factor $\delta \in [0, 1]$. Further, in this period we assume a competitive oligopoly ensues among the $J + 1$ firms, and there are no licensing agreements, as the patent has expired. However, the profits and/or market shares are not equal as the order of entry matters, i.e., one of the generic products has a first mover advantage over the other generics. For the base case we assume that the second through the last generic entrants all earn the same profit (which is less than that of the first generic entrant).⁵

Based on the rules above, the game is as follows. The patent can be challenged in any of the Γ_j stages by a generic challenger j . In the first stage Γ_1 of the game, generic G_1 can choose to stay out of the market, in which case the monopoly continues and the game ends, or challenge entry. If it contests entry, the brand makes an offer of X_1 to G_1 to stay out of the market. If the offer is rejected, litigation ensues. If it is not rejected, the process is repeated with the second challenger. For exposition, the game is depicted in the Figures (1) and (2) below for the special case of $J = 2$ when there are only two potential challengers. The game and payoffs differ slightly for the first versus the second challenger, and hence we show these two cases explicitly, but the generalization to $J > 2$ challengers is similar to the second challenger case and we discuss that later.

Continuing with the example of just two potential challengers ($J = 2$), we denote equilibrium profits due to the sales of the branded or generic drugs in any period by $\Pi_j^M, \Pi_j^{D\#}$ and $\Pi_j^{T\#}$ where M, D , and T stand for profits in monopoly, duopoly and triopoly market structures respectively, and the subscripts $j \in \{0, 1, 2\}$ are for the brand and first and second generic entrants. The superscript ‘#’ is set to 1 or 0 to indicate when an authorized generic has been launched either by the branded firm itself (self-AG) or via one of the paid off generic firms in a P4D deal, and accounts for the possibility of price coordination and joint profit maximization. While we do not specify the exact values of the profits here, we assume that monopoly profits are greater than industry profits in a duopoly, which are in turn greater than industry profits in a triopoly. Further, profits are negatively correlated with entry order, and thus in a triopoly, the branded firm has the highest profits followed by those of the first and then the second generic entrant. Note that the j th generic challenger is not necessarily the same as j th entrant since a generic firm can choose to stay out of a market, and

⁴This is a simplification but follows the example from Shire-Barr deal mentioned above. An alternative is to randomize.

⁵We later relax this assumption and allow the successful generic one to earn more than other generics if it enters in period one (i.e. to model an incumbency advantage).

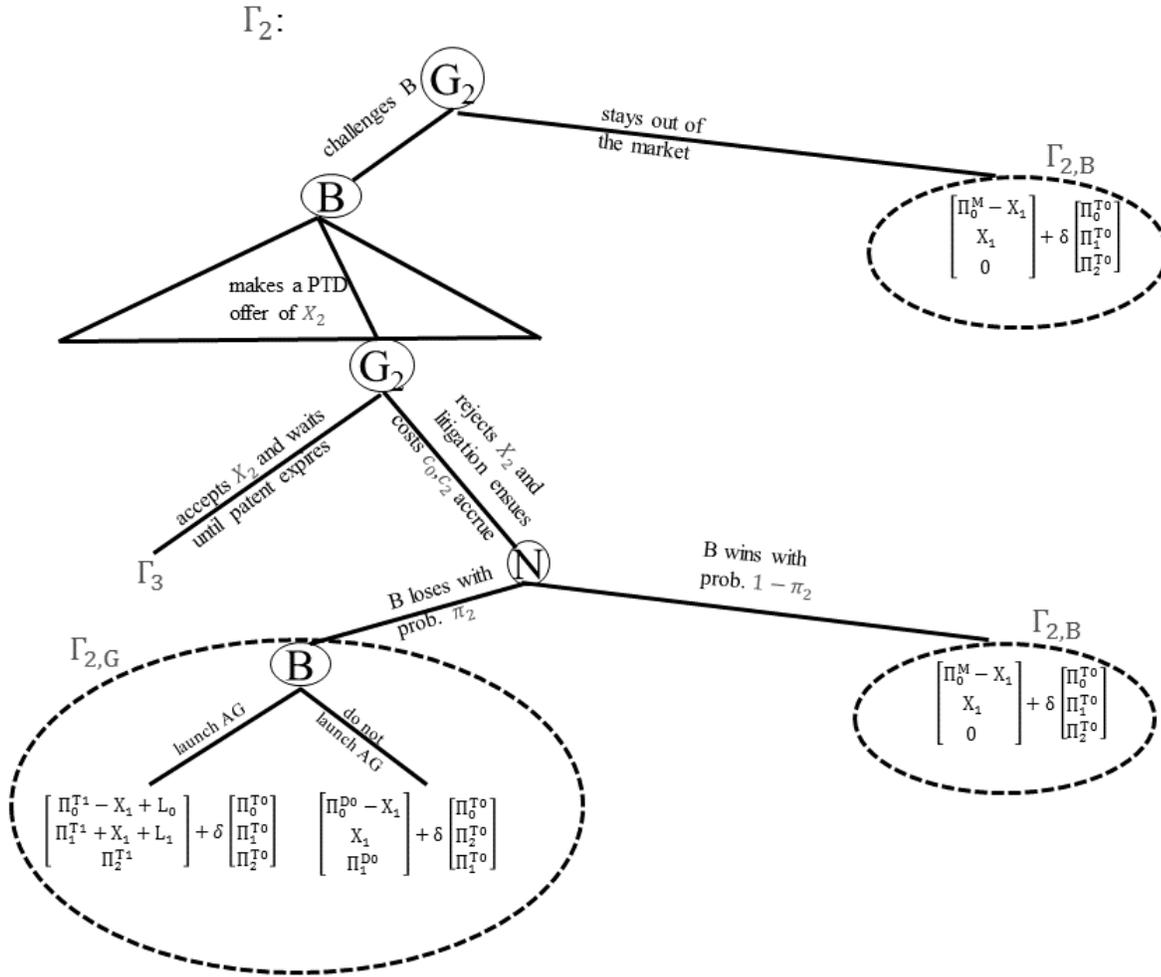


Note that only equilibrium profits from sales are shown in the nodes. The final payoffs include also litigation costs and AG costs as indicated along the branches.

FIGURE 1. Game Tree (Γ_1)

hence we denote the profits of the j th *player* by V_j . For example, consider the case where generic 1 has been paid off and hence agrees to stay out of the market and generic 2 enters the market and duopoly ensues between the brand and the second generic firm. Then, the equilibrium profits for the three players in the first period would be given by $(V_0^D, V_1^D, V_2^D) = (\Pi_0^D, 0, \Pi_1^D)$. Similarly, L_j is the adjustment to the final payoffs of the j th player due to any licensing agreements for an AG and we use the notation $\tilde{V}_j^{T1} = V_j^{T1} + \delta V_j^{T\#} + L_j$ to indicate sum of equilibrium profits from the two periods plus any licensing fee (note that we use the superscript $T1$ on the sum of profits even if the second period is not necessarily T1, as long as the first period is T1). Also, since we assume that if an AG is launched it is only via the first challenger, we can simplify the notation to $L_1 = -L_0 = L$ and $L_2 = 0$.

If at any of the two stages the generic rejects the offer, litigation ensues and the involved parties incur the costs of c_0 and c_j (to be paid at the end of Γ_j). We assume c_0 is sufficiently low for B to always



Note that only equilibrium profits from sales, P4D payments and licensing fees are shown in the nodes. The final payoffs also include litigation costs and AG costs as indicated along the branches.

FIGURE 2. Game Tree (Γ_2)

prefer litigation over unopposed entry and the ensuing competition. The outcome of the litigation is modeled by the fictitious player (N, Nature), who decides randomly with probabilities $1 - \pi_j$ and π_j , respectively whether the brand B is successful with its lawsuit over patent infringement or not.⁶

As shown in Figures (1) and (2), the brand has the option of launching an AG at several of its decision nodes. For convenience, we will denote the subgames that start at these nodes as $\Gamma_{j,y}$, where j denotes the challenger and $y = \{B, G\}$ denotes the relevant path of the game: $y = B$ if either the brand wins the case or if the generic stays out, and $y = G$ if the generic wins. Note also that in the first stage when G_1 is the current challenger, the branded firm has the option to launch AG itself, whereas in the later stages, the option to launch an AG is only via the first paid-off generic firm. Hence, the first P4D deal contains - unlike the successive P4D deals - an (implicit) option to become

⁶As a special case we let $\pi_j = \pi$ for all j where π then represents the strength of the patent with $\pi = 0$ being a very strong patent and $\pi = 1$ being a very weak patent.

an AG producer.⁷ Alternatively, if a self-AG is not launched and the generic does not challenge (as in $\Gamma_{1,B}$), the order of entry between the two generics for the second period is randomized (and hence the profits are depicted as expected generic profits). Further, if the branded firm launches a generic itself, the firm incurs a fixed entry cost θ of entering a generic market (or $\delta \cdot \theta$ if the generic is launched in the second period). If both generic challengers have accepted P4D payments, the game ends at the Γ_3 node with payoffs given by $(\Pi_0^M - X_1 - X_2, X_1, X_2) + \delta(\Pi_0^{T0}, \Pi_1^{T0}, \Pi_2^{T0})$ (which note is similar to $\Gamma_{2,B}$ with an adjustment of X_2 payment to the second challenger).

The final payoff to a player along a path of the game Γ consists of the corresponding (continuation) profit in the ensuing market structure adjusted by the P4D payments and/or litigation costs received and/or paid along the path. Except for some specific values of the parameters, the finite game Γ has a unique subgame perfect equilibrium (SPE) that can be readily computed by backward induction. In particular, we can compute the minimum offer that G_j , $j = 1, 2$, will accept in the SPE from the condition,

$$u_j(\Gamma_{j+1}) + X_j = \pi_j u_j(\Gamma_{j,G}) + (1 - \pi_j) u_j(\Gamma_{j,B}) - c_j, \quad (1)$$

where $u_j(\cdot)$ is the expected payoff to player j in the unique SPE of the subgame. The condition (1) makes the (risk neutral) player G_j indifferent between accepting X_j - and getting the left hand side (lhs) of (1) - and rejecting it - and expecting the right hand side (rhs) of (1). The brand B (player 0) will make the offer X_j in equilibrium, whenever its expected SPE payoff $u_0(\Gamma_{j+1})$ after paying X_j (receiving X_j if it is negative) exceeds its expected payoff from the litigation, i.e., when,

$$u_0(\Gamma_{j+1}) - X_j > \pi_j u_0(\Gamma_{j,G}) + (1 - \pi_j) u_0(\Gamma_{j,B}) - c_0. \quad (2)$$

By combining (1) and (2), we obtain the condition for an agreement in Γ_j and the implied P4D payment as stated in the next proposition.⁸

Proposition 1. Under the take-it-or-leave-it offer, if the condition,

$$u_0(\Gamma_{j+1}) + u_j(\Gamma_{j+1}) > \pi_j \left(u_0(\Gamma_{j,G}) + u_j(\Gamma_{j,G}) \right) + (1 - \pi_j) \left(u_0(\Gamma_{j,B}) + u_j(\Gamma_{j,B}) \right) - c_0 - c_j \quad (3)$$

holds, then the brand B and the generic G_j , agree in Γ_j on the P4D payment,

$$X_j = \pi_j u_j(\Gamma_{j,G}) + (1 - \pi_j) u_j(\Gamma_{j,B}) - c_j - u_j(\Gamma_{j+1}). \quad (4)$$

Otherwise no P4D payment is made and a court litigation between B and G_j ensues.⁹

Proof. See [Appendix B](#).

⁷Note also that in $\Gamma_{1,G}$ if the brand launches its own AG in the first period, the first and second generics' profits in post patent period are set equal to $\delta \Pi_2^{T1}/2$, i.e., they split the profits associated with a third product in a triopoly as there is no incumbency advantage for generic one, even though it enters in first period. The alternative extreme would be to assign $\delta \Pi_2^{T1}$ to generic one and zero economic profit to generic two with similar adjustments in $\Gamma_{j,G}$ for $j > 1$ cases. We consider the outcomes from such incumbency advantage in a later section.

⁸Note that our model allows X_j to be negative, in which case it is not a 'reverse' payment, or the usual P4D deal, but rather a 'forward' payment. This can happen, for instance, if the profits for a generic from being the first generic in the post-patent period are large enough so that it makes a payment to reserve this position.

⁹Without the loss of generality, we assume that B and G_j disagree in equilibrium when the condition in the proposition holds with equality (and, hence, B and G_j are indifferent between agreement and disagreement). The equilibrium payoffs are the same if we assume an agreement.

In this game of perfect information, G_j is able to compute the condition (3) and, in case it is satisfied, the P4D payment given in equation (4). Hence, it can rationally decide whether to challenge B or not. The following corollary gives a condition under which G_j challenges B .

Corollary 1. In the SPE, the generic G_j , challenges the brand B in Γ_j if,

$$X_j + u_j(\Gamma_{j+1}) > u_j(\Gamma_{j,B}), \quad (5)$$

where X_j is defined in (4).

Proof. See [Appendix B](#).

The corollary simply states that G_j 's post-challenge continuation payoff must exceed G_j 's outside option.

3.1. Licensing Fee. As mentioned earlier, when a branded firm launches an authorized generic, it would charge a licensing fee. The authorized generic, however, is only launched if it increases the profit of the branded firm relative to an alternative outcome, but also increases the profit of the generic. In our game tree described above, and with just two challengers, this would be in the sub-game $\Gamma_{2,G}$, where the second challenger G_2 rejects payment X_2 to stay out of the market, and the court decides in favor of the generic. In this case, in period one the brand's options are either to earn Π_0^{D0} (duopoly with no AG) or to earn Π_0^{T1} (triopoly with an AG) plus a licensing fee.

We model licensing fees arising out of a Nash Bargaining Solution where the brand and the generic reach a fee schedule by splitting the net surplus from the launch (i.e., they have equal bargaining power). Then by launching an AG (T1 configuration), the profits due to the sales of the branded drug are $V_0^{T1} = \Pi_0^{T1} + \delta\Pi_0^{T0}$, where the second part is from sales in the post-patent period, and similarly, those due to sales of the authorized generic are $V_1^{T1} = \Pi_1^{T1} + \delta\Pi_1^{T0}$. On the other hand, by not launching the AG, the profits for the two products are $V_0^{D0} = \Pi_0^{D0} + \delta\Pi_0^{T0}$ and $V_1^{D0} = 0 + \delta\Pi_2^{T0}$, respectively (see [Figure 2](#)). Note that we are explicitly accounting for the entry order of the challengers, where the first paid-off challenger either makes a profit $\delta\Pi_1^{T0}$ or $\delta\Pi_2^{T0}$ in the post-patent period, depending on whether it was launched in period one or not.

