Hitotsubashi University
Institute of Innovation Research

Pharmaceutical Patents and Generic Entry Competition: A New View on the Hatch-Waxman Act

Jiangyun(Yunyun) Wan
Kaz Miyagiwa

IIR Working Paper WP#15-18

Revised in Jul. 2016

Institute of Innovation Research
Hitotsubashi University

Tokyo, Japan
http://www.iir.hit-u.ac.jp

Pharmaceutical patents and generic entry competition: A new view on the Hatch-Waxman Act

Kaz Miyagiwa[*] and Yunyun Wan[†]

July 11, 2016

Abstract

This paper presents a model of generic entry competition that captures salient features of the Hatch-Waxman Act, enacted with two incompatible objectives of restoring incentives for new drug development and stimulating generic competition. For the latter objective, Hatch-Waxman eases FDA approval procedures for generics, and encourages patent challenges by granting marketing exclusivity rights to the first patent challenger. We show that the two entry-promoting measures complement each other in restoring innovative incentives. While branded and generic drug manufacturers' interests are opposed in general, there are cases in which they all benefit from marketing exclusivity. The welfare implications are also examined.

Keywords: pharmaceutical innovation and patents, generic entry competition, Hatch-Waxman, marketing exclusivity,

JEL Classification Codes: I18, K23, L13

Corresponding author: Kaz Miyagiwa, Department of Economics, Florida International University, Miami, FL 33199, U.S.A. Email: kmiaygiw@fiu.edu

---

[*] Department of Economics, Florida International University, U.S.A.

[†] Institute of Innovation Research, Hitotsubashi University, Japan

1. Introduction

This paper presents a dynamic model of pharmaceutical patents and generic entry competition under the Hatch-Waxman Act.[1] Hatch-Waxman was enacted in 1984 with the twin objectives of restoring incentives for new drug development and stimulating generic entry. The first objective reflects the concern that incentives to develop new drugs may not be sufficient in the United States.[2] The reason is that, while a patent grants an inventor twenty-year exclusive rights to his or her innovation, for pharmaceutical innovations the effective patent life (time remaining after a product launch) is shortened by 10 to 12 years, as the FDA (U.S. Food and Drug Administration) review process requires lengthy and extensive preclinical and clinical trials for efficacy and safety.

Hatch-Waxman has addressed this concern by extending the patent life for innovative drugs by up to five years. However, extended protection of branded drug patents delays generic entry and keeps the costs of medicines high.[3] To cope with this issue, Hatch-Waxman contains two provisions designed to promote generic entry. First, Hatch-Waxman allows generic drug companies to use the clinical data the branded drug manufacturer has already submitted to the FDA regarding the branded drug's safety and efficacy.[4] Generic drug manufacturers are now required to submit ANDA (Abbreviated New Drug Application) only to demonstrate the bioequivalence of their drugs to the original branded drug. It is estimated that a preparation of an ANDA costs around $1 million (Hemphill and Lemley, 2011, footnote 13). In contrast, the average cost to develop a new drug and obtain marketing approval was estimated to be $231 million (in 1987 dollars) and $802 million (in 2000 dollars).[5] Thus, easing of FDA approval procedures constitutes enormous cost reductions for generic entry.

---

[1] See Bulow (2004) for an overview of key Hatch-Waxman provisions. Morton and Kyle (2012) review competition and regulations of the pharmaceutical markets.

[2] See Mossinghoff (1999) for the details of the FDA review process.

[3] Iizuka (2012) uses micro panel data from the Japanese pharmaceutical markets to demonstrate the sensitivity of generic entry to the prescription pattern, especially, to physicians' failure to internalize cost differences offered by generics.

[4] This is subject to "data exclusivity," which prohibits use of the branded drug's data for five years since the launch of branded drugs. For details, see http://www.fda.gov/Drugs/DevelopmentApprovalProcess/SmallBusinessAssistance/ucm069962.htm.

[5] These estimates are taken from DiMasi et al. (1991) and DiMasi et al. (2003), respectively. Adams and Brantner (2006) put the figures between $500 and $2,000 million.

To further promote generic entry, Hatch-Waxman also reward the first generic firm that successfully challenges the patent allegedly covering the branded drug with a marketing exclusivity right for 180 days, during which no other generics are allowed to compete.

There is little disagreement regarding the effect of Hatch-Waxman on generic entry. The generic drug share of U.S. prescription drug volume has increased from 18.6% in 1984 to over 40% in 1996 and to 74.5 % in 2009 (Rosenthal 2002, Berndt and Aitken 2010). Patent challenges have also increased significantly. Hemphill and Sampat (2012), for example, find that two-thirds of the samples they examined have had patent challenges and that patent challenges have shortened the brand-name drug patent life by nearly four years on average.[6] Grabowski and Kyle (2007) obtain similar results.

If marketing exclusivity is successful in inducing generic entry and patent challenges, however, it must also harm the patent holder. However, marketing exclusivity also benefit the branded drug manufacturer by inhibiting competition among generics for 180 days. Because it has both the pro-competitive and the anti-competitive effect, it is not entirely clear what the net effect of marketing exclusivity. Thus, the primary objective of the present paper is to investigate the role marketing exclusivity plays for incentive restoration.

Second, since Hatch-Waxman encourages generic entry not only with marketing exclusivity but also with easing of marketing approval procedures for generics, the relationship between these two entry-promotion measures is not obvious. If easing of FDA approval procedures result in substantial generic entry cost reductions, would it be possible that marketing exclusivity is redundant and even harmful to society because of its anti-competitive effect? Our formal analysis can address this and other welfare-related issues.

We address these questions in a dynamic model of generic entry competition that captures many features of Hatch-Waxman. The model has one incumbent (branded drug company) and two potential entrants (generic companies). Time runs continuously from zero to infinity. The

[6] According to Hemphill and Sampat (2012), the pioneer drug has an average effective patent life of 15.9 years but actual patent life of 12.2. The authors attribute the difference of nearly four years due to patent challenges.

incumbent's patent expires at a given date. We assume that both generic firms can enter profitably after the patent expires. This assumption allows us to focus on when generic firms challenge the patent.