Thus, in this sub-game with just two challengers, the net surplus from launching an AG is $(\Pi_0^{T1} + \Pi_1^{T1} - \Pi_0^{D0}) + \delta(\Pi_1^{T0} - \Pi_2^{T0})$, where the second term in the parenthesis is due to the relative gain in profits of the first challenger in the post-patent period due to entering first or entering second. Consequently, two period profits inclusive of a licensing fee for the three firms are (if an AG is launched post losing a court case),

$$\begin{aligned} \tilde{V}_0^{T1} &= (1/2) \times (\Pi_0^{T1} + \Pi_1^{T1} + \Pi_0^{D0}) + \delta\Pi_0^{T0} + (1/2) \times \delta(\Pi_1^{T0} - \Pi_2^{T0}) \\ \tilde{V}_1^{T1} &= (1/2) \times (\Pi_0^{T1} + \Pi_1^{T1} - \Pi_0^{D0}) + \delta\Pi_2^{T0} + (1/2) \times \delta(\Pi_1^{T0} - \Pi_2^{T0}) \\ \text{and } V_2^{T1} &= \Pi_2^{T1} + \delta\Pi_2^{T0} \end{aligned} \quad (6)$$

where \tilde{V}_j^{T1} as defined earlier, is the sum of profits from the two periods adjusted by the licensing fee.¹⁰

3.2. Extension to $J > 2$ and related cases. The game naturally extends to more than two challengers where Γ_j is repeated for $j \in \{2, \dots, J\}$ with the only difference being that the Π_j payoffs will be based on oligopoly profits $\Pi_j^{N\#}$ rather than for instance by triopoly profits given by $\Pi_j^{T\#}$. [Appendix C](#) provides an example of profits under a oligopoly with $J + 1$ firms along with the structure of Γ_j and where we maintain the assumption that if the brand reaches an agreement with all the challengers, then in the post patent period the order of entry is given by the order of challengers (i.e., first paid off firm gets the first mover advantage). [Appendix C](#) also provides other important extensions. Notably changes in listed profits in $\Gamma_{j,G}$ – the payoffs if j -th generic wins the case – allows us to model (i) the FF system (exclusivity available to only the first generic), (ii) no exclusivity (if a generic wins, all challengers can enter immediately and there is no duopoly period), (iii) an incumbency advantage for an entering challenger (if a generic wins and enters in period one, it continues to earn more than other generics in post patent period even, if an AG is launched), and (iv) finally when a brand is prevented from launching an AG if an independent generic wins a court case.

Observe that solving Γ_j , i.e., finding out whether G_j challenges B and computing X_j , requires the solution to the game Γ_{j+1} first. Hence, SPE payoffs in Γ_j and all payments X_J, \dots, X_j are found by a recursive computation that uses equation (4) and Corollary 1 at each step J, \dots, j . For example, if this computation yields that the generics G_J, \dots, G_j challenge B and agree on the P4D payments X_J, \dots, X_j , then the brand's expected SPE payoff in Γ_j is,

$$u_0(\Gamma_j) = u_0(\Gamma_{J+1}) - \sum_{s=j}^J X_s \quad (7)$$

where $u_0(\Gamma_{J+1})$ is the payoff to the brand B after the game ends with J P4D agreements. If all these P4D payments are positive, the condition (2) for a fixed j will be eventually violated when the number of generics J is sufficiently large. In this case, B and G_j will go to court. On the other hand, a condition for a universal agreement on P4D deals is specified in the next proposition.

Proposition 2. If all generics G_1, \dots, G_J challenge the brand B , then B will agree in the SPE on the P4D payments X_1, \dots, X_J if for all $j = 1, \dots, J$,

$$u_0(\Gamma_{J+1}) - \sum_{s=j}^J X_s > \pi_j u_0(\Gamma_{j,G}) + (1 - \pi_j) u_0(\Gamma_{j,B}) - c_0, \quad (8)$$

where X_j is defined in equation (4).

Proof. See [Appendix B](#).

¹⁰An alternative is to be agnostic about the fee setting process and set the profits inclusive of the licensing fees to be $\tilde{V}_0^{T1} = \Pi_0^{T1} + \rho \Pi_1^{T1}$ (for the branded firm), $\tilde{V}_1^{T1} = (1 - \rho) \Pi_1^{T1}$ (for generic 1) and, $V_2^{T1} = \Pi_2^{T1}$ (for generic 2), where $\rho \in (0, 1)$ is an exogenously set parameter based on the relative bargaining power.

4. RESULTS

4.1. First Mover Advantage. To get intuition into the way the game unfolds, we derive equilibrium profits under alternative market structures (monopoly, duopoly, triopoly, oligopoly) and where the order of entry establishes a first mover advantage (FMA). All else equal, in a duopoly the branded drug earns more than the first generic so $\Pi_0^D \geq \Pi_1^D$ due to differences in demand, and similarly in a triopoly, profits are ordered as $\Pi_0^T \geq \Pi_1^T \geq \Pi_2^T$.

To obtain such profit orders, and to systematically study the effect of FMA, we model demand for differentiated products using a quadratic utility function $U(\mathbf{q}) = \boldsymbol{\alpha}\mathbf{q} - \frac{1}{2}\mathbf{q}'\boldsymbol{\Sigma}\mathbf{q}$, where the vector $\boldsymbol{\alpha}$ specifies maximum willingness-to-pay (WTP) for the brand, generic 1, generic 2, and so on (details are in [Appendix D](#)). Thus, for instance, in a triopoly $\boldsymbol{\alpha} = (\alpha_0^T, \alpha_1^T, \alpha_2^T)$ and based on utility maximization, derived demand is linear involving intercepts (a_0^T, a_1^T, a_2^T) and slope coefficients. Similarly, in a duopoly $\boldsymbol{\alpha} = (\alpha_0^D, \alpha_1^D)$ and $\boldsymbol{\alpha} = \alpha_0^M$ in a monopoly.¹¹ We also fix the market size across alternative market structure, i.e. potential number of patients is the same across monopoly, duopoly or triopoly (no new patients are discovered if a generic enters the market though the actual realized demand may be different due to different prices). The two key parameters in our model are $\kappa \in [0, 1]$ and $\lambda > 0$ which model the relative magnitudes of WTP parameters and hence the magnitude of first mover advantages. Specifically, we model FMA between the first and second generic entrant via $\kappa \in [0, 1]$, where $\kappa = 0$ implies WTP for first and second generic is the same i.e. $\alpha_1^T = \alpha_2^T$, and $\kappa = 1$ means that the first generic entrant has the maximum advantage relative to the second generic entrant (where we set it to be the same as that for the branded drug, i.e., $\alpha_0^T = \alpha_1^T$). The second parameter λ sets the relative market size between the generic and branded segments of the market, and is also determined by the WTP for generics relative to that of the branded product. When $\lambda = 1$, the total size of the generic market is fixed and set equal to the branded market ($a_1^{(T)} + a_2^{(T)} = \lambda a_0^{(T)}, \lambda = 1$).

We derive equilibrium outcomes (prices, quantities and profits) using Nash-Bertrand price competition with differentiated products with the possibility of joint profit maximization if a generic is launched as an AG.

4.2. Equilibrium Outcomes. For selected values of parameters, we compare equilibrium outcomes for different values of κ and λ under alternative configurations (monopoly, duopoly, triopoly and with and without AGs).¹² The first three panels of [Figure \(3\)](#) show equilibrium prices, quantities and profits in a competitive triopoly (T0, no authorized generics) with two generic firms. To provide insight here we only show outcomes from one period ($\delta = 0$) and results for the general case of $J > 2$ challengers and profits summed over two periods ($\delta > 0$) are given later. These are

¹¹Also, $\boldsymbol{\Sigma}$ is a symmetric positive definite matrix which we parameterize with just two terms, β on the leading diagonal and γ as the term on off-diagonals and in the case of a duopoly, $\boldsymbol{\Sigma}$ is a two by two matrix with similar terms, while in the case of a monopoly, it is a scalar equal to β .

¹²Our results are not sensitive to the specific values set for the parameters and are based on the restrictions discussed earlier. In these graphs, we set $\beta = 1$, $\gamma = 0.5$, $\alpha_0^M = 50/\gamma$, and $\lambda = 1$. Further, we set constant marginal costs to zero.

plotted as functions of $\kappa \in [0, 1]$, our parametrization of relative first mover advantage for the first generic.

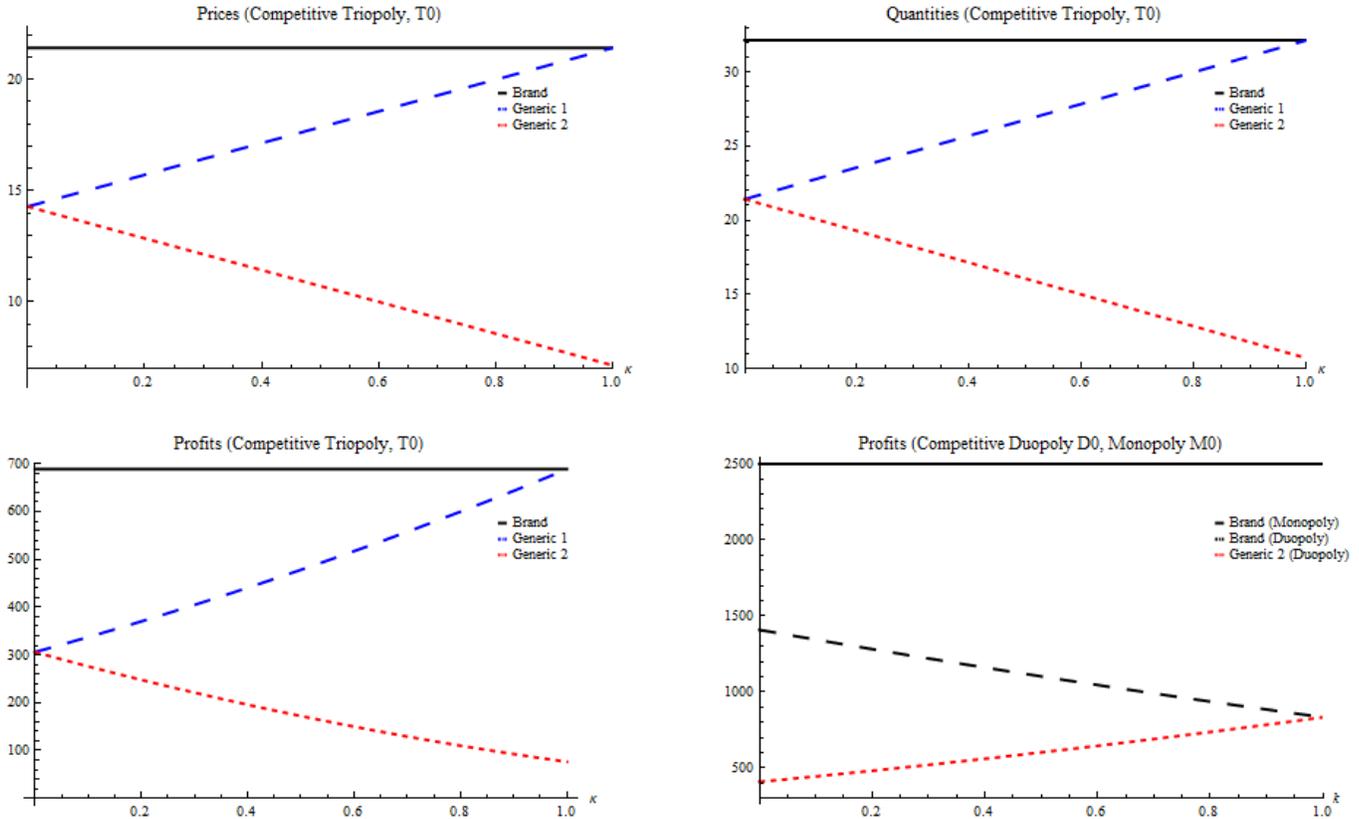


FIGURE 3. Non-Collusive Triopoly, Duopoly, and Monopoly

Since κ measures the relative first mover advantage by changing the WTP of these drugs, but not that of the branded drug, prices, quantities or profits of the latter firm do not vary over κ . For generic 1 (if it is the first entrant), as κ increases, prices, quantities and profits increase while for generic 2 (if it is the second entrant) they decrease. At $\kappa = 0$, neither of the generics have an advantage over the other, and hence at this value, prices, quantities and profits of the two generics are equal to each other. Since by construction the total size of the generic market is fixed and set equal to the branded market ($a_1^{(T)} + a_2^{(T)} = \lambda a_0^{(T)}, \lambda = 1$), a generic firm's output and profits become equal to those of the branded firm only when κ is one, because at this extreme value, the willingness-to-pay for the generic is equal to that of the branded drug.

By contrast, the last panel (bottom right) shows the single period profits in a monopoly as a function of κ , and those of the brand and a generic in a competitive duopoly, where for instance, the first challenger does not enter the market due to a P4D deal, but the second challenger wins the patent litigation and a duopoly follows (i.e., $V_0^{D0} = \Pi_0^{D0}$ and $V_2^{D0} = \Pi_1^{D0}$). Note that in a duopoly at $\kappa = 0$, the WTP for the branded drug is below its WTP in a monopoly, and falls to that in a triopoly as κ increases to one. Consequently, the profits for the branded firm in a duopoly fall in κ , while those

of the generic entrant increase. Further, under our parametrization of the first mover advantage, in D0 configuration they would be equal only at $\kappa = 1$ and be well below those of the branded firm in a monopoly.

The most important aspect of these graphs is the general monotonic increase/decrease of profits in κ as depicted in Figure 3, as they do not depend on the precise values of $\beta, \gamma, \alpha_0^M$ and λ (changing these values only changes the relative magnitudes but not the shapes). In that respect, the general shapes of the curves are invariant to the selected values of the parameters.

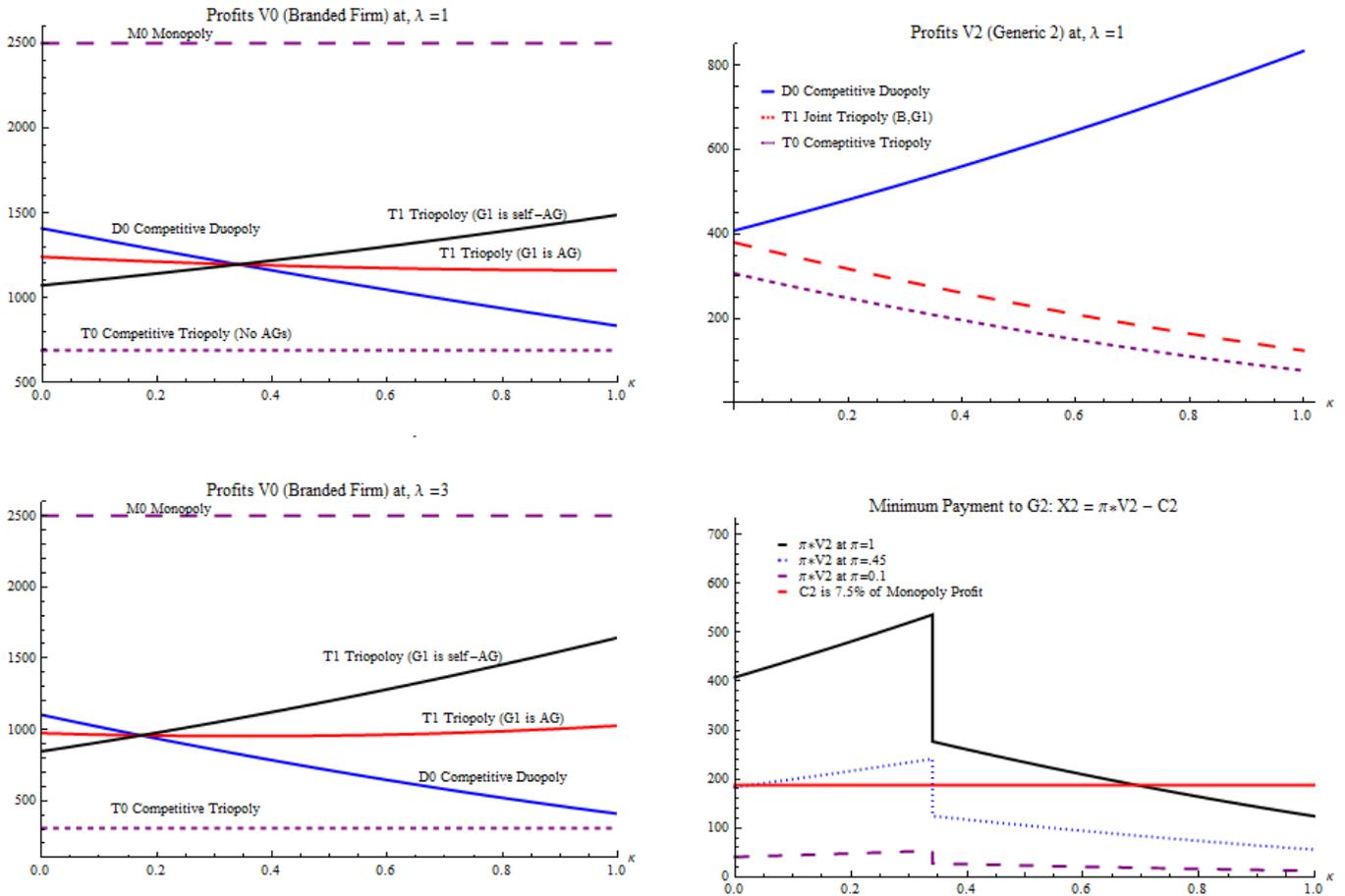


FIGURE 4. Profits of Branded and Second Generic Challenger

Next, to better understand the amount of payments in P4D deals and when these will lead to later generic challengers staying out of the market, we compute and plot equilibrium profits for a given firm as a function of first mover advantage (κ) for different market structures. Figure 4 shows the profits of the branded firm and second challenger under a number of alternative cases, and are inclusive of licensing fees if an AG is launched.

Specifically, in the top left panel of Figure 4 are the profits of the branded firm in five different situations: (i) a monopoly, (ii) a T0 competitive triopoly with no AGs, (iii) a T1 triopoly when generic 1 is an AG and the brand earns a licensing fee, (iv) a T1 triopoly but where the generic

1 has been launched by the branded firm at zero cost, i.e., $\theta = 0$ and thus the brand earns two profits, and (v) a D0 competitive duopoly. In this panel the branded firm's profits are depicted for the special case where $\lambda = 1$. Similarly, the top right panel shows the profits of the generic that is competing against the brand in either a D0 competitive duopoly, or in a T0 or a T1 triopoly with AG, but the generic under consideration is not the firm that enters first and grabs the first mover advantage in the triopoly. The third panel (bottom left), is the same as the first one but with a higher value of λ , i.e., when generics have a larger share of the total market due to the WTP values discussed earlier. Finally, the last panel shows expected profits for a second generic under alternative values of π , the probability of winning a court case. We discuss each of these cases below.

4.3. Credible Threat. As can be seen from the first panel, as well as from the third panel for a higher value of λ , the most desirable position from the perspective of the branded firm is the monopoly profit, and the least desirable is the competitive triopoly profit, neither of which depend on κ (see the figures above). More generally when $\delta \neq 0$, note that $\tilde{V}_0^{D0} = \Pi_0^{D0} + \delta\Pi_0^{T0}$ is decreasing in κ over the entire range while \tilde{V}_0^{T1} , which is inclusive of the licensing fee, starts below \tilde{V}_0^{D0} , but eventually is greater than \tilde{V}_0^{D0} (these are marked “D0 Competitive Duopoly” and “T1 Triopoly (G1 is AG)” in the graphs for the branded firm above). We label the intersection point of these two curves as the credible threat point κ^* , such that for all $\kappa \geq \kappa^*$, the branded firm's profits are higher in a triopoly with an authorized generic than in a competitive duopoly, i.e., $\tilde{V}_0^{T1}(\kappa) \geq \tilde{V}_0^{D0}(\kappa)$.

Similarly, if the branded firm can launch its own generic (at some cost θ) then it can earn two profits from the sales of the branded and the generic drug and the sum of these profits is also increasing in κ . This is because the profits of the first generic in a triopoly increase in κ (see bottom left panel in [Figure 3](#)). In fact, for $\theta = 0$, it crosses the D0 Competitive Duopoly line at exactly the same point (see the curve marked “T1 Triopoly (G1 is self-AG)”) and hence if $\kappa \geq \kappa^*$, the branded firm's profit in a T1 triopoly with a self launched generic is greater than its profit in a competitive duopoly.¹³ However, for positive values of θ this last curve shifts to the right (not shown) and for large enough values of θ the threat by a branded firm to launch its own generic may not be credible. Further, to launch own generic after losing the case requires that cost θ be lower than some threshold value ($\theta \leq \theta^{**}$) and that $\kappa \geq \kappa^*$.

Note that the general shapes of these curves do not change much due to the selected parameter values. For instance, if λ is increased from 1 to 3, (see third panel above), it increases the market share of generics relative to the brand and consequently moves the credible threat point to the left.

¹³The fact that for $\theta = 0$, the intersection points are the same is not a coincidence. This is because \tilde{V}_0^{T1} includes a licensing fee and is computed via Nash Bargaining Solution with equal bargaining power for both parties. Thus, \tilde{V}_0^{T1} is computed as what the branded firm would have earned without launching an AG plus $\rho = .5$ times the net surplus from launching. If instead the branded firm had full bargaining power so that $\rho = 1$, then the line would pivot at the intersection point and overlap the self-AG line where the branded firm would own all profit (or loss) incurred the generic, i.e., it would be equal to the sum of two profits.

We provide below the condition under which a threat point κ^* exists for launching an external AG, as well as the threshold value of θ above which the branded firm cannot launch its own generic.

Proposition 3. If there exists a $\kappa^* \in [0, 1]$ such that $\Pi_0^{T1}(\kappa^*) + \Pi_1^{T1}(\kappa^*) + (\Pi_1^{T0}(\kappa^*) - \Pi_2^{T0}(\kappa^*)) = \Pi_0^{D0}(\kappa^*)$, then under a Nash-Bargaining solution for the licensing fee, the threat to launch an AG is credible for all $\kappa \geq \kappa^*$ if $\partial\Pi_0^{T1}/\partial\kappa \approx 0$.

Proof. See [Appendix B](#).

The condition $\partial\Pi_0^{T1}/\partial\kappa \approx 0$ – that the equilibrium profit for the branded firm in T1 is nearly constant, i.e., not changing with the relative first mover advantage between generic one and generic two – is stronger than needed. What we need for net surplus to be increasing in κ is the condition $|\partial\Pi_0^{T1}/\partial\kappa| < |\partial\Pi_1^{T1}/\partial\kappa|$, i.e., the branded firm’s equilibrium profit is decreasing in first mover advantage at a slower rate than the increase in the equilibrium profit of the first generic entrant so that the overall net surplus still keeps on increasing in κ (recall that Π_0^{T0} does not change with κ , but Π_0^{T1} can decrease in κ due to price coordination between the brand and the AG, see lower-left panel in [Figure 3](#) for the shape of Π_0^{T0}).

Proposition 4. Let $\theta^{**}(\kappa) = (\Pi_0^{T1} + \Pi_1^{T0} - \Pi_0^{D0}) + \delta \cdot \theta^*(\kappa)$ where $\theta^*(\kappa) = (\Pi_0^{T1} + \Pi_1^{T1} - \Pi_0^{T0})$. If $\theta \leq \theta^{**}(\kappa)$ then branded firm would prefer to launch its own generic in period one if it loses litigation to the generic challenger. Alternatively, if it wins the challenge, it would prefer to launch in period two (post patent expiration) if $\theta \leq \theta^*(\kappa)$.

Proof. See [Appendix B](#).

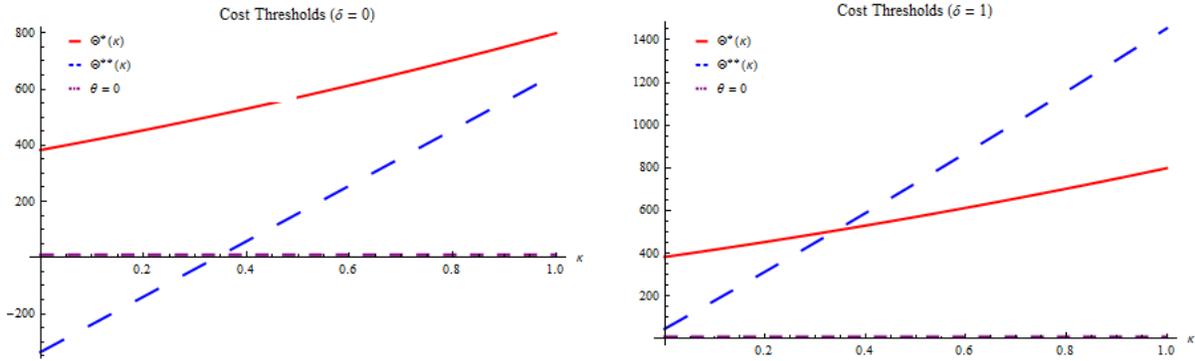


FIGURE 5. Cost Thresholds

Note that $\theta^{**}(\kappa)$ and $\theta^*(\kappa)$ are both increasing functions of κ , and if $\delta = 1$, then $\theta^{**}(\kappa)$ cuts $\theta^*(\kappa)$ from below (see [Figure 5](#)). However if $\delta = 0$, then θ^{**} is always below θ^* (since $\Pi_0^{D0} \geq \Pi_0^{T0}$) and more importantly for low values of κ the threshold $\theta^{**}(\kappa)$ will be negative. This in turn implies that even if $\theta = 0$, for low values of κ the brand will not find it profitable to launch its own generic and only credibly threaten to launch its generic against a winning challenger if the first mover advantage is large. We next discuss the payments to generic challengers.

4.4. P4D Payments. Consider first the subgame $\Gamma_{2,G}$ above. In this subgame, the first challenger (generic 1) has been paid-off an amount X_1 to drop the patent challenge, and in return will be allowed to enter first for the second period (providing the patent is not invalidated) while the second challenger (generic 2) is contesting the patent validity. If $\kappa \geq \kappa^*$, then the brand will always find it profitable to allow generic 1 to enter in period 1 as an AG rather than be in a competitive duopoly if the second challenger wins the court case.¹⁴ Further, when the threat is credible, as the top right panel (Figure 4) shows, the second challenger's profits would be much lower than when it was not credible, i.e., they would be Π_2^{T1} rather than Π_2^{D0} , since if the branded firm launches an AG, it can make sure that the AG enters first and claims the first mover advantage. In this case, generic 2 may well choose to stay out of the market and not challenge entry if its incremental expected profit post entry is less than its litigation costs, i.e., if $\pi\Pi_2^{T1} < c_2$ (this is the incremental expected profit since in period two, either way generic 2 would still earn Π_2^{T0}).

This situation is depicted in the fourth panel in Figure 4 above which shows the expected profits for the second generic for different values of π , which can be read as the strength of the patent ($\pi = 1$ is a weak/invalid patent and $\pi = 0$ is a strong/iron-clad patent) and when the litigation costs are set to 7.5% of the monopoly profits. Note that for $\kappa < \kappa^*$ they are increasing in κ , then drop to a much lower value at κ^* , and decrease thereafter and eventually can fall below c_2 for a large enough value of κ . In this case, the second generic chooses to stay out of the market. Alternatively, if the first mover advantage is not large, i.e. $\kappa < \kappa^*$, then the branded firm's preferred outcome is D0 duopoly over a T1 triopoly with an AG. In this case, the second challenger may well prefer to enter over the option of staying out since it can enter as a duopolist and grab the first mover advantage given that the first challenger has opted to stay out. However, a low value of κ also implies that the generic firm's profits are also small (for the given parameter values, they are roughly around 400 at $\kappa = 0$ if it were to succeed in invalidating the patent) while the brand has much to lose (roughly 1400 in D0 instead of 2500 in a monopoly) and will prefer to pay off the second challenger as well in the amount of $\pi\Pi_2^{D0} - c_2$, than lose its monopoly position.

Thus with just two potential challengers ($J = 2$), either both will stay out of the market (an unchallenged monopoly) if the patent is strong (π is low) and/or cost of litigation is high, or the branded firm can always pay off both firms in P4D deals to maintain its monopoly in period 1. For a given litigation cost, whether it pays off both or only the first challenger, and the second optimally stays out, depends on π and κ with the possibility of paying off only the first firm starting at $\kappa \geq \kappa^*$.

It can be verified that Proposition (1) implies the payment to the second challenger for a P4D deal (with $J = 2$) is

$$X_2 = \begin{cases} \pi\Pi_1^{D0} + \pi\delta(\Pi_1^{T0} - \Pi_2^{T0}/(J-1)) - c_2 & \text{if } \kappa < \kappa^* \\ \pi\Pi_2^{T1} - c_2 & \text{otherwise.} \end{cases} \quad (9)$$

¹⁴Note that in this subgame, X_1 will be subtracted from both $\tilde{V}_0^{D0}(\kappa)$ and $\tilde{V}_0^{T1}(\kappa)$ hence the value of X_1 will not matter in the comparison.

Thus, for $\kappa \leq \kappa^*$ the challenger must be paid its expected profit in period one (as a duopolist) plus the expected premium due to first mover advantage in period two minus the litigation costs, while if $\kappa \geq \kappa^*$, the payment to stay out falls to the expected profit of the second generic entrant in a triopoly minus the litigation costs. Note also the payment X_2 first increases in κ up to κ^* , and then falls in κ after that. See bottom right panel in [Figure 4](#).

Similarly, the payments to the first challenger are also based on expected profits in period one (net of litigation costs) either as first entrant in a duopoly ($\pi\Pi_1^{D0} - c_1$) if $\theta > \theta^{**}(\kappa)$ or as a second entrant in a triopoly ($\pi\Pi_2^{T1} - c_1$) if $\theta \leq \theta^{**}(\kappa)$ plus the expected premium in period two, due to the first mover advantage. The exact values (given below) further depend on whether the branded firm can profitably launch an AG in the post patent period even if its generic does not enter in period one, i.e., if θ is less than or greater than $\theta^*(\kappa)$, and hence

$$X_1 = \begin{cases} (\pi\Pi_1^{D0} - c_1) + \delta(1 - \pi)[(\Pi_1^{T0} + \Pi_2^{T0})/J - \Pi_1^{T0}] & \text{if } \theta > \theta^{**} \text{ and } \theta > \theta^* \\ (\pi\Pi_1^{D0} - c_1) + \delta(1 - \pi)[\Pi_2^{T1}/J - \Pi_1^{T0}] & \text{if } \theta > \theta^{**} \text{ and } \theta \leq \theta^* \\ (\pi\Pi_2^{T1} - c_1) + \delta(\pi\Pi_2^{T1}/J - \Pi_1^{T0}) + \delta(1 - \pi)[(\Pi_1^{T0} + \Pi_2^{T0})/J] & \text{if } \theta \leq \theta^{**} \text{ and } \theta > \theta^* \\ (\pi\Pi_2^{T1} - c_1) + \delta(\Pi_2^{T1}/J - \Pi_1^{T0}) & \text{if } \theta \leq \theta^{**} \text{ and } \theta \leq \theta^* \end{cases} \quad (10)$$

where θ^* and θ^{**} are defined in Proposition (4). The net surplus from the payments to the two challengers is calculable from the expression in Proposition (1) (i.e., difference between the lhs and rhs) and the brand will pay off both firms as long as this net surplus is positive (equivalently, if the condition in Proposition (1) holds).