Formally, two generic firms play a game of timing, simultaneously deciding, at each instant of time, whether to enter the market, conditionally on not having done so to date. Entry requires a fixed entry cost representing all the expenses necessary for obtaining FDA approval. If entry occurs before patent expiration, the incumbent files suit for patent infringement, which triggers an automatic stay of FDA approval for 30 months, ensuring that the incumbent remains a monopoly. During stay a court deliberates on the validity of the patent and announces its decision at the end of stay. If the patent is found valid, generic firms enter when the patent expires, If the patent is invalid, marketing exclusivity is granted to the patent challenger, delaying entry by the second generic firm for 180 days.

We first solve the game with marketing exclusivity, and then turn to the counterfactual scenario, in which the FDA stops granting marketing exclusivity. Comparisons of the two versions of the model yield a number of interesting results. First, marketing exclusivity raises the profit to the branded drug company only if entry costs have been substantially reduced. Thus, Hatch-Waxman's two entry promotion measures – easing of approval procedures and granting of marketing exclusivity – have complementary effects. Second, the effect of marketing exclusivity also depends on the strength of the patent allegedly covering the branded drug. The weaker the patent, the more likely marketing exclusivity benefits the incumbent. Interestingly enough, although the benefit to the branded drug company generally comes at the expense of generic firms' profits, there are cases in which marketing exclusivity benefit all three firms.

Turning to the welfare results, we find that market exclusivity generally harms consumers. An exception occurs only if entry costs are so high that generic firm would not enter without marketing exclusivity. In such cases marketing exclusivity has a pro-competitive effect, inducing generic entry and improving consumer welfare. As for social welfare, which comprises consumer surplus and the industry net profit, it is difficult to draw general conclusions. We can show that social welfare is definitely lower with marketing exclusivity when entry costs are substantially low. However, our welfare calculus presupposes the existence of branded drugs. For drugs yet to be discovered, the prospect of greater profits due to marketing exclusivity can speed up discoveries of new drugs. When this R&D effect is taken into account, it is quite possible that marketing exclusivity benefits every one including consumers, though it is harmful to society in the short run.

The remainder of this paper is organized as follows. Section 2 describes the environment. Section 3 presents the model with marketing exclusivity. Section 4 examines the counterfactual scenario, in which marketing exclusivity is absent. Section 5 compares the results obtained in sections 3 and 4. Section 6 examines the welfare effect of marketing exclusivity. Section 7 applies the model to examine issues concerning "orphan drugs" and "me-too drugs." The final section concludes.

2. Model environment

In this section we describe the environment of our model, introduce key features of Hatch-Waxman, and show how we incorporate them into our model. The model has three firms: an incumbent and two potential entrants. The incumbent is a branded drug manufactures and a patent holder. Two potential entrants are generic drug companies, which must decide when to enter. Entry requires incurrence of the one-time fixed cost f (> 0), which stands for all the cost incurred for FDA approval.

With two potential entrants there can be three market structures: monopoly (M), duopoly (D) or triopoly (T). The market demand is stationary over time. Let Π^J denote the incumbent's present-discounted sum of profits when the market structure remains J (= M, D, T), and let π^J denote a generic firm's discounted sum of profits under market structures J (= D, T). Note that these profits are stock values. The corresponding momentary profits equal $r\Pi^J$ and $r\pi^J$, where r is the instantaneous rate of interest. These profits satisfy the following standard conditions.

Assumption 1: (A) $\Pi^M > \Pi^D > \Pi^T$ and $\pi^D > \pi^T$. (B) $\Pi^D \geq \pi^D$ and $\Pi^T \geq \pi^T$.

(C) $\pi^T > f$. (D) $2\pi^T > \pi^D$.

Assumption 1A says that profits decrease as competition intensifies.[7] Assumption 1B introduces possible asymmetry between the incumbent and the entrants, as many consumers tend to prefer a brand to generics even though generics are equally efficacious to and yet less expensive than a branded drug.[8] Assumption 1C implies that both generic firms will enter when the patent expires, if they have not done so already. This assumption allows us to focus on the central question of this paper: generic entry competition before patent expiration. Finally, assumption 1D implies that two firms cannot profitably merge into a single firm if an industry has three (or more) firms. This well-known property gives rise to the merger paradox, a subject that has received much attention in the industrial organization literature.[9]

Time, indexed by T, runs continuously from zero to infinity.[10] At T = 0 the incumbent holds the patent on a branded drug. The patent is assumed to expire at a given time P > 0. Then assumption 1C says that generic firms enter without infringement at P, if they have not done so by the. If a generic firm enters before P, however, there is a possibility that it infringe the incumbent's patent. Therefore, a generic firm is required to certify that the incumbent's patent is either invalid or not infringed by its generics. Upon receiving notice that such a certification – called a paragraph IV certification – has been submitted, the incumbent has forty-five days to decide whether to file

[7] Frank and Salkever (2004) find increases in branded drug prices following generic entry, but this is not unexpected if users view branded and generic drugs as vertically differentiated.
[8] The first entrant may have a first-mover advantage vis-à-vis the second entrant. However, this inter-generic asymmetry does not influence our results qualitatively as long as the incumbent's profit is unaffected by the order of generic entry. Thus, we assume symmetry between entrants, to keep the notation compact.
[9] The merger paradox pertains to the difficulty in accounting for profitable merger with standard oligopoly models. The seminal work is Salant, Switzer and Reynolds (1983), who showed that merger is unprofitable unless at least 80 percent of the firms in the industry participate in it. See, e.g., Belleflamme and Peitz (2015) for a review of the literature.
[10] T = 0 does not correspond to the branded drug launch time but represents some time after "data exclusivity" expired.

infringement suit against the generic firm. To keep things simple, we assume however that the incumbent makes its decision immediately.

If the incumbent files suit, there is a 30-month automatic stay of FDA approval, during which no generics can be marketed.[11] In our analysis, C denotes the length of this period. During this period a court deliberates and announces its decision at the end.[12] It is assumed that a court will find the patent valid with probability $\alpha \in (0, 1)$.[13,14] If the patent is found valid, the generic firms must delay entry till P.[15] On the other hand, if the patent is invalid, the FDA awards a 180-day marketing exclusivity privilege to the patent challenger. In our analysis X denotes the length of the exclusivity period. If it happens that both generic firms challenge the patent simultaneously, it is assumed that both firms get a marketing exclusivity right with equal probabilities.[16]

If the incumbent does not file suit, the challenger's ANDA is approved and its generic is launched immediately.[17] Even in this case, the incumbent still reserves the right to file infringement suit later. However, "if the patent owner were to wait to sue until the generic manufacturer began marketing its drug weeks or months later, it would be able to obtain a stay only if it convinced a court to issue a preliminary injunction in accordance with the familiar criteria, including proof of a likelihood of success on the merits. In the meantime, the patent owner would be fairly certain to suffer some immediate harm in the form of lost profits" (Cotter 2004, page 1079). Because the action triggers a thirty-month automatic stay of approval, the patent holder has a very strong incentive to file suit even if its case is weak. We therefore assume in this paper that the incumbent

[11] Actually, a stay on FDA approval lasts for 30 months or until a final judicial finding is announced regarding the patent's validity, or until the patent expires, whichever comes first.
[12] This is true if litigation starts at $T < P - C$. If litigation starts later than $P - C$, the patent expires before a court settles, allowing generic entry at P.
[13] The stochastic nature of patent litigation is much emphasized in the economic and legal studies literature. For example, Lemley and Shapiro (2005) write: "When the patent holder asserts the patent against an alleged infringer, the patent holder is throwing the dice. If the patent has been found invalid, the property right has been evaporated" (p. 75).
[14] According to a 2002 FTC study entitled *Generic Drug entry Prior to Patent Expiration*, generic applicants prevailed 73 per cent of the cases in which courts have resolved the patent dispute. This suggests a relatively small value of α on average.
[15] That is to say the firms respect the court finding and do not challenge the patent again; see Choi (1988) for the same assumption.
[16] We follow Choi (1988) in assuming here that two cases are litigated in a single suit.
[17] In reality, the FDA takes about two years to review and approve generic drug applications (Mossinghoff 1999). To focus on patent challenges, however, we assume immediate FDA approval.

always file suit when there is a patent challenge. We, in Appendix A, justify this assumption by proving that doing so is the dominant strategy for the incumbent.

We assume next that filing suit entails vanishingly small legal fees. In fact, litigation expenses, usually at several million dollars, are dwarfed by the profits branded drugs can generate. Especially, if branded drugs turn out to be blockbusters, they can generate revenues of more than a billion dollars per year.[18] If legal fees are large enough, the incumbent will choose not to file suit. We examine the implications of non-negligible legal fees for our analysis in Appendix A.

Throughout our analysis we also assume that the following are true:

Assumption 2: (A) $P > C + X$. (B) $C > X$

Suppose one generic firm challenges the patent at $T = 0$. Then, a court will announce its finding at $T = C$. If the patent is invalid, the other generic firm cannot enter before $T = C + X$. If assumption 2A does not hold, the patent expires before the exclusivity period ends, preventing entry by the second generic firm before patent expiration. This possibility is ruled out by Assumption 2A. Assumption 2B just restates the fact that X stands for 6 months (180 days) while C stands for 30 months. It is worth emphasizing at this point that C and X are lengths of intervals while P is a fixed date.

3. The Hatch-Waxman Act

We begin with the analysis of Hatch-Waxman with marketing exclusivity. Since the profits Π^J and π^J and the entry cost f are independent of generic entry dates, our model can be presented as a simple timing game, in which two generic entrants decide when to enter at each instant in time,

[18] See http://www.bbc.com/news/business-28212223.

conditionally on not having done so.[19] In solving the model we pay attention to pure-strategy subgame-perfect equilibrium outcomes.[20]

Suppose that one generic firm enters at $T_1 \geq 0$. We call this firm the leader and the other firm the follower. We begin by considering the follower's optimal entry decision time T_2, given T_1. If the patent is found invalid at $T_1 + C$, the follower enters at $T_2 \geq T_1 + C + X$, receiving the present discounted sum of operating profits

$$\int_{T_2}^{\infty} e^{-rT} r\pi^T \, dT - e^{-rT_2} f = \left(e^{-rT_2} - e^{-rP}\right)\pi^T - e^{-rT_2} f + e^{-rP}\pi^T .$$

Here we find it convenient to use the shorthand notation: q = exp (-rQ), where q and Q are dummies. Using this notion, we can express the above profit as $(t_2 - p)\pi^T - t_2 f ++ p\pi^T$, where t_2 = exp (- rT_2) and p = exp (-rP). Since this profit expression is increasing in t_2 (decreasing in T_2), the follower's best response is to enter as soon as the exclusivity ends. Thus, the follower enters at $T_1 + C + X$, provided that $T_1 + C + X < P$, earning the profit $(t_1cx - p)\pi^T - t_1cxf ++ p\pi^T$, where t_1cx = exp[- r(T_1 + C + X)]. If the patent is valid, the follower enters at P, earning the profit $p\pi^T - pf$.

Now, let $F(t_1)$ denote the expected present-discounted sum of net profits the follower receives over the interval [0, P] when the leader enters at T_1 and the follower chooses the entry date optimally. When the leader chooses T_1 the follower does not know what a court will find, so

(1) $$F_1(t_1) = (1 - \alpha)[(t_1cx - p)\pi^T - t_1cxf] - \alpha pf. \qquad (T_1 < P - C - X)$$

To understand this equation, note first that by assumption 1C both generic firms enter at P if they have not done so. Thus, the total profits $p\pi^T$ accrue to each generic firm from time P on, regardless of what happens before P. For this reason we can ignore those profits for our analysis without loss of generality. Thus, the first term on the right-hand side of (1) represents the sum of net profits the follower receives when the patent is invalid, whereas the second term equals the the net profit when the patent is found valid.

[19] See Katz and Shapiro (1987) and Fudenbeg and Tirole (1990).
[20] See Hendricks, Weiss and Wilson (1988) and Simon and Stinchcombe (1990) for details of timing games.