4.5. Agreement Simulations. We evaluated the game tree over combinations of κ and π values between zero and one and with alternative parameter values. [Figure 6](#) shows the type of agreement outcomes (litigation, P4D deals etc.) from four selected cases but where λ, β, γ and $\tilde{\alpha}_0$ are as before (see footnote 12). The litigation costs for all firms are set equal to 7.5% of the monopoly profits.

In the first panel, there are only two challengers and we have set both $\theta = 0$ and $\delta = 0$ ($\theta = \text{Low}$). If the patent is strong ($\pi \approx 0$) and litigation costs are high, the challengers choose to stay out (labelled ‘I – Unchallenged Monopoly’). If the patent is weak ($\pi \approx 1$), the branded firm prefers to pay off the challengers and is able to do so rather than take its chances in a court (labelled ‘II – P4D, Pay All’). Further, the boundary between whether both the challengers are paid off or if they choose to not challenge the brand is marked by a tradeoff between the strength of the patent, and the relative first mover advantage: The actual payments to the two challengers are identical and increasing in κ until $\kappa = \kappa^*$, and are based on expected profits from entering in duopoly ($\pi\Pi_1^{D0} - c_j$). Starting at the threshold value, the payoffs to the challengers drop down to the expected profit of second generic entrant in a triopoly ($\pi\Pi_2^{T0} - c_j$) and thereafter further decrease in κ . Thus, as κ increases, the patent can be weaker and the two challengers can still be paid-off by the generic firm.

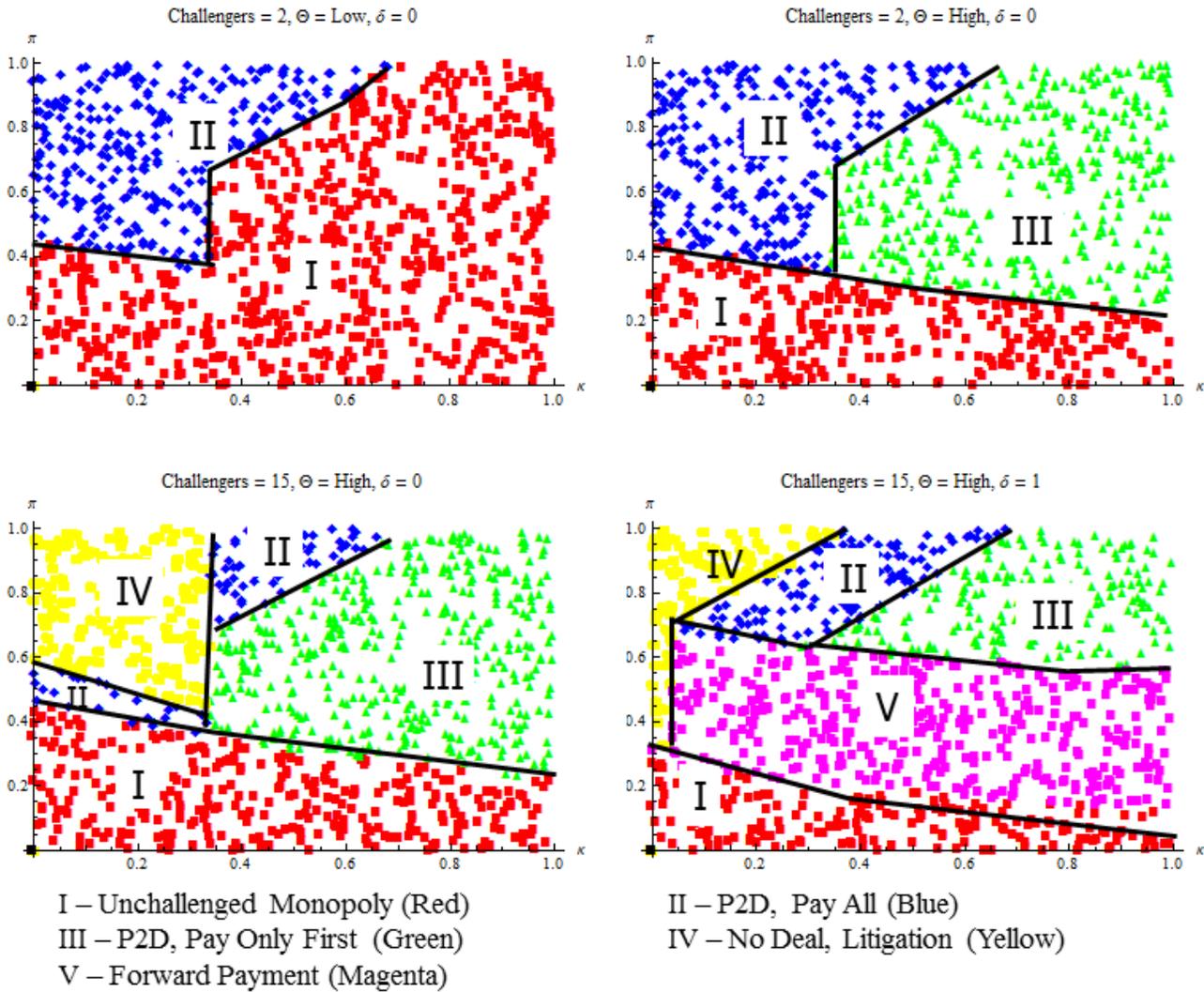


FIGURE 6. Agreement outcomes

An interesting case appears when the branded firm cannot launch its own generic. This is shown in the next panel (top right), where we set $\theta > \theta^{**}$ and $\theta > \theta^*$ ($\theta = \text{High}$). As before, if the patent is strong ($\pi \approx 0$), neither of the generics challenge and the monopoly continues. For somewhat weaker patents, say $\pi \approx 0.6$, as we move in the direction of increasing the first mover advantage, level of payments and nature of P4D deals change. For $\kappa \leq \kappa^*$, the branded firm pays off both challengers and each is paid based on expected duopoly profits ($\pi \Pi_1^{D0} - c_j$). The magnitude of the payments becomes larger as κ increases. However, when $\kappa > \kappa^*$, the payments to the first challenger continue as before and keep increasing with κ , but the payments to the second challenger drop off to the level of second entrant in a triopoly ($\pi \Pi_2^{T0} - c_j$) and decrease with κ . This is because the second challenger can now be threatened with the launch of an AG via the first challenger. For a large enough value of κ (or equivalently for high litigation cost) the second challenger optimally stays out of the market (this area is labelled ‘III – P4D, Pay Only First’).

When there are just two challengers then the branded firm can pay off both, but this is not always possible for a large number of challengers. This situation is depicted in the third panel (bottom left) for $J = 15$ case. When there are many potential challengers ($J > 2$), the payments necessary to maintain the monopoly retain the form given above. Specifically, every challenger from the second one onwards must be paid off expected profits in D0 or T1 (depending on whether κ less than equal to, or greater than κ^*) minus their litigation cost, and hence $X_j = X_2$ for $j = 3, \dots, J$.

However, the net surplus with P4D deals with J challengers becomes negative making it impossible to pay off all the firms, i.e. condition in Proposition (2) is violated for large J . In this case, rather than paying off all the challengers, litigation ensues and the ‘Pay All’ region starts changing to ‘IV – No Deal, Litigation’ as shown in the third panel of [Figure 6](#). The payments to later challengers are based on what they would earn in a duopoly (i.e., are $\pi\Pi_1^{D0} - c_j$) if $\kappa < \kappa^*$, but are based on what they would earn as second entrants in a triopoly (i.e., $\pi\Pi_2^{T1} - c_j$) if $\kappa \geq \kappa^*$. Since duopoly payments are larger than triopoly payments, the area to the left of κ^* converges faster in J to ‘IV – No Deal, Litigation’ compared to the area to the right of the credible threat point, but in the limit, all of the earlier ‘Pay All’ region becomes ‘No Deal’ region.

Observe also that increasing the number of challengers does not change the outcomes in the region earlier identified as ‘III - P4D, Pay Only First’. Specifically, with a large number of challengers, the branded firm cannot afford to pay off all the firms. However, it can pay off the first challenger and the second onwards will not challenge as long as (i) $\kappa \geq \kappa^*$ and (ii) the patent is neither too strong (in which case no one challenges) nor too weak where the brand anticipates a large number of small payments that exceed its ability to pay off and hence it does not offer P4D to any firm.

The last panel (bottom right) extends the forgoing analysis to the case when second period profits are not discounted. While the payment formulas X_j are more complicated, the logic of agreement outcomes over the π, κ range is clear and the intuition is similar to when $\delta = 0$. Two main changes from the earlier case are that the threshold κ^* moves to the left, and that a new type of agreement outcome, ‘V – Forward Payment’ appears on the graph. The threshold moves to the left because \tilde{V}_0^{T1} has increased in magnitude more than \tilde{V}_0^{D0} (see equation (6)).

The new region is where the branded firm offers a *negative* payment to the first challenger to stay out of the market in period one and the challenger accepts this payment i.e., the net surplus from the deal is positive. We call this a forward payment region (opposite of ‘reverse payment’) because the generic firm makes a payment to the branded firm and stays out in the current period, but is able to enter first in the post patent period and grab the first mover advantage associated with its entry order. This payment goes to zero if either the future is discounted or if the branded firm has no ability to decide the order of entry. Since (in our model) the branded firm can always launch

an AG just before the patent expiration to help a generic firm grab the first mover advantage, the firm is willing to pay to obtain that position.¹⁵

5. ALTERNATIVE PAYOFFS AND POLICY OPTIONS

5.1. **Incumbency Advantage.** In the FSC system above, the second challenger, were it to enter successfully in period one, does not have an advantage over other generics in post patent period who do not enter in the first period. For instance, in our payoff specification above, if the j -th generic wins the litigation and the brand launches AG (say via the first challenger), then the winning generic earns $\delta(\Pi_2^{T0}/(J - 1))$ in the post patent period, which is the same as what other non-entering generics earn in the second period (and the AG earns $\delta\Pi_1^{T0}$). An alternative is that the winning generic earns more than other generics who do not enter, and in an extreme case captures the entire generic residual market. This may be an important factor for some drugs.¹⁶ Thus, we consider the other extreme where the incumbency advantage is at its maximum and the winning j -th generic earns all of the $\delta\Pi_2^{T0}$ in the post-patent period if it enters in period one, while other generics earn zero profits. This requires that the payoffs in Γ_{jG} are adjusted accordingly, which are shown in Appendix C.

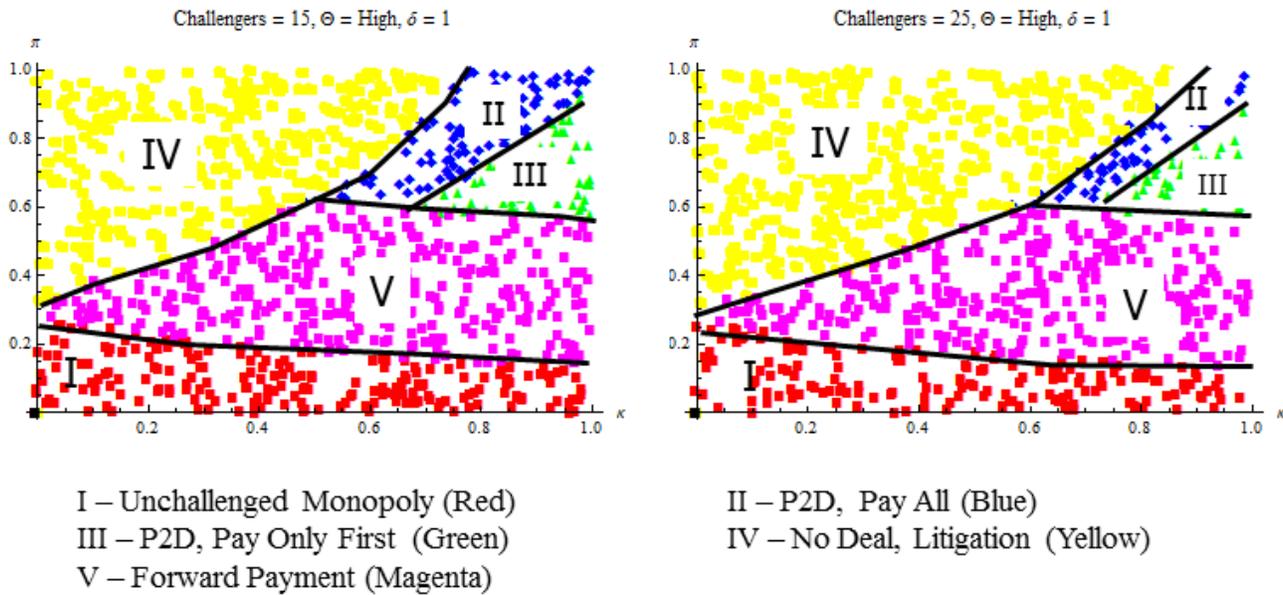


FIGURE 7. Agreement outcomes with Incumbency Advantage

With this change in payoffs, note that the outcomes depicted for the first three panels in Figure 6 do not change (because $\delta = 0$ in these cases). However, the payments X_j increase slightly. For

¹⁵In our model in case no generic firm challenges in period one, the order of entry among the J generics in the post patent period is randomly decided. In an alternative version of the model where we assign entry order in period two to be non-random and arbitrarily given to the first challenger, this new region never arises.

¹⁶A report by FTC [2011b, see graph on p. 104] shows pronounced market shares for second generic (in their graph first-filer) relative to late entrants in post-exclusivity period when the AG and first-filer both enter in the first period.

instance, X_j for $j \geq 2$ changes from $X_2 = \pi\Pi_2^{T1} - c_2$ as given in (9) to

$$X_2 = \pi\Pi_2^{T1} - c_2 + \pi\delta(\Pi_1^{T0} - \Pi_2^{T0}) \quad \text{if } \kappa \geq \kappa^* \quad (11)$$

but are the same for $\kappa < \kappa^*$ (and similar changes in X_1 for the two subcases when $\theta \leq \theta^{**}$). Thus, the second (or j -th challenger) must be paid an additional amount equal to the discounted expected value of incumbency, $\pi\delta(\Pi_1^{T0} - \Pi_2^{T0})$ with similar small increase in the X_1 payment for relevant subcase. Compared to the no incumbency advantage case, this change increases the parameter space over which P4D deals are not possible (area marked as ‘IV - No Deal’ increases) since the threshold moves to the right. But more importantly, parameter space over which P4D deals with one challenger (region ‘III- P2D Pay Only First’) does not shrink as J increases as was the case when there was no incumbency advantage (and as in the previous case, region II shrinks and IV increases with J). Thus P4D deals are still possible under the FSC system, though less so if there is also an incumbency advantage.