It is also conceivable that the leader challenges the patent at $T_1 \leq P - C - X$. In this case, the patent expires before the exclusivity period ends. Therefore, the follower enters at P, receiving the net profit equal to

$$F_2(t_1) = -pf. \qquad (P - C - X \leq T_1)$$

Turning to the leader's profits, we again ignore the post-expiration profit $p\pi$, and focus on its profit over the interval [0, p]. Let $L(t_1)$ denote the expected sum of net profits to the leader over that interval, if the leader enters at T_1 and expects the follower to behave optimally. From the preceding paragraph, if the leader enters at $T_1 < P - C - X$, the follower enters at $T_1 + C + X$. Therefore,

$$L_1(t_1) \quad = -t_1f + (1 - \alpha)[(t_1c - t_1cx)\pi^D + (t_1cx - p)\pi^T] \qquad (T_1 < P - C - X).$$

This equation says that the leader receives the duopoly profit during the exclusivity period and the triopoly profit in the post-exclusivity period, provided that the patent is invalid. If $P - C - X \leq T_1 < P - C$, the patent expires before the exclusivity period ends. Since exclusivity is forfeited at P, there is duopoly from $T_1 + C$ to P, with the follower entering at P. Thus, the leader's profit equals

$$L_2(t_1) \quad = -t_1f + (1 - \alpha)(t_1c - p)\pi^D \qquad (P - C - X \leq T_1 < P - C).$$

Lastly, if the leader enters at $T_1 > P - C$, the patent expires before FDA stay is expected to end. As both generic firms market their products at P, the leader's profit is

$$L_3(t_1) \quad = -t_1f \qquad (P - C \leq T_1 \leq P).$$

We also characterize the profits resulting from simultaneous generic entry. If both enter at T_1, each generic firm is equally likely to receive a marketing exclusivity right, so we can write, $B(t_1)$, the present discounted sum of profits to each generic firm, by replacing π^D with $\pi^D/2$ in $L(t_1)$:

$$B_1(t_1) = -t_1f + (1 - \alpha)[(t_1c - t_1cx)\pi^D/2 + (t_1cx - p)\pi^T] \qquad (T_1 < P - C - X)$$

$$B_2(t_1) \quad = -t_1f + (1 - \alpha)(t_1c - p)\pi^D/2 \qquad (P - C - X \leq T_1 < P - C)$$

$$B_3(t_1) \quad = -t_1f \qquad (P - C \leq T_1 \leq P)$$

One can verify that L, F, and B are continuous everywhere between T = 0 and P.

The model yields a large number of taxonomical cases, depending on parameter values. To eliminate some uninteresting cases we assume:

Assumption 3:

(A) $L_1(1) > L_3(p) = -pf$

(B) $(1 - cx)\pi^T > (c - cx)\pi^D$.

Assumption 3A says that a leader's profit is greater at T = 0 (t_1 = 1) than at P. If this assumption does not hold, no firm wants to lead, and hence there are no patent challenges. Assumption 3B says that the present discounted sum of tripoly profits earned from T = 0 to C + X is greater than the duopoly profit earned from C to C + X. This is not an unreasonable assumption because C +X is about 36 months long while X is only 6 month long.

Substituting for $L_1(1)$ and rearranging, we can express this assumption as

(2) $$f < \Gamma(\alpha) \equiv \frac{(1-\alpha)[(c-cx)\pi^D + (cx-p)\pi^T]}{1-p}.$$

It is also easy to prove that:

(3) $F(1) > F(p) = -pf$.

This result is intuitive, because the follower will enter at T = 0, only if that is more profitable than entering at P.

We illustrate the graphs of L and F in figure 1. To understand how they are drawn, note first that, as T runs from zero to infinity, t = exp(-rT) runs from 1 to zero. Next, differentiation yields

$$dL_1/dt_1 = -f + (1-\alpha)[(c - cx)\pi^D + cx\pi^T] \qquad (T_1 < P - C - X).$$

$$dL_2/dt_1 = -f + (1-\alpha)c\pi^D \qquad (P - C - X \le T_1 < P - C)$$

$$dL_3/dt_1 = -f \qquad (P - C \le T_1 < P)$$

To determine the sign of dL_1/dt1, note that

$$(1-\alpha)[(c-cx)\pi^D + cx\pi^T] - \Gamma(\alpha)$$

$$= [(1-cx)\pi^T - (c-cx)\pi^D](1-\alpha)p/(1-p) > 0$$

where the inequality follows form assumption 3B. Since $f < \Gamma(\alpha)$ by (2), this proves that $dL_1/dt_1 > 0$, that is, $dL_1/dT_1 < 0$. This explains why the graph of L_1 declines as t_1 runs from 1 to p/cx. We also have that

$$dL_2/dt_1 - dL_1/dt_1 = (1-\alpha)cx(\pi^D - \pi^T) > 0.$$

Hence.

$$dL_2/dt_1 > dL_1/dt_1 > 0 > dL_3/3.$$

Thus, L declines between t = 1 and t = p/c and slopes up between p/c and p, as in figure 1.

As for the graph of $F(t_1)$, differentiation yields

$$dF_1/dt_1 = (1-\alpha)cx(\pi^T - f) > 0 \qquad (T_1 < P - C - X)$$

$$dF_2/dt_1 = 0, \qquad (P - C - X \leq T_1)$$

This explains why F slopes down between t = 1 and t = p/cx (T = P – C – X) and then remain constant at – pf.

Turning to the graph of B, we know that $L_1 > B_1$, $L_2 > B_2$ and $L_3 = B_3$ by definition. We can also show that $dB_2/d\tau_1 > dB_2/d\tau_1$, although there is no guarantee that they are downward sloping. However, as we will see shortly, the slopes of B do not affect our results.

We next calculate the values of L, F and B at t = 1 (T = 0). Straightforward calculations yield

(4) $$L(1) - F(1) = (1-\alpha)[(c - cx)\pi^D] - [1 - \alpha p - (1-\alpha)cx]f$$

(5) $$B(1) - F(1) = (1-\alpha)[(c - cx)\pi^D/2] - [1 - \alpha p - (1-\alpha)cx]f.$$

Define the function $\phi : (0, 1) \rightarrow \mathbb{R}_+$ by

$$\phi(\alpha) \equiv \frac{(1-\alpha)(c-cx)\pi^D}{1-\alpha p-(1-\alpha)cx}.$$

Then, (4) and (5) imply the following results:

$$f \leq \phi(\alpha)/2 \quad \Rightarrow \quad L(1) > F(1) \text{ and } B(1) > F(1)$$

$$\phi(\alpha)/2 < f \leq \phi(\alpha) \quad \Rightarrow \quad L(1) \geq F(1) \text{ and } B(1) \leq F(1)$$

$$f > \phi(\alpha) \quad \Rightarrow \quad L(1) < F(1) \text{ and } B(1) < F(1)$$

Since L(1) > B(1), these inequalities lead to the following three possibilities.

Case 1: $f \leq \phi(\alpha)/2$

In this case, we have that L(1) > B(1) ≥ F(1). This is the case depicted in figure 1. (3) implies that B(1) > - pf. This case occurs when the entry cost is sufficiently low so each generic firm prefers to enter at T = 0, even if the rival firm also enters at T = 0. On the other hand, since F(1) ≤ B(1), if both enters, no one wishes to deviate. Thus, there is a pure strategy equilibrium in which both firms enter at T = 0. Figure 1 shows the case in which B_1 slopes down. Even if it slopes upward, B_2 must slope down. Hence, the graph of B stays between those of L and F from t = 1 to t = p/cx. Thus, the shape of B does not affect the equilibrium outcome.

Case 2: $\phi(\alpha)/2 < f \leq \phi(\alpha)$

In this case, we have that L(1) ≥ F(1) > B(1). This case is illustrated in figure 2. Since L is declining and L(1) > - pf, entry at T = 0 dominates entry as any other time. Thus, there is at least one generic firm that enters at T = 0. Next, F(1) > B(1) implies that following instead of entering simultaneously is the other firm's best response. The equilibrium outcome therefore has the leader entering at T = 0 with probability 1 and the follower entering at C + X if the patent is invalid and at P if the patent is valid. It is verified that the equilibrium outcome is unaffected by the shape of B.

Case 3: $f > \phi(\alpha)$

This case implies that F(1) > L(1) > B(1) and is illustrated in figure 3. Since the follower's profit is greater than the leader's at T = 0, both firms prefer to follow than lead at T = 0. The consequence is a war of attrition, which continues until the profit from following equals that of leading.[21] Setting $L_1(t_1^*) = F_1(t_1^*)$ yields

$$t^*_1 = \frac{\alpha pf}{f[1-(1-\alpha)cx]-(1-\alpha)(c-cx)\pi^D} > 0.^{22}$$

[21] Hendricks, Weiss and Wilson (1988) discuss timing games resulting in wars of attrition.
[22] The denominator is positive given the fact that $f > \phi(\alpha)$ in region 3.

From the properties of the profit functions, t_1* must be between 1 and p/cx. Thus, in equilibrium the leader enters at T_1* = - r ln (t_1*) and the follower enters at T_1* + C + X if the patent is found invalid and at P if the patent is valid. Since $L_1(t_1^*) = F_1(t_1^*)$, a war of attrition results in rent equalization, as is familiar in the literature (e.g., Fudenberg and Tirole 1985). As in other cases, the slope of B does not affect the equilibrium outcome.

In figure 4 we represent the above three cases in (α, f) space. Note that φ(α) is decreasing and convex downward, since

$$\phi'(\alpha) = \frac{-(c-cx)\pi^D}{[1-\alpha p-(1-\alpha)cx]^2} < 0 \text{ and } \phi''(\alpha) = \frac{2(c-cx)(c-p)\pi^D}{[1-\alpha p-(1-\alpha)cx]^3} > 0$$

φ(α) is downward sloping because of the fundamental tradeoff between α and f. The probability α denotes the likelihood that a court will find the patent valid, and hence it is an inverse measure of expected profits for generic firms. The higher this probability, the smaller the expected profit and hence the less incentive to challenge the patent. Restoration of incentives to challenge the patent requires a decrease in the entry cost.

We summarize the findings of this section in

Proposition 1. Fix α. The following results hold in the presence of marketing exclusivity.

(A) If $f \in (0, \phi(\alpha)/2]$, both firms enter (challenge the patent) at t = 0.

(B) If $f \in (\phi(\alpha)/2, \phi(\alpha)]$, one firm challenges the patent at t = 0 while the second one enters at t = C + X (when the exclusivity period ends) with probability (1 - α) and at t = P with probability α

(C) If $f \in (\phi(\alpha), \Gamma(\alpha))$, there is a war of attrition. One firm challenges the patent at T* > 0 while the other firm enters at T* + C + X with probability (1 - α) and at P with probability α.

(D) If f ≥ $\Gamma(\alpha)$, there are no patent challenges.

We end this section with the following remark. As we saw in proposition 1 and figure 4, the equilibrium outcomes hinge crucially on the generic entry cost f and the strength of the patent (probability α). Prior to Hatch-Waxman, entry costs were presumably very high, since "there were 150 drugs that went off-patent (after 1962), but for which there were no generics because generic companies simply would not spend the time and money doing the clinical trials to get to the market" (Mossinghoff 1999). Thus we are likely to have been in the region above the schedule Γ(α) in figure 4. As mentioned in the introduction, after Hatch-Waxman generic firms are required only to prove the bioequivalence of their drugs to the branded drug. Furthermore, trials data of the branded drug, previously kept as trade secrets, are made available to generic drug manufacturers after five years of data exclusivity. The fact that there are so many generics in the market today implies that Hatch-Waxman has been quite successful in reducing entry costs for generics.

4. Counterfactual: Hatch-Waxman without marketing exclusivity

In this section we consider the counterfactual scenario: Hatch-Waxman without marketing exclusivity. The analysis closely follows that of the preceding section. Suppose that the leader challenges the patent at T_1 (< P – C). The incumbent files suit, resulting in stay of FDA approval until a court's announcement at T_1 + C. If the patent is found invalid, the leader markets its drug immediately, but so does the follower in the absence of marketing exclusivity. Since the follower sinks the entry cost whenever it enters, the follower's expected present-discounted sum of profits equals

$$\hat{F}_1(t_1) = (1-\alpha)[(t_1c - p)\pi^T - t_1cf] - \alpha pf \qquad (T_1 < P - C)$$

Alternatively, if the leader enters at T_1 (≥ P – C), the patent expires before a judicial announcement. Hence, the follower enters at P, receiving the net profit

$$\hat{F}_2(t_1) = -pf \qquad (P - C \leq T_1).$$

Turn next to the leader's profits. Given the follower's optimal strategy, if the leader enters at T_1 (< P – C), the leader's present discounted sum of profits equals

$$\hat{L}_1(t_1) = -t_1f + (1-\alpha)(t_1c - p)\pi^T \qquad (T_1 < P - C).$$

Alternatively, if the leader enters sometime between P – C and P, the patent expires before stay ends. Hence, there is no marketing of generics before P. Thus, the leader's net profit equals

$$\hat{L}_2(t_1) \; = -t_1 f \qquad (P - C \leq T_1 \leq P)$$

It is straightforward to show that

$$\hat{F}_1(t_1) > \hat{L}_1(t_1) \text{ and } \hat{F}_2(t_1) = \hat{L}_2(t_1).$$

In particular, we have that

$$\hat{L}_1(1) - \hat{F}_1(1) = -f\,[1 - (1-\alpha)c - \alpha p] < 0.$$

Thus, the follower's profit never falls below the leader's, and strictly exceeds the latter at t = 1 (T = 0).