5.2. No Exclusivity or FF Exclusivity. We next consider the cases where exclusivity is either not available to anyone (No Exclusivity) or restricted to the first filer only (the FF system). The outcomes from the two cases are similar so we treat them together.

No Exclusivity. Consider the payoffs if the j th challenger wins the court case and all the remaining $J - j$ challengers can enter immediately in period one for free (i.e., without any litigation costs). Then building on our earlier specification where the profits for firms can be approximated as in a triopoly (the brand and the first entrant earn profits of the first two firms in a triopoly $\Pi_0^{T\#}$ or $\Pi_1^{T\#}$ and the profit of all the remaining entrants is equal to the profit of the third firm in a triopoly divided by the number of $J - j$ remaining entrants $\Pi_2^{T\#}/(J - j)$), only the payoffs in the subgame Γ_{jG} change. Specifically, if the brand does not launch an AG but all other challengers can enter in period one, the potential profits for the winning j th challenger change from $(\Pi_1^{D0} + \delta\Pi_1^{T0})$ to $(\Pi_1^{T0} + \delta\Pi_1^{T0})$, while if an AG is launched, they change from $(\Pi_2^{T1} + \delta\Pi_2^{T0}/(J - 1))$ to $(\Pi_2^{T1}/(J - j + 1) + \Pi_2^{T0}/(J - 1))$. The remaining challengers also earn positive amounts rather than zero in the first period and the associated payoffs of all the players under this policy option are shown in [Appendix C](#).

Since the expected profit of the challenger reduces from duopoly based rents to a competitive triopoly, this in turn lowers the payment required to keep the challenger out of the market (see [Figure 4](#)). Similarly, if the branded firm does not launch an AG, its profits also decrease from Π_0^{D0} in period one to Π_0^{T0} . However, $\Pi_0^{T0} + \delta\Pi_0^{T0} \leq \tilde{V}_0^{T1}$ for all values of κ even if it does not charge a licensing fee since it can coordinate on the price with an AG. Effectively, as before, the brand chooses between having one more firm that produces the drug as the first entrant AG with first mover advantage, or one less firm in an oligopoly but with no option to coordinate on price or charge a licensing fee. Consequently, the threat to launch an AG is credible for all values of κ and it is cheaper to pay off a challenger, making P4D deals still possible. The outcomes with $J = 15$ challengers and with $\delta = 0$ or $\delta = 1$ are shown in [Figure 8](#). In both cases, P4D deals are still possible and infact the area of unchallenged monopoly increases.

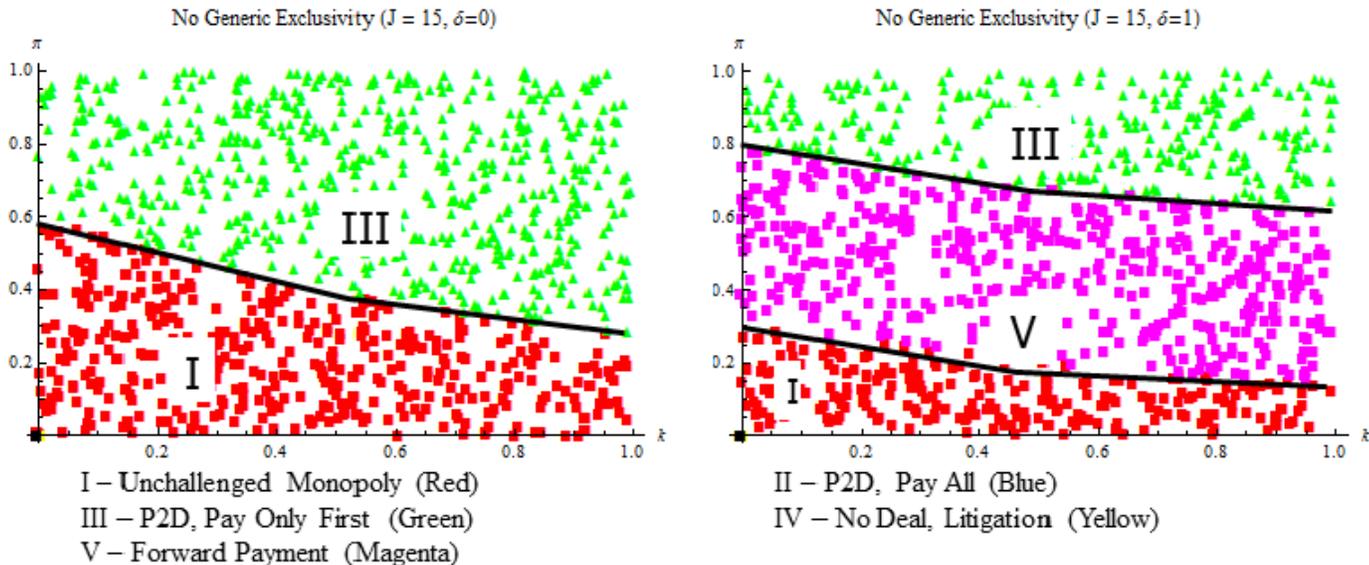


FIGURE 8. No Exclusivity

FF Exclusivity. In this case if the j th firm wins the litigation, it's entry is still blocked if AG is launched (since the first generic has the exclusivity). Thus if the AG is launched the profit for the first generic changes from $(\Pi_1^{T1} + \delta\Pi_1^{T0})$ to $(\Pi_1^{D1} + \delta\Pi_1^{T0})$ plus X_1 payment and minus licensing fee in both cases while the payoff for the winning challenger changes from $(\Pi_2^{T1} + \delta\Pi_2^{T0}/(J - 1))$ to $(0 + \delta\Pi_2^{T0}/(J - 1))$. If however the AG is not launched (post a win by the j -th generic) then exclusivity is lost to all due the MMA 2003 forfeiture clause, and the winning challenger and remaining firms enter immediately. In this case the payoff for the j -th firm is $\Pi_1^{T0} + \delta\Pi_1^{T0}$ and of the remaining generics without P4D deals is $\Pi_2^{T0}/(J - j + 1) + \delta\Pi_2^{T0}/(J - 1)$. See [Appendix C](#) for the full description of the payoffs.

The outcomes (and logic) in this case is similar to the previous one, where the expected profit of the j -th challenger reduces from $(\Pi_2^{T1} + \delta\Pi_2^{T0}/(J - 1))$ to $(0 + \delta\Pi_2^{T0}/(J - 1))$ if an AG is launched making it easier to pay off this challenger. The only difference we observe in the corresponding figures is that boundaries between the regions shown in [Figure 8](#) shift downwards, indicating that P4D deals are available over a larger parameter space (and hence for brevity the figure is omitted).

5.3. No AG against a winning challenger. The branded firm's ability to credibly threaten to launch an AG in case a challenger wins a court case gives rise to the P4D deals. If this option is not available – and hence the threat is never credible, then with enough challengers in the market, a P4D deal will never be reached. In the US, this would mean amending the Hatch-Waxman Act so that it *also* applies to the branded firm: if no other generic firm can enter for 180 days as a reward for invalidating the patent, then the branded firm can also not launch an AG prior to the exhaustion of the 180-day exclusivity by the successful challenger. To understand the implications of such a policy, with the same parameters as before, we modified the tree imposing that the brand

is (legislatively) prevented from launching an AG, and then resolved when there are ten or twenty challengers.

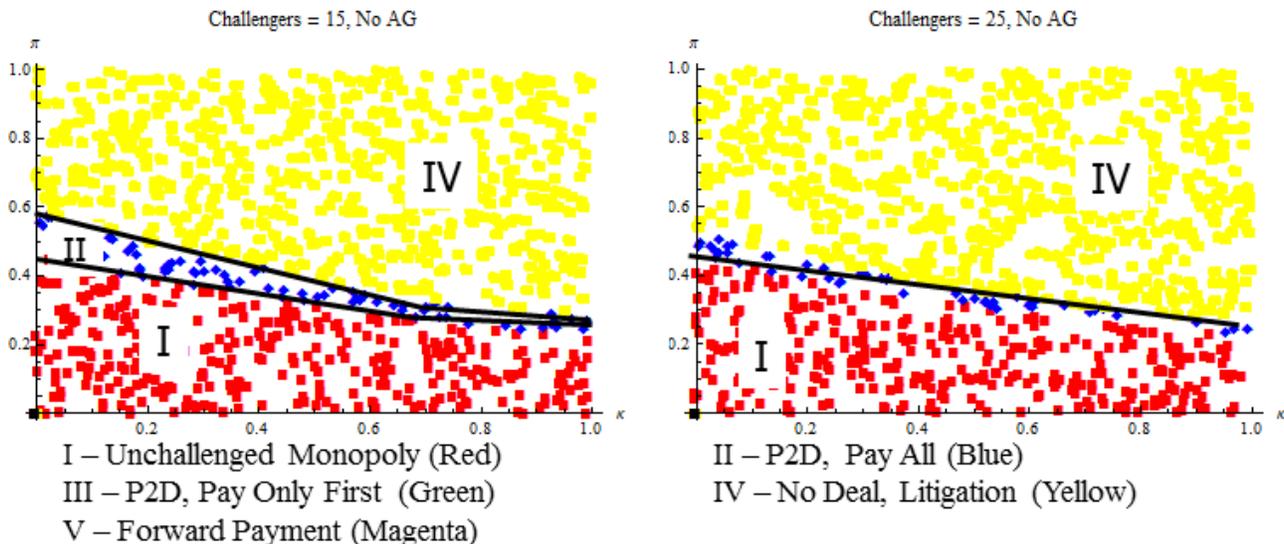


FIGURE 9. No Option to Launch AG

As shown in Figure 9, with no AG option, either by itself or via a third party, the branded firm either has to pay off all the challengers (in case there are few challengers) or if there are many challengers, it may fail to reach an agreement with any of them. This is because after paying off the first challenger, the remaining $J - 1$ challengers never optimally choose to stay out of the market, and hence the region marked as ‘III – P4D Pay only First’ never occurs. The only exception is when even the first firm does not consider challenging the branded firm’s patent because it is too strong ($\pi \approx 0$) relative to the litigation costs. All in all, removing the AG option for the brand leads to either an unchallenged monopoly for relatively strong patents, or a court decision rather than an out of court settlement if there are enough challengers.

6. SUMMARY AND DISCUSSION

The model employed in this paper allows us to study the stability of reverse payment agreements between brand and generic challengers that lead to extended monopoly periods. Prior literature has focused on the welfare effects of out of court settlements with and without reverse payments, and under what conditions they may be anti or pro-competitive. We focused instead on when ex ante pay-for-delay deals would be observed in equilibrium, which are equivalent to exclusivity awarded to the first successful challenger, and how it compares to the current system in the US which awards exclusivity to the first filer.

Our model combines the first mover advantage for the first generic entrant with the ability of the branded manufacturer to launch an authorized generic to describe the conditions under which such deals are an equilibrium outcome. We do not explore all the other possible explanations for this

phenomena. We show that compared to the FF system, P4D deals are more difficult under the FSC system. However even under this system P4D deals can occur. We also show that P4D deals can occur even without any exclusivity period, as they do in Europe.

The model also shows that the payment to stay out increases not only in the ‘weakness’ of the underlying patent, but also in the extent of the first mover advantage. This is important because both the US Supreme Court in the case against *Actavis*, and the European Commission (DG Competition) in announcing the €147 fine against Lundbeck and the agreeing generics in a P4D case, cite the size of the payment as a “workable surrogate” for the weakness of the underlying patent, but ignore the role of the first mover advantage.¹⁷

The first mover advantage arises due to differing willingness to pay for first versus the later generics. If the first mover advantage is large enough, the branded firm can make a credible threat to later challengers of launching an authorized generic via the first challenger. This, in turn, can deter later generics from contesting entry until patent expiration, even if the patent is weak. Further, the payment to the first challenger is correlated not only to the weakness of the patent, but also to the extent of the first mover advantage. Importantly, from a policy perspective, it is not the duopoly period (exclusivity period) that makes such deals possible, but rather the combination of the two factors outlined above, i.e., the first mover advantage and authorized generics (witness the pay-for-delay deals in Europe). We have shown that the removal of the possibility to launch the authorized generic option can be a successful policy to check P4D deals. In this regard greater scrutiny should be given to branded firms’ ability to launch authorized generics.

While our sequential game is stylized, it is sufficiently flexible to capture the relevant strategic interactions. For example, it allows for a wide range of payoff profiles that cover several policy-relevant scenarios, and for application-specific parameters, such as litigation costs and patent strength. Our framework can accommodate different bargaining powers between the brand firm and the generic challengers, when these players negotiate the P4D and the licensing fee. As part of robustness check we experimented with different bargaining environments and parameterizations, and observed the emergence of different shapes and sizes of the five outcome regions I-V discussed in this paper. Generally, this exercise confirmed our main findings on the existence and robustness of the five main regions.

¹⁷See p.19 [US Supreme Court \[2013\]](#) and comments by the Director General (DG Competition) of EC, p.9 [Italianer \[2013\]](#).