Lastly, we calculate the discounted sum of profits $\hat{B}$ obtainable if both generic firms enter simultaneously at T_1. In this case, since they both launch their generics at T_1 + C with probability (1 – α), the profits are identical to those of the leader's;

$$\hat{B}_1(t_1) = \hat{L}_1(t_1) \qquad (T_1 < P - C)$$

$$\hat{B}_2(t_1) \; = \; \hat{L}_2(t_1) \qquad (P - C \leq T_1 \leq P)$$

Further, it can be checked that $\hat{L}$, $\hat{F}$, and $\hat{B}$ are continuous.

We now depict the profit functions. Differentiation gives

$$d\hat{F}_1/dt_1 = (1-\alpha)c(\pi^T - f) < 0 \qquad (T_1 < P - C)$$

$$d\hat{F}_2/dt_1 = 0 \qquad (P - C \leq T_1 \leq P).$$

We can also show that $\hat{F}_1(1) > 0$. Thus, the graph of $\hat{F}$ declines till T = P – C (t = p/c) and remains constant at – pf thereafter, as shown in figure 5. Figure 5 also depicts the graph of $\hat{L}$. To see it is correctly drawn, differentiate $\hat{L}_1$ to obtain

$$d\hat{L}_1/dt_1 = -f \; + (1-\alpha)c\pi^T > 0,$$

where the inequality comes from the fact that $(1-\alpha)c\pi^T \; > \; \Gamma(\alpha) > f$, where the second inequality is by (2). Thus, the graph of $\hat{L}_1$ is sloping downward, as shown in figure 5 till t = p/c. Since, $d\hat{L}_2/dt_1 < 0$, $\hat{L}_2$ increases between p/c and p.

The last thing we need to check is the sign of the expression:

$$\hat{L}_1(1) - \hat{L}_2(\mathrm{p}) = \hat{L}_1(1) + \mathrm{pf} = (1-\alpha)(\mathrm{c}-\mathrm{p})\pi^{\mathrm{T}} - (1-\mathrm{p})\mathrm{f}.$$

If we define $\psi(\alpha)$ by

$$\psi(\alpha) \equiv \frac{(1-\alpha)(c-p)\pi^T}{1-p},$$

$f < \psi(\alpha)$ implies that $\hat{L}_1(1) > \hat{L}_2(p) = -pf$. The graph of L in figure 5 depicts this case. It follows that there is the time $\hat{T}_1 \in (0, P - C)$ such that

$$L_1(\hat{t}_1) = -pf,$$

where $\hat{t}_1 = \exp(-r\hat{T}_1)$. Suppose that one firm chooses to enter at every instant between some date between 0 and $\hat{T}_1$ and wait till P at each instant between $\hat{T}_1$ and P, provided that no one has entered before. Then, since $\hat{L} < -pf$ for $\in(\hat{T}_1, P]$, the other firm has no incentive to lead, once $\hat{T}_1$ is reached without any firm having entered. Then, since $\hat{L}(1) > -pf$, it is optimal for one firm to enter at T = 0. Thus, the model has a unique subgame-perfect equilibrium outcome, in which one firm enters at T = 0 (t = 1).

Consider now the alternative case. $f \geq \psi(\alpha)$ implies that $\hat{L}_1(1) \leq -pf$. Since $\hat{L}$ is increasing in t_1, the entire graph of $\hat{L}$ lies below the horizontal line at $-pf$.[23] On the other hand, the value of $\hat{F}$ never falls below $-pf$. Therefore, there is no incentive to lead at any time; both firms prefer to wait till P to enter.

Proposition 2: Without marketing exclusivity,

(A) if $f < \psi(\alpha)$, one firm challenges the patent at T = 0;

(B) if $f \geq \psi(\alpha)$, no firm challenges the patent.

To compare the results in propositions 1 and 2, we need to examine the relationships among the conditioning functions. We show, in Appendix B, that

$$\Gamma(\alpha) > \psi(\alpha) > \phi(\alpha)/2.$$

[23] The only exception occurs if $\hat{L}_1(1) = -pf$.

However, we cannot uniquely rank $\psi(\alpha)$ and $\phi(\alpha)$. To keep the analysis compact, for the remainder of the analysis we focus on the following case.

Assumption 4: $\psi(\alpha) \geq \phi(\alpha)$.

Appendix C discusses the case in which $\psi(\alpha) < \phi(\alpha)$. There, you will find that this alternative case does not add anything new to what we analyze below.

Under assumption 4, $\psi(\alpha)$ dissects region 3 into two subregions 3A and 3B as shown in figure 5. WE can now summarize the findings of this section.

Proposition 3: Under assumption 4, the following results hold without marketing exclusivity.
(A) In regions 1, 2, and 3A, where $f < \psi(\alpha)$, one generic firm challenges the patent at T = 0 while the other firm enters conditionally at T = C or at P.
(B) In region 3B, where $f > \psi(\alpha)$, both firms enters at P.

5. The effect of marketing exclusivity

We are in a position to examine the effect of marketing exclusivity. We will compare the equilibrium outcomes with and without marketing exclusivity for a given value of α, which determines $\phi(\alpha)$, $\psi(\alpha)$ and $\Gamma(\alpha)$].

Region 1: $f \leq \phi(\alpha)/2$.

In this region both generic firms challenge the patent at T = 0 with marketing exclusivity. The incumbent is a monopoly until C with certainty. From then on through P, it remains a monopoly with probability α; otherwise, it faces one generic competitor until C + X and two thereafter Thus, the incumbent's equilibrium profit through P equals:

(6) $$(1 - c)\Pi^M + \alpha(c - p)\Pi^M + (1 - \alpha)[(c - cx)\Pi^D + (cx - p)\Pi^T].$$

After P the incumbent earns the triopoly profits, $p\Pi^T$, regardless of what happens previously. We disregard this term without losing generality. The outcome without marketing is similar, except that both generic firms enter at C with probability $(1 - \alpha)$. Thus, the incumbent's profit equals

(7) $$(1 - c)\Pi^M + \alpha(c - p)\Pi^M + (1 - \alpha)(c - p)\Pi^T.$$

A comparison shows that the incumbent is better off with marketing exclusivity, as it faces one generic competitor during the exclusivity period.