REFERENCES

- Appelt, Silvia**, “Authorized generic entry prior to patent expiry: reassessing incentives for independent generic entry,” *The Review of Economics and Statistics*, July 2015, 97 (3), 654–666.
- Barr Laboratories, Inc.**, “Barr and Shire sign three agreements: a product acquisition agreement for Adderall; a product development agreement; and a settlement and license agreement for Adderall XR,” August 14 2006. [Press Release] August 14.
- Berndt, Ernst R., Richard Mortimer, Ashoke Bhattacharjya, Andrew Parece, and Edward Tuttle**, “Authorized generic drugs, price competition, and consumers welfare,” *Health Affairs*, May/June 2007, 26 (3), 790–799.
- Bernheim, B. Douglas**, “Strategic deterrence of sequential entry into an industry,” *RAND Journal of Economics*, Spring 1984, 15 (1), 1–11.
- Branstetter, Lee G., Chirantan Chatterjee, and Matthew Higgins**, “Regulation and welfare: Evidence from paragraph IV generic entry in the pharmaceutical industry,” *RAND Journal of Economics*, Winter 2016, 47 (4), 857–890.
- Bulow, Jeremy**, “The gaming of pharmaceutical patents,” in Adam B. Jaffe, Scott Stern, and Josh Lerner, eds., *Innovation Policy and the Economy*, Vol. 4, MIT Press, 2004, pp. 145–187.
- Caves, Richard E., Michael D. Whinston, and Mark A. Hurwitz**, “Patent expiration, entry, and competition in the U.S. pharmaceutical industry,” *Brookings Papers on Economic Activity*, 1991, 1991, 1–66.
- Crane, Daniel A.**, “Exit payments in settlement of patent infringement lawsuits: Antitrust rules and economic implications,” *Florida Law Review*, 2002, 54 (4), 747–797.
- Dickey, Bret, Jonathan Orszag, and Laura Tyson**, “An economic assessment of patent settlements in the pharmaceutical industry,” *Annals of Health Law*, Winter 2010, 19 (2), 367–400.
- Drake, Keith M., Martha A. Starr, and Thomas G. McGuire**, “Do “reverse payment” settlements constitute an anticompetitive pay-for-delay?,” *International Journal of the Economics of Business*, 2015, 22 (2), 173–200.
- EC**, “Pharmaceutical Sector Inquiry, Final Report,” European Commission (Directorates General, Competition) Brussels, Belgium July 2009.
- _____, “3rd Report on the Monitoring of Patent Settlements,” European Commission (Directorates General, Competition) Brussels, Belgium July 2012.
- Eswaran, Mukesh**, “Cross-licensing of competing patents as a facilitating device,” *The Canadian Journal of Economics*, August 1994, 27 (3), 689–708.
- Frank, Richard G.**, “The ongoing regulation of generic drugs,” *The New England Journal of Medicine*, November 15 2007, 357 (20), 1993–1996.
- Friedman, Daniel and Donald Wittman**, “Litigation with symmetric bargaining and two-sided incomplete information,” *The Journal of Law, Economics, & Organization*, April 2007, 23 (1), 98–126.
- FTC**, “Agreements Filed with the Federal Trade Commission under the Medicare Prescription Drug, Improvement, and Modernization Act of 2003: Summary of Agreements Filed in FY 2010,” A Report by the Bureau of Competition Washington, D.C. May 2011.
- _____, “Authorized generic drugs: Short-term effects and long-term impact,” Federal Trade Commission Report Washington, D.C. August 2011.
- Gallini, Nancy T.**, “Deterrence by market sharing: A strategic incentive for licensing,” *American Economic Review*, December 1984, 74 (5), 931–941.
- Glowicka, Elzbieta, Szabolcs Lorincz, Enrico Pesaresi, Lluís Sauri Romero, and Vincent Verouden**, “Generic entry in prescription medicines in the EU: main characteristics, determinants and effects,” July 8 2009. http://ec.europa.eu/dgs/competition/economist/prescription_medicines.pdf.
- Grabowski, Henry G. and John M. Vernon**, “Brand loyalty, entry and price competition in pharmaceuticals after the 1984 drug act,” *Journal of Law and Economics*, October 1992, 35 (2), 331–350.
- _____, and **Margaret Kyle**, “Generic competition and market exclusivity periods in pharmaceuticals,” *Managerial and Decision Economics*, August 2007, 28 (2-5), 491–502.
- Graham, Stuart J. H., Bronwyn H. Hall, Dietmar Harhoff, and David C. Mowery**,

- “Post-issue patent “quality control”: A comparative study of US patent re-examinations and European patent oppositions,” NBER Working Paper 8807, National Bureau of Economic Research February 2002.
- Gratz, Linda**, “Economic analysis of pay-for-delay settlements and their legal ruling,” Munich Discussion Paper 2012-6, University of Munich, Department of Economics 2012.
- Gürkaynak, Gönenç, Ayşe Güner, and Janelle Filson**, “The global reach of FTC v. Actavis Will Europe differ from the US approach to pay-for-delay agreements?,” *IIC - International Review of Intellectual Property and Competition Law*, March 2014, 45 (2), 128–160.
- Hancher, Leigh**, “The EU pharmaceuticals market: parameters and pathways,” in Elias Mossialos, Govin Permanand, Rita Baeten, and Tamara K. Hervey, eds., *Health systems governance in Europe: the role of European Union law and policy*, European Observatory on Health Systems and Policies, Cambridge: Cambridge University Press, 2010, chapter 15, pp. 635–682.
- Harhoff, Dietmar**, “Economic cost-benefit analysis of a unified and integrated European patent litigation system,” Final Report MARKT/2008/06/D, Institute for Innovation Research, Technology Management and Entrepreneurship 2009.
- Helland, Eric and Seth A. Seabury**, “Are Settlements in Patent Litigation Collusive? Evidence from Paragraph IV Challenges,” NBER Working Paper 22194, National Bureau of Economic Research April 2016.
- Hemphill, C. Scott**, “Drug patent settlements between rivals: A survey,” March 12 2007.
- , “An aggregate approach to antitrust: Using new data and rulemaking to preserve drug competition,” *Columbia Law Review*, May 2009, 109 (4), 629–688.
- and **Bhaven N. Sampat**, “Evergreening, patent challenges, and effective market life in pharmaceuticals,” *Journal of Health Economics*, March 2012, 31 (2), 327–339.
- and **Mark A. Lemley**, “Earning exclusivity: generic drug incentives and the Hatch-Waxman Act,” *Antitrust Law Journal*, January 2011, 77 (3), 947–989.
- Hollis, Aidan**, “The importance of being first: evidence from Canadian generic pharmaceuticals,” *Health Economics*, December 2002, 11 (8), 723–734.
- , “The anti-competitive effects of brand-controlled “pseudo-generics” in the Canadian pharmaceutical market,” *Canadian Public Policy - Analyse de Politiques*, 2003, 29 (1), 21–32.
- Hovenkamp, Erik and Jorge Lemus**, “Reverse settlement and holdup at the patent office,” 2016. Available at [SSRN:https://ssrn.com/abstract=2814532](https://ssrn.com/abstract=2814532).
- Hovenkamp, Herbert J., Mark D. Janis, and Mark A. Lemley**, “Anticompetitive settlement of Intellectual Property disputes,” *Minnesota Law Review*, June 2003, 87 (6), 1719–1766.
- Italianer, Alexander**, “Competitor agreements under EU competition law,” 40th Annual Conference on International Antitrust Law and Policy, Fordham Competition Law Institute, New York, 26 September 2013 2013. Comments by Director General for Competition, European Commission.
- Lakdawalla, Darius and Tomas Philipson**, “Does intellectual property restrict output? an analysis of pharmaceutical markets,” *Journal of Law and Economics*, February 2012, 55 (1), 151–187.
- Leffler, Keith and Cristofer Leffler**, “Efficiency trade-offs in patent litigation settlements: Analysis gone astray,” *University of San Francisco Law Review*, Fall 2004, 39, 33–56.
- Lemley, Mark A. and Carl Shapiro**, “Probabilistic patents,” *Journal of Economic Perspectives*, Spring 2005, 19 (2), 75–98.
- McGuire, Thomas, Keith Drake, Einer Elhauge, Raymond Hartman, and Martha Starr**, “Resolving reverse-payment settlements with the smoking gun of stock price movements,” *Iowa Law Review*, 2016, 81 (4), 1581–1599.
- Mulcahy, Andrew**, “Patent policy and entry: Evidence from pharmaceutical patent challenges.” PhD dissertation, University of Pennsylvania January 2011.
- Nevo, Aviv**, “Identification of the oligopoly solution concept in a differentiated-products industry,” *Economics Letters*, 1998, 59, 391–395.
- Regibeau, Pierre**, “Pay-for-Delay: What do we disagree on?,” *Competition Policy International*, Autumn 2013, 9 (2), 116–127.
- Reiffen, David and Michael R. Ward**, “‘Branded Generics’ as a strategy to limit cannibalization of pharmaceutical markets,” *Managerial and Decision Economics*, 2007, 28 (4-5), 251–265.
- Rockett, Katharine E.**, “Choosing the competition and patent licensing,” *RAND Journal of*
-

- Economics*, Spring 1990, 21 (1), 161–171.
- Scott Morton, Fiona M.**, “Pay-for-delay,” *Competition Policy International*, Autumn 2013, 9 (2), 128–136.
- _____ and **Margaret Kyle**, “Markets for pharmaceutical products,” in Mark V. Pauly, Thomas G. McGuire, and Pedro P. Barros, eds., *Handbook of Health Economics*, Vol. 2, Amsterdam, North-Holland: Elsevier Science Inc., 2011, chapter 12, pp. 763–823.
- Shajarizadeh, Ali, Paul Grootendorst, and Aidan Hollis**, “Newton’s first law as applied to pharmacies: why entry order matters for generics,” *International Journal of the Economics of Business*, June 2015, 22 (2), 201–217.
- Shapiro, Carl**, “Antitrust analysis of patent settlements between rivals,” *Antitrust*, Summer 2003, pp. 70–77.
- _____, “Antitrust limits to patent settlements,” *RAND Journal of Economics*, Summer 2003, 34 (2), 391–411.
- Singh, Nirvikar and Xavier Vives**, “Price and quantity competition in a differentiated duopoly,” *Rand Journal of Economics*, Winter 1984, 15 (4), 546–554.
- US Supreme Court**, “*FTC v. Actavis, Inc.*,” 570 U.S. (2013), Slip Opinion,” Supreme Court of the United States, No 12-416 2013.
- Willig, Robert D. and John P. Bigelow**, “Antitrust policy toward agreements that settle patent litigation,” *The Antitrust Bulletin*, Fall 2004, 49, 655–698.
- Yu, Yu and Sachin Gupta**, “Pioneering advantage in generic drug competition,” *International Journal of Pharmaceutical and Healthcare Marketing*, 2014, 8 (2), 126–150.

APPENDIX A. INSTITUTIONAL DETAILS

The Food and Drug Agency (FDA) in the US and the European Medicines Agency (EMA) in the EU (or national medicinal agencies) are responsible for granting market authorization (MA) for drugs.¹⁸ When applying to the EMA a firm can choose to apply via the community authorization procedure (CAP), where a single application can be used for authorization in multiple jurisdictions, or it can choose to obtain market authorization from a national agency directly and obtain authorization for that member state only. Alternatively, if the drug is already approved in one member state, the firm can apply for the mutual recognition procedure (MRP) at the EMA to gain marketing approval in other member states. Finally, if no national market authorization exists, the firm can also use the decentralized procedure (DCP) at the EMA, which allows for submission of the application in select multiple member states, and where one country is designated as a reference member state. All in all, there are three different procedures (CAP, MRP and DCP) for gaining marketing authorization with the EMA or via 27 national medical agencies.

In the first instance, original drugs are protected from direct competition from generics via patents, which are granted for 20 years and confer monopoly rights to the originators. In the US, the originator lists the relevant patents with the FDA when filing for a New Drug Application (NDA), while in the EU a similar ‘full application’ is filed with the EMA but without any patent linkage. The

¹⁸Information regarding market authorization rules, patent litigation, and other regulation in the pharmaceutical industry is well documented and hence we don’t provide individual citations. Readers interested in further details on EU/US regulations and differences there in, are referred to (among others) [Graham et al. \[2002\]](#), [Harhoff \[2009\]](#), [Glowicka et al. \[2009\]](#), [Hancher \[2010\]](#) and [Gürkaynak et al. \[2014\]](#).

drug approval process for new drugs lasts several years, involving multiple phases of clinical trials establishing safety and efficacy. However, since these cut the effective exclusive market life of the patented drug significantly, both the US and EU provide non-patent exclusivity to the originator to compensate for these delays. In the US, a market exclusivity period for the originator was introduced as part of the Hatch-Waxman Act, where the originator is protected from generic competition via the ‘data exclusivity’ period – a period during which a generic firm cannot rely on the original drug’s safety and efficacy to file its own application. As per the provisions of the Act, a generic can forego clinical trials, citing safety and efficacy already established by the originator’s reference drug, and file instead for bio-equivalence under the abbreviated new drugs application (ANDA) procedure, but not during the data exclusivity period. Testing and establishing bio-equivalence is also expensive and time consuming, but not as much as the clinical trials required when filing an NDA application (see [Appelt \[2015\]](#) for some recent estimates).

The Act also allows the generic firm to use the patented drug for testing bio-equivalence and developing an ANDA application without infringing the patent so that the ANDA application can be filed on the day the data exclusivity expires (this is the so-called Bolar exception). The data exclusivity period is five years for drugs classified as New Molecular Entity (NME), three years for new formulations (which also carry a patent but not on the molecule), and seven years for orphan drugs. Six month extensions can be added on as pediatric exclusivity when the firm conducted and submitted pediatric studies in response to request from the FDA. The five year exclusivity is cut to four years if the generic files under paragraph IV citing that the patent is either not valid, or will not be infringed.¹⁹ In the latter case, the FDA informs the originator, and if the originator objects on grounds of patent infringement within 45 days, a one time 30-month stay order for generic entry comes into effect to allow the courts time to resolve patent litigation. Thus, generic entry typically takes place after resolution of patent litigation or settlement. Finally, as mentioned earlier, the first filer is entitled to a 180-day market exclusivity period against other generics and is entitled to it even if it settles with the branded firm (also, the exclusivity does not pass over to a later ANDA filer if the first filer settles or loses the case). Further, as documented by [Hemphill \[2009\]](#), the first filer may not be just a single generic firm, as all firms that file on the same first day are awarded the 180-day exclusivity against other generics. Multiple filings on the same day can happen due to the Bolar exception since generics can start preparing for the ANDA filing during the exclusivity period.

In the EU there are two routes available to the originators to extend the exclusive marketing of their products from generic competition. The first, available since 1992, is the Supplementary Protection Certificate (SPC) available for medicinal products, which allows originators to extend the original

¹⁹For NME exclusivity, an ANDA application cannot be filed for first five years (or four years if it is para IV challenge), but for other cases, it can be filed but not approved by the FDA until the exclusivity period runs out.

patent for up to five years after the expiration of the original patent, or fifteen years from the first marketing authorization in the EU, whichever is less. While all member states provide SPC, there is no cross-border recognition, and hence the application has to be filed in each country where the originator wants to enforce and extend the patent life [Hancher, 2010, Graham et al., 2002]. Further, patent infringement and validity fall under the jurisdiction of national courts, and hence patent-holders (or parties seeking to revoke granted patents) may have to enter into litigation in multiple countries resulting in duplication of cases [Harhoff, 2009, EC, 2009].²⁰

Second, like the US, there is a data exclusivity period which was also introduced in 1984 as part of the mutual recognition procedure for drug approval in the EU (prior to that, drug approval was at the national level and with varying rules), and similar to ANDA, generics can file an ‘abridged’ application. Initially, data exclusivity extended either to six years from initial market authorization date, or ten years, depending on the member state, and did not include the Bolar provision to allow for use of patented drug for clinical studies.²¹ Further, some member states opted not to allow for data exclusivity to extend beyond the patent expiration of the original product. In 2005, a new ‘8+2(+1)’ exclusivity period was introduced which, (i) added the Bolar provision, and (ii) provided unified rules of exclusivity across member states – eight years of data exclusivity during which a generic cannot file for an abridged application, plus two additional years of market exclusivity, i.e., the generic may file the abridged application but not market the drug, and a final one additional year of market exclusivity for new indication(s) if they constitute a significant clinical benefit.

Thus both the US and EU provide data exclusivity periods during which the generic drugs cannot enter the market but entry can take place afterwards as long as there are no patents protecting the drug. In turn this implies that while there is no 30-month automatic stay order with the EMA, the branded firm can obtain interim injunction from the national court(s) to prevent generic entry until the litigation case is resolved. In the *Pharmaceutical Sector Inquiry* (SI) by the DG Competition, the average patent litigation was 2.8 years, and interim injunctions were granted in 44% of the cases lasting on average 18 months [EC, 2009, pp.229-238]. Even if there are no injunctions, the report also notes that some health authorities responsible for pricing and reimbursement of medical products can require certification from the drug manufacturers that no patent is infringed, and hence in the EU (as in the US) generic entry would mostly take place *after* patent litigation is resolved.

²⁰However, patent ‘opposition’ can be filed at EU level at the European Patent Office (EPO) but must be initiated within the first nine months from the grant of the patent.

²¹Austria, Denmark, Finland, Ireland, Portugal, Spain, Greece, Poland, the Czech Republic, Hungary, Lithuania, Latvia, Slovenia, Slovakia, Malta, Estonia, Cyprus, Norway, Liechtenstein and Iceland provided six years of data exclusivity while Belgium, Germany, France, Italy, the Netherlands, Sweden, the UK and Luxembourg had ten year of exclusivity for the originator.

Finally, while there is also no automatic 180-day exclusivity period for the first generic entrant in the EU, delays in the drug approval process at the EMA or by national authorization agencies may provide the first generic entrant a short lived duopoly period. As pointed out in the SI report, this can happen when a national authorization agency has to act as a reference member state in MRP/DCP application with the EMA, and the work load at the national agency is high enough for it not to be able to process additional applications for another one or two years. The report notes that in 2008-09, several national agencies were already ‘fully booked’, that according to some generic companies they had to ‘book 18 months in advance to get a slot for a product’, and if they experienced any delay in development, they had to miss a whole year [EC, 2009, p.465]. Some generic firms also reported that these bottlenecks were due to ‘misuse of procedures by some applicants, who make “unnecessary” or parallel bookings, possibly also to delay access for other applicants’. While these delays apply to all firms and not just the second generic challengers, they can, nonetheless, create a wedge between the entry dates of the first and later generic challengers so as to create short duopoly periods for the first generic entrant.

APPENDIX B. PROOFS

Proof of Proposition 1. The condition in the proposition obtains as the sum of (1) and (2). If this condition holds, then the net agreement surplus, i.e., the total continuation payoff to B and G_j after agreement minus their total payoff after disagreement, given by,

$$\left(u_0(\Gamma_{j+1}) + u_j(\Gamma_{j+1})\right) - \left(\pi_j(u_0(\Gamma_{j,G}) + u_j(\Gamma_{j,G})) + (1 - \pi_j)(u_0(\Gamma_{j,B}) + u_j(\Gamma_{j,B})) - c_0 - c_j\right),$$

is positive, and both parties will rationally agree. As B makes a take-it-or-leave-it offer in Γ_j , it will extract the entire net surplus. This post-agreement sharing rule is implemented by the P4D payment (4). If the net surplus is negative, i.e., the condition in the proposition does not hold, B prefers the litigation to the agreement. Hence, an unacceptable offer (below X_j) will be made by B , rejected by G_j , and litigation will ensue. \square

Proof of Corollary 1. After challenging B , the generic G_j expects the payoff $X_j + u_j(\Gamma_{j+1})$ in case of agreement with B . This amount is equal to the expected G_j ’s payoff after disagreement as the substitution from (4) shows,

$$X_j + u_j(\Gamma_{j+1}) = \pi_j u_j(\Gamma_{j,G}) + (1 - \pi_j) u_j(\Gamma_{j,B}) - c_j.$$

Hence, G_j ’s expected payoff $X_j + u_j(\Gamma_{j+1})$ after challenging B does not depend on the outcome of the bargaining stage in Γ_j . On the other hand, if G_j does not challenge B , its continuation payoff is $u_j(\Gamma_{j,B})$. A rational G_j will challenge B if the former payoff is greater than the latter. \square

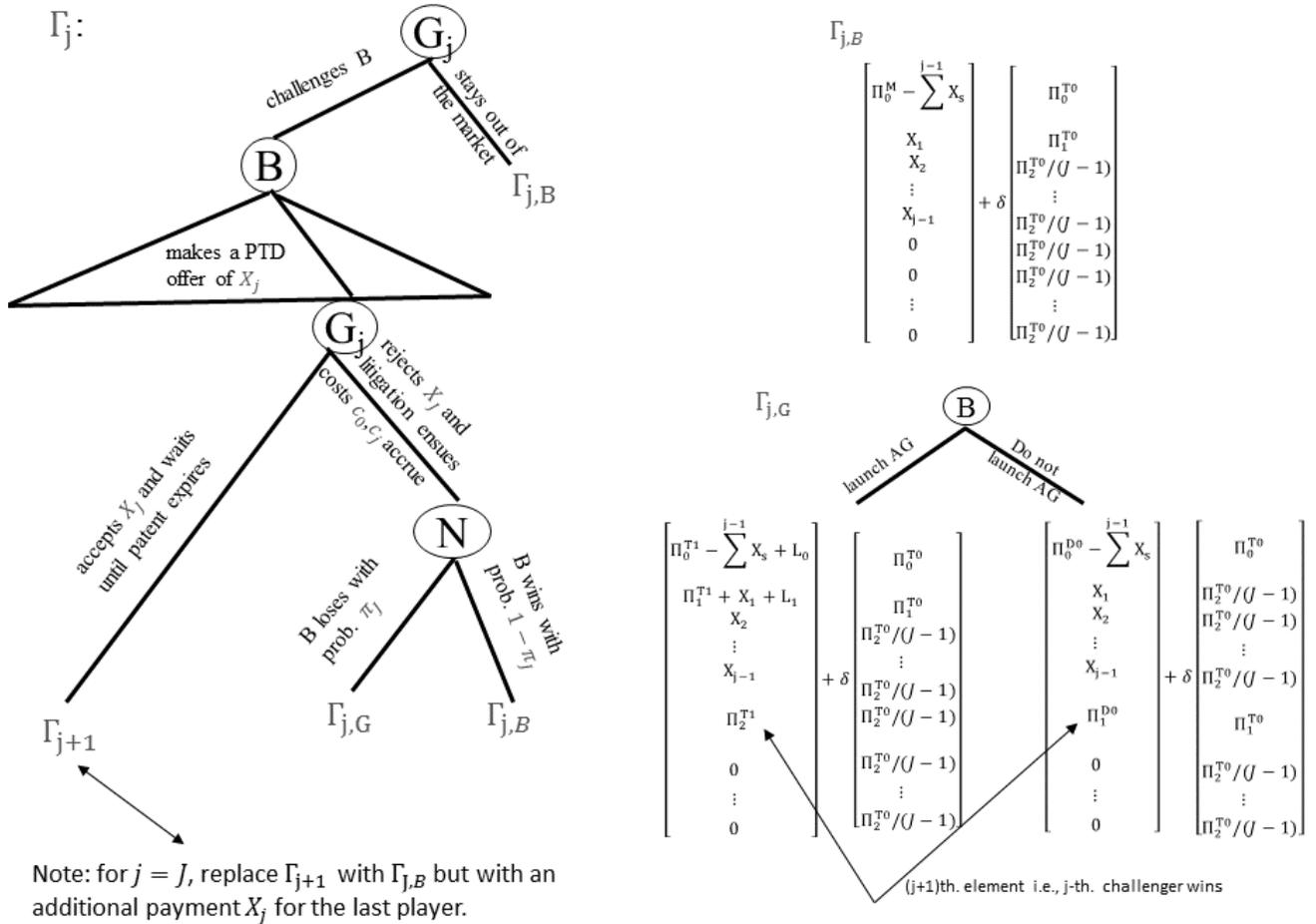
Proof of Proposition 2. For each $j = 1, \dots, J$, the left hand side of (8) is the SPE payoff to B upon agreement with G_j in Γ_j and subsequent agreements with G_{j+1}, \dots, G_J (for $j < J$). Hence, B anticipates in Γ_j that it will make equilibrium P4D payments to G_j and all subsequent challengers if (8) holds for j, \dots, J . The rhs of (8) is B 's expected payoff from litigating G_j (and avoiding the payments X_j, \dots, X_J). Hence, B will agree with all challengers if the former payoff is greater than the latter for all $j = 1, \dots, J$. \square

Proof of Proposition 3. The net surplus from launching an AG with a previously paid off firm when the brand has lost the litigation to a challenger is $(\Pi_0^{T1} + \Pi_1^{T1} - \Pi_0^{D0}) + \delta(\Pi_1^{T0} - \Pi_2^{T0})$. Note that as long as $\partial \Pi_0^{T1} / \partial \kappa \approx 0$, the net surplus is increasing in κ : $(\Pi_1^{T0} \geq \Pi_2^{T0})$ for all values of κ with equality only when there is no first mover advantage, i.e., $\kappa = 0$, while Π_1^{T1} and Π_0^{D0} are respectively monotonically increasing and decreasing in κ . Thus, with the net surplus equal to zero at κ^* , it is positive for all $\kappa > \kappa^*$ and hence the threat is credible for $\kappa \geq \kappa^*$. \square

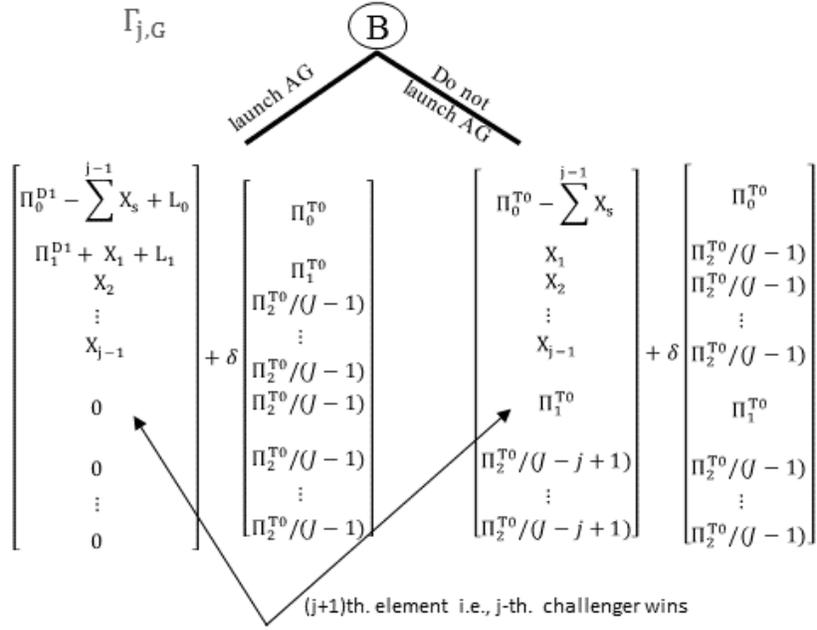
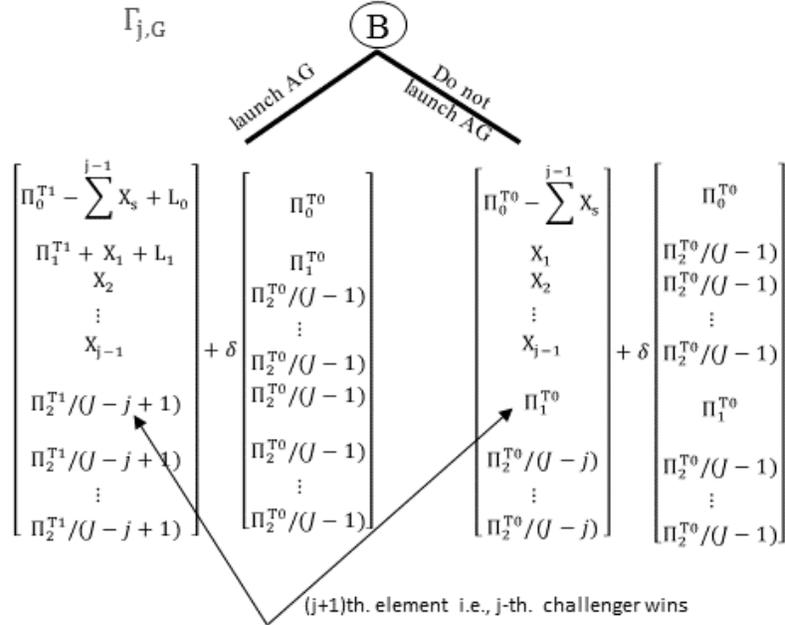
Proof of Proposition 4. The first part follows directly from the subgame $\Gamma_{1,G}$ (see payoffs given in figure 1). After loosing to G_1 , brand would launch a generic if $\Pi_0^{T1} + \Pi_1^{T1} - \theta + \delta(\Pi_0^{T1} + \Pi_1^{T1}) \geq \Pi_0^{D0} + \delta \Pi_0^{T0}$. Rearranging the terms gives the required result $\theta \leq (\Pi_0^{T1} + \Pi_1^{T1} - \Pi_0^{D0}) + \delta \cdot (\Pi_0^{T1} + \Pi_1^{T1} - \Pi_0^{T0}) = \theta^{**}(\kappa)$. Similarly, the second part follows from the subgame $\Gamma_{1,B}$. After winning against G_1 , the brand launches a generic in post-patent period if $\Pi_0^M + \delta(\Pi_0^{T1} + \Pi_1^{T1} - \theta) \geq \Pi_0^M + \delta(\Pi_0^{T0})$. Rearranging gives the required result $\theta \leq (\Pi_0^{T1} + \Pi_1^{T1} - \Pi_0^{T0}) = \theta^*(\kappa)$. \square

APPENDIX C. EXTENSION TO THE GAME TREE

C.1. Payoffs with $J > 2$ Firms. In the game with $J > 2$ challengers, let the equilibrium profits of the j th player from sales of its product be given by $\Pi_j^{N\#}$ (see Figure 10). We model these similarly to those in the triopoly where the first two players earn profits equal to that of the brand and the first generic in a triopoly, and all the later entrants equally share profits associated with the third player in a triopoly (an alternative is to set the profits of later entrants to zero which did not change our results in any significant way). Thus, for instance, in the post patent period with no AGs, the profits would be given by $(\Pi_0^{T0}, \Pi_1^{T0}, \Pi_2^{T0}/(J-1), \dots, \Pi_2^{T0}/(J-1))$ and hence the final payoffs are accounted using the values $\Pi_j^{T\#}$ depending on the entry order. Then the Γ_j subgame would be as shown in the figure below.


 FIGURE 10. Game Tree (Γ_j) with $J > 2$ players

Note that if B loses to the j -th challenger ($j > 1$), then the choice to launch AG or not in the $\Gamma_{j,G}$ is the same as before. Further, if AG is launched, the first mover advantage does not go to the winning challenger. The latter earns Π_2^{T1} in the current period and $\delta\Pi_2^{T0}/(J-1)$ in period 2, while


 FIGURE 12. $\Gamma_{j,G}$ Under FF Exclusivity

 FIGURE 13. $\Gamma_{j,G}$ Under No Exclusivity

APPENDIX D. EQUILIBRIUM PROFITS

In this appendix we model demand with differentiated products and parameterize the first mover advantage (FMA) for the first generic via the demand curves and derive equilibrium profits.