As for the generic firms, with marketing exclusivity both firms enter at T = 0, eraning the combined profits:

$$2B_1(1) = -2f + (1 - \alpha)[(c - cx)\pi^D + (cx - p)2\pi^T].$$

By contrast, without marketing exclusivity one generic firm enters while the other does conditionally, yielding the combined profits equal:

$$\hat{L}(1) + \hat{F}(1) = (1 - \alpha)(c - p)2\pi^T - f - (1 - \alpha)cf - \alpha pf.$$

Hence,

$$2B_1(1) - [\hat{L}(1) + \hat{F}(1)] = (1 - \alpha)(c - cx)(\pi^D - 2\pi^T) - f[1 - (1 - \alpha)c - \alpha p] < 0;$$

the inequality hold since $\pi^D < 2\pi^T$ by assumption 1D and $1 > (1 - \alpha)c - \alpha p$. Thus, marketing exclusivity harms the generic firms.

Claim 1: In region 1, marketing exclusivity increases the incumbent's expected profit, and decreases the generic firms' profits.

Market exclusivity softens generic competition by delaying entry by the follower. Thus, it is obvious that the incumbent is better off with marketing exclusivity. To explain the negative effect on the generic firms, it is useful to consider their (gross) profits and the entry costs separately. Since there is triopoly after C + X with or without marketing exclusivity, the combined profits differ only during the preceding interval between C and C + X, during which they jointly earn $(c - cx)\pi^D$ with marketing exclusivity and $(c - cx)2\pi^T$ without it. Since $\pi^D < 2\pi^T$ by assumption 1D, the generic firms' profits are greater without marketing exclusivity. On the entry cost side, marketing exclusivity results in excess entry competition. Although one generic is marketed during the exclusivity period, both generic firms enter at T = 0, incurring the entry cost upfront. In contrast, without marketing exclusivity, only the leader enters at T = 0, while the follower incurs the entry cost later, at C with probability (1 – α) and at P with probability α. Delayed entry and avoidance of litigation risk means savings in interest. With a greater sum of gross profits and a smaller discounted sum of entry costs, the generic firms are better off without marketing exclusivity.

Region 2: $f \in (\phi(\alpha)/2, \phi(\alpha)]$

In this region, with or without marketing exclusivity only one generic firm challenges the patent at T = 0. Therefore, the incumbent's profits are exactly the same as in region 1. Marketing exclusivity benefits the incumbent. As for the generic firms, their combined profits equal

$$L_1(1) + F_1(1) = (1-\alpha)[(c-cx)\pi^D + (cx-p)2\pi^T - f \; - (1-\alpha)cxf - \alpha pf,$$

with marketing exclusivity, and

$$\hat{L}_1(1) + \hat{F}_1(1) = (1-\alpha)(c-p)2\pi^T - f \; - (1-\alpha)cf - \alpha pf \tag{8}$$

without marketing exclusivity. Hence,

$$[L_1(1) + F_1(1)] - [\hat{L}_1(1) + \hat{F}_1(1)] = (c-cx)(1-\alpha)(\pi^D - 2\pi^T + f). \tag{9}$$

Thus, t generic firms are better off with marketing exclusivity if and only if $\pi^D - 2\pi^T + f > 0$.

To understand the above result intuitively, we again look at the effect on the profit side and the entry cost side. The profit side is the same as in region 1; market exclusivity reduces the generic firms' combined earnings by $(c - cx)(\pi^D - 2\pi^T)$. On the entry cost side, in both scenarios the follower enters conditionally on a judicial decision. The difference is that with marketing exclusivity the follower enters later, at C + X rather than at C, collecting the difference in interest earnings equal to $(c - cx)f > 0$. Therefore, if $f > 2\pi^T - \pi^D$, the interest earnings exceeds the loss of profits, making the generic firms better off with marketing exclusivity. This has a more intuitive explanation. The condition $\pi^D - 2\pi^T + f > 0$ is written as $\pi^D - \pi^T > \pi^T - f$. The right-hand side of this inequality, $\pi^T - f$, represents the gain from entry to the follower, while the left-hand side, $\pi^D - \pi^T$, measures the harm done to the leader when the follower enters. Thus, if $\pi^D - 2\pi^T + f > 0$, the generic firms would be better off if the follower could pre-commit to delay entry in exchange for a side payment from the leader. Although such pre-commitment is of course untenable, marketing exclusivity serves as a commitment device by forcing the follower to delay entry for 180 days. As a result, the ex ante profits to the generic firms are greater with marketing exclusivity if $f > 2\pi^T - \pi^D$. Naturally, the incumbent always welcomes delayed generic entry.

Claim 2. In region 2:

(A) Marketing exclusivity increases the incumbent's profit.

(B) Marketing exclusivity increases a generic firm's profit if $f > 2\pi^T - \pi^D$.

Subregion 3A: $f \in (\phi(\alpha), \psi(\alpha)]$

By proposition 1, there is a war of attrition with marketing exclusivity, delaying a patent challenge till $T_1^* \in (0, P - C - X)$. Prolonged monopoly yields the incumbent's profit equaling

(10) $$(1 - ct^*_1)\Pi^M + \alpha(ct^*_1 - p)\Pi^M + (1 - \alpha)[t^*_1 (c - cx)\Pi^D + (t^*_1cx - p)\Pi^T].$$

Without marketing exclusivity, by contrast, the leader challenges the patent at T = 0, so the incumbent's profit is given in (7). It is evident that the incumbent is better off with marketing exclusivity. As for the generic firms, a war of attrition yields the combined profits

$$L_1(t^*_1) + F_1(t^*_1)$$
$$= (1 - \alpha)[(t^*_1c - t^*_1cx)\pi^D + (t^*_1cx - p)2\pi^T] - f[t^*_1 + (1 - \alpha)t^*_1cx + \alpha p].$$

Without marketing exclusivity, the leader enters at T = 0; the combined profits are given in (8). Hence,