D.1. Market Demand Curves. Following [Singh and Vives \[1984\]](#), we use a quadratic (strictly concave) utility function for a representative consumer to derive linear demand functions for differentiated products, but where differentiation exists up to the third product (second generic product), i.e., products $2, \dots, J + 1$ are homogenous with respect to each other. Thus, let

$$U(\mathbf{q}) = \boldsymbol{\alpha}\mathbf{q} - \frac{1}{2}\mathbf{q}'\boldsymbol{\Sigma}\mathbf{q} \quad (12)$$

where the vector $\boldsymbol{\alpha}$ specifies the maximum willingness-to-pay (WTP) for the brand, generic 1, generic 2, and so on. In a triopoly $\boldsymbol{\alpha} = (\alpha_0^{(T)}, \alpha_1^{(T)}, \alpha_2^{(T)})$, while in a monopoly $\boldsymbol{\alpha} = \alpha_0^{(M)}$ (the branded firm), and similarly $\boldsymbol{\alpha} = (\alpha_0^{(D)}, \alpha_1^{(D)})$ in a duopoly between the branded and the generic entrant. When there are more than three firms in the market, we make the simplifying assumption that the market structure is approximated by a triopoly where the second generic is a collective sum of all the remaining identical generic firms, and thus $\alpha_2^{(T)} = \sum_{j=2}^J \alpha_j^{(N)}$. Similarly, $\boldsymbol{\Sigma}$ is a symmetric positive definite matrix and we parameterize it with just two terms, β on the leading diagonal, and γ as the term on off-diagonals so that, in a triopoly,

$$\boldsymbol{\Sigma} = \begin{bmatrix} \beta & \gamma & \gamma \\ \gamma & \beta & \gamma \\ \gamma & \gamma & \beta \end{bmatrix} \text{ where } \beta > 0 \text{ and } \gamma > 0.$$

As such γ can be negative, positive or zero corresponding to complementary, substitute or unrelated products but we focus on the case when the drugs are substitutes. In the case of a duopoly, $\boldsymbol{\Sigma}$ is a two by two matrix with similar terms, while in the case of a monopoly, it is a scalar equal to β . While $\boldsymbol{\Sigma}$ appears very restrictive with just two parameters, it suffices for our purpose, as we wish to highlight the role of the first mover advantage for the first generic in determining the outcomes in the earlier game, which we capture via the WTP parameters $\alpha_1^{(T)}$ and $\alpha_2^{(T)}$ in relation to $\alpha_0^{(T)}$ for the branded firm. Our motivation for this choice of modeling comes from the fact that patients (and physicians and pharmacists) may view the branded drug to be of a different quality than the generic, but without a price differential they may be less willing to switch from the first to the second generic, i.e., inherently view the latter generic(s) to be of lower quality [[Hollis, 2002](#)]. An alternative would be to model FMA by changing either the parameters that directly affect the demand sensitivity of own price (so that the leading diagonals are not all equal to β but instead given by β_j) or by not making all the off-diagonals equal, particularly $\gamma_{01} = \gamma_{10} \neq \gamma_{02} = \gamma_{20}$. However, these latter parameters are better suited to capture the degree of product differentiation via price effects, and hence we keep this matrix simple, and simply note that the price elasticities will be defined by both sets of parameters (i.e., $\boldsymbol{\alpha}$ and $\boldsymbol{\Sigma}$) and hence the cross-price effects need not be symmetric.

To derive demand functions that correspond to a utility maximization problem, it must be true that Σ is positive definite, which in turn requires that

$$\beta - \gamma > 0 \quad \text{and} \quad \beta + 2\gamma > 0 \quad (13)$$

where the restrictions arise because $|\Sigma| = (\beta - \gamma)^2(\beta + 2\gamma)$ and the eigenvalues are $\{\beta - \gamma, \beta - \gamma, \beta + 2\gamma\}$. The inverse and direct demand functions are then given by $\mathbf{P}(\mathbf{q}) = \boldsymbol{\alpha}' - \Sigma\mathbf{q}$ and $\mathbf{D}(\mathbf{p}) = \Sigma^{-1}(\boldsymbol{\alpha}' - \mathbf{p})$. Solving explicitly, the inverse and direct demand functions for the triopoly are,

$$\begin{aligned} p_0 &= \alpha_0^{(T)} - \beta q_0 - \gamma q_1 - \gamma q_2, & q_0 &= a_0^{(T)} - bp_0 + cp_1 + cp_2 \\ p_1 &= \alpha_1^{(T)} - \gamma q_0 - \beta q_1 - \gamma q_2, & q_1 &= a_1^{(T)} + cp_0 - bp_1 + cp_2 \\ p_2 &= \alpha_2^{(T)} - \gamma q_0 - \gamma q_1 - \beta q_2, & q_2 &= a_2^{(T)} + cp_0 + cp_1 - bp_2. \end{aligned} \quad (14)$$

In the equation above, the parameters a, b, c represent the relative size of the market and price coefficients and are related to the primitives of the model by

$$\begin{aligned} a_0^{(T)} &= [\alpha_0^{(T)}(\beta + \gamma) - \gamma(\alpha_1^{(T)} + \alpha_2^{(T)})]/d \\ a_1^{(T)} &= [\alpha_1^{(T)}(\beta + \gamma) - \gamma(\alpha_0^{(T)} + \alpha_2^{(T)})]/d \\ a_2^{(T)} &= [\alpha_2^{(T)}(\beta + \gamma) - \gamma(\alpha_0^{(T)} + \alpha_1^{(T)})]/d \end{aligned} \quad (15)$$

$$\text{where } b = (\beta + \gamma)/d, \quad c = \gamma/d, \quad \text{and } d = (\beta - \gamma)(\beta + 2\gamma).$$

Since d is positive (see restriction (13)), it also implies that $b > 0$ and $c > 0$. Note that if we allowed complementarities in the model so that $\gamma < 0$ and hence $c < 0$, we would then explicitly require $\beta + \gamma > 0$ for downward sloping demand curves. The demand equations in the case of duopoly and monopoly are similar to the linear structure above but omitted in the interest of space. An additional condition under duopoly is that (13) is modified to $\beta - \gamma > 0$ and $\beta + \gamma > 0$ rather than $\beta + 2\gamma > 0$ under triopoly (but these are automatically satisfied in a duopoly if they are already satisfied in a triopoly).

D.2. Willingness to Pay. To ensure positive demand curves, the intercepts $a_j^{(T)}$ must be positive (equivalently, we can impose second order conditions for profit maximizing which would impose similar restrictions on demand parameters). Positive demand implies that WTP for the two generics $\{\alpha_1^{(T)}, \alpha_2^{(T)}\}$ be such that

$$\begin{aligned} \alpha_2^{(T)} &< \left(\frac{\beta + \gamma}{\gamma}\right) \alpha_0^{(T)} - \alpha_1^{(T)}, & \alpha_2^{(T)} &< \left(\frac{\beta + \gamma}{\gamma}\right) \alpha_1^{(T)} - \alpha_0^{(T)} \text{ and,} \\ \alpha_2^{(T)} &> \left(\frac{\gamma}{\beta + \gamma}\right) \alpha_0^{(T)} + \left(\frac{\gamma}{\beta + \gamma}\right) \alpha_1^{(T)}. \end{aligned} \quad (16)$$

The shaded region in [Figure 14](#) shows the allowed range for WTP parameters for the two generics given the WTP for the branded drug $\alpha_0^{(T)}$ (outside the range the problem is not of any economic

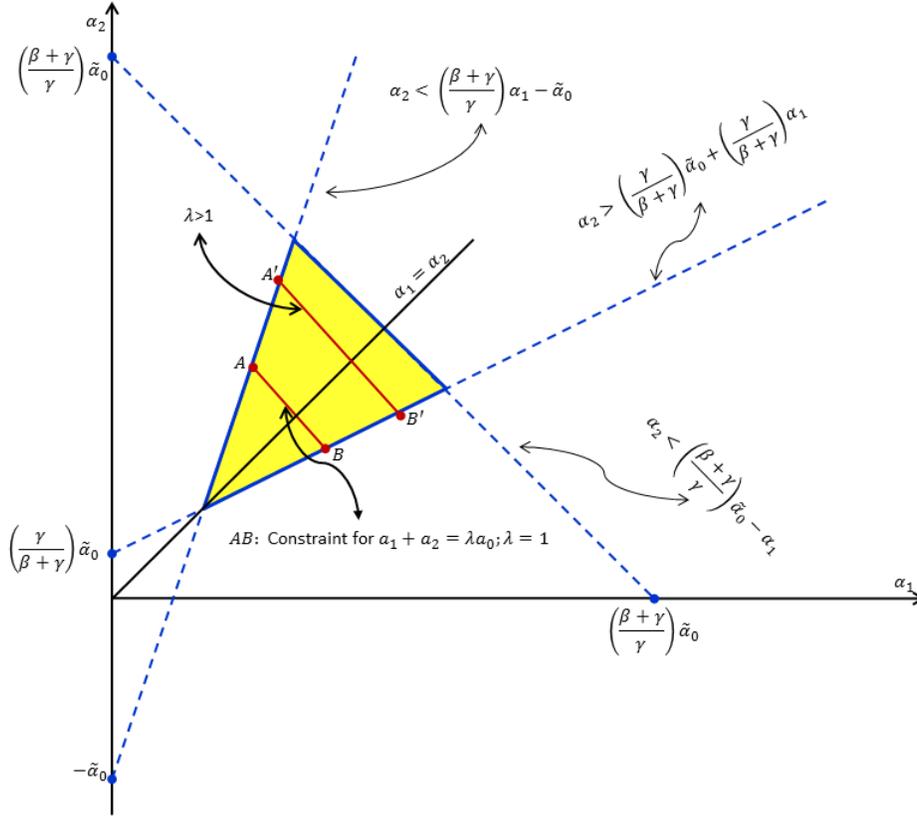


FIGURE 14. WTP for Generics in a Triopoly

interest). The 45° line (given by $\alpha_1^{(T)} = \alpha_2^{(T)}$ but within the region), indicates that a patient's willingness to pay for the two drugs is equal, but increasing relative to the branded drug as we move further away from the origin. This in turn implies that the potential market size for the generics is equal (i.e., $a_1^{(T)} = a_2^{(T)}$, see (15)) on the line, but increases in magnitude as we move further away from the origin. All points off the 45° line increase the WTP for one or the other generic (and consequently imply a larger potential market for that generic). We choose movements along line segments such as $A'B'$ to parameterize first mover advantage, where all points on the line segment fix total potential market size of generics as a proportion of the branded market. Specifically, along all points of $A'B'$, we have $a_1^{(T)} + a_2^{(T)} = \lambda a_0^{(T)}$, where $\lambda > 0$. Then in terms of WTP of the branded drug, points on $A'B'$ are parameterized as

$$\begin{aligned}\alpha_1^{(T)} &= (1 - \kappa) \left[\frac{(2 + \lambda)\gamma + \beta\lambda}{2(\beta + \lambda\gamma)} \right] \alpha_0^{(T)} + \kappa \left[\frac{\gamma + \beta\lambda}{\beta + \lambda\gamma} \right] \alpha_0^{(T)}, \\ \alpha_2^{(T)} &= (1 - \kappa) \left[\frac{(2 + \lambda)\gamma + \beta\lambda}{2(\beta + \lambda\gamma)} \right] \alpha_0^{(T)} + \kappa \left[\frac{(1 + \lambda)\gamma}{\beta + \lambda\gamma} \right] \alpha_0^{(T)},\end{aligned}\tag{17}$$

for $\kappa \in [0, 1]$, where $\kappa = 0$ implies neither generic has a first mover advantage (corresponds to the point on the 45° line) and $\kappa = 1$ means that the first generic has the maximum first mover advantage (allowing negative values of κ up to negative one allows for modeling second mover advantage but is

not of interest to us). On the other hand, the λ parameter sets the relative market size between the generic and branded segments of the market, and is determined by the WTP for generics relative to that of the branded product.

To compare outcomes (prices, quantities, and profits) across market structures (triopoly, monopoly or a duopoly), we impose the restriction that the total (potential) market size under the three structures is the same. Thus, we assume that the introduction of generics to the market does not increase the potential set of patients *per se*, meaning no new patients exist that can use the drug, though in equilibrium the actual number of patients that consume the drug may increase due to lower prices if existing patients were originally priced out and hence, for comparison, we impose

$$a_0^{(T)} + a_1^{(T)} + a_2^{(T)} = (1 + \lambda)a_0^{(T)} = a_0^{(D)} + a_1^{(D)} = a_0^{(M)}. \quad (18)$$

In turn, this implies that if the WTP for the branded drug in a monopoly is normalized to $\alpha_0^{(M)} = \tilde{\alpha}_0$, then in a triopoly,

$$\alpha_0^{(T)} = \frac{\beta + \gamma\lambda}{\beta(1 + \lambda)}\tilde{\alpha}_0 \quad (19)$$

i.e., the willingness-to-pay for the branded drug would be lower in a triopoly.²² In a duopoly, we allow the solo generic to have the same WTP as the first generic entrant in a triopoly, while making sure that the potential market size is constant. Specifically, let $\alpha_1^{(D)} = \alpha_1^{(T)}$ and set $\alpha_0^{(D)}$ such that (18) holds, which gives

$$\alpha_0^{(D)} = \frac{\beta + \gamma}{\beta}\tilde{\alpha}_0 - \alpha_1^{(D)} \quad \text{and} \quad \alpha_1^{(D)} = \alpha_1^{(T)}. \quad (20)$$

D.3. Price Competition. We model competition as Nash-Bertrand with differentiated products. Consider first the case where all three firms engage in a price competition, a competitive triopoly, i.e. configuration T0, and where there are no authorized generics. Then the profit maximizing equilibrium prices are determined by

$$p = c + \Omega^{-1}D(p_0, p_1, p_2) \text{ and where } \Omega \text{ is a three by three matrix such that} \quad (21)$$

$$\Omega_{ij} = -O_{ij} \frac{\partial D_j(\cdot)}{\partial p_i}.$$

In the equation above, O_{ij} are terms of the ‘ownership’ matrix, set equal to the identity matrix for the base-line case of a competitive triopoly [Nevo, 1998]. Triopoly outcomes in other cases (authorized generics) are computed similarly but by adjusting the terms of the ownership matrix. For instance, when the branded firm launches an AG via the first challenger and competes with the second challenger (T1), equilibrium prices are computed by setting the off-diagonal terms for

²²This follows from the inverse demand function in monopoly defined equivalently as $p_0 = \alpha_0^{(M)} - \beta q_0$, which gives the demand function as $q_0 = a_0^{(M)} - b_0^{(M)} p_0$ where $a_0^{(M)} = \alpha_0^{(M)}/\beta$ and $b_0^{(M)} = 1/\beta$ and then using substitution and simplification from earlier relations. Note that as long as $\beta > \gamma$, the WTP of the branded in triopoly is always lower than that in monopoly for all $\lambda > 0$. Further, it is decreasing function of λ .

the branded and the first generic equal to one in the ownership matrix to allow for joint profit maximization between these two firms.²³ In a duopoly, the pricing equation is similar except that dimensionality is reduced by one, and the ownership matrix is either equal to an identity matrix (in the D0 competitive duopoly case) or all terms are equal to one (in the D1 duopoly where the branded firm has launched an AG). Computation of equilibrium prices allows computation of quantities and firm profits.

²³Similarly, our model allows for a fully collusive triopoly, i.e., the branded firm launches two AGs, and is in a ‘T2’, all terms of the ownership matrix are set to one.