(11) $$[L_1(t^*_1) + F_1(t^*_1)] - [\hat{L}(1) + \hat{F}(1)]$$
$$= (1 - \alpha)[(t^*_1 c - t^*_1cx)(\pi^D - 2\pi^T) - (c - t^*_1c)2\pi^T] + f[(1 - t^*_1) + (1 - \alpha)(c - t^*_1cx)].$$

The first term on the right-hand side is negative, so the gross profits are smaller with marketing exclusivity as before. On the other hand, the second term is positive so the discounted sum of entry costs is also smaller with marketing exclusivity. Hence the net effect is ambiguous in general. However, (11) is identical to (9) when evaluated at $t^*_1 = 1$, so the right-hand side of (11) equals $(c - cx)(1 - \alpha)(\pi^D - 2\pi^T + f)$ at $t^*_1 = 1$. Further, $[L_1(t^*_1) + F_1(t^*_1)]$ is increasing in t^*_1 (falling in T^*_1). Therefore, we conclude that

$$[L_1(t^*_1) + F_1(t^*_1)] - [\hat{L}(1) + \hat{F}(1)] < (c - cx)(1 - \alpha)(\pi^D - 2\pi^T + f).$$

Therefore, if $\pi^D - 2\pi^T + f \leq 0$, the generic firms are worse off with marketing exclusivity. Unlike in region 2, they may be worse off even if this inequality is reversed, because the gross profit is smaller than in region 2 due to a war of attrition.

Claim 3: In subregion 3A, marketing exclusivity increases the incumbent's profit. If $\pi^D - 2\pi^T + f \leq 0$, the generic firms are worse off with marketing exclusivity.

Subregion 3B: $f \in (\psi(\alpha), \Gamma(\alpha)]$

In this subregion, while the leader enters at $T_1^* < P$ with marketing exclusivity, there are no patent challenges without it. Thus, the incumbent is worse off with marketing exclusivity. Since $L_1(t^*_1) + F_1(t^*_1) > -2pf$, the generic firms are better off.

Claim 4: In subregion 3B, marketing exclusivity harms the incumbent and benefits the generic firms.

We summarize claims 1 – 4 in the next proposition:

Proposition 4: Hold α. Then:

(A) Marketing exclusivity increases the incumbent's expected profit when the entry cost is relatively low (regions 1, 2, 3A) and decreases it when the entry cost is high (region 3B).

(B) Marketing exclusivity harms the generic firms when entry costs are sufficiently low (region 1) but benefits them at high entry costs (subregion 3B). When the entry costs are in the intermediate range (regions 2 and 3A), the generic firms are worse off with marketing exclusivity if $\pi^D - 2\pi^T + f < 0$.

(C) When the entry costs are in the intermediate range (region 2), if $\pi^D - 2\pi^T + f > 0$, marketing exclusivity makes all three firms better off.

5. The welfare effect of marketing exclusivity

In this section we investigate the welfare implications of marketing exclusivity. We begin with introduction of new terms. CS^J (J = M, D, T) denote the present discounted sum of the consumer surpluses under market structure J. Adding the total industry profit (gross of entry costs), defines the corresponding gross social surpluses S^J, which increases as the market becomes more competitive; i.e.,

$$S^M (= \Pi^M + CS^M) < S^D (= \Pi^D + \pi^D + CS^D) < S^T (= \Pi^T + 2\pi^T + CS^T).$$

Finally, subtracting from S^J the discounted sum of generic entry costs yields social welfare.

From the preceding analyses we know that in regions 1, 2, and 3A marketing exclusivity delays a launch of the second generic, making consumers worse off. This is due to marketing exclusivity's anti-competitive effect. On the other hand, in region 3B marketing exclusivity induces entry before patent expiration, when there is no incentive to enter without it. Due to this pro-competitive effect consumers are better off with marketing exclusivity. Then, by proposition 4A, makes consumers worse off in regions 1, 2 and 3A and better of in region 3B.

Proposition 5: Consumers are worse off whenever the incumbent benefits from marketing exclusivity.

Welfare calculations are more complicated because we have to include the generic firms' net profits. With marketing exclusivity, one generic firm challenges the patent at $T_1 = 0$ in regions 1 and 2. The equilibrium social surplus equals

$$S(1, 2) = (1 - c)S^M + \alpha(c - p)S^M + (1 - \alpha)[(c - cx)S^D + (cx - p)S^T]$$

Note that this is the social surplus between T = 0 and P as we disregard, without loss of generality, the discounted sum of social surpluses accruing after the patent expires (= pS^T). Regions 1 and 2 differ only because of the difference in entry dates. In region 1, both firms enter at T = 0, so the social welfare is

$$W(1) = S(1,2) - 2f.$$

In region 2, the second firm enters, at the end of marketing exclusivity with probability (1 - α) and at P with probability α, so the welfare level equals

$$W(2) = S(1,2) - f - (1- \alpha)cxf - \alpha pf.$$

In region 3, there is monopoly until $T_1^* + C$. If the patent is valid, monopoly continues through P. Otherwise, there is duopoly until $T_1^* + C + X$, when the follower enters and triopoly begins. Thus, we have:

$$W(3) = (1 - t^*_1c)S^M + \alpha(t^*_1c - p)S^M + (1 - \alpha)[(t^*_1c - t^*_1cx)S^D + (t^*_1cx - p)S^T$$
$$- [t^*_1 + (1 - \alpha) t^*_1cx + \alpha p]f.$$

Without marketing exclusivity, we know that the leader enters at T = 0 in regions 1, 2, and 3A, so the welfare equals

$$\hat{W}(1, 2, 3A) = (1 - c)S^M + \alpha(c - p)S^M + (1 - \alpha)(c - p)S^T - [1 + (1 - \alpha)c + \alpha p]f.$$

In subregion 3B, there are no generic patent challenges, so the equilibrium social welfare is

$$\hat{W}(3B) = (1 - p)S^M - 2pf.$$

Equipped with these expressions, we are ready to make welfare comparisons. Again it is helpful to consider the effect on social surplus and on the entry cost separately.

Region 1: $f \in (0, (\phi(\alpha)/2]$

In this region,

$$W(1) - \hat{W}(1, 2, 3A) = (1 - \alpha)(c - cx)(S^D - S^T) - f\{1 - [\alpha p + (1 - \alpha)c]\} < 0.$$

Thus, welfare is smaller with marketing exclusivity. The first term is negative, making social surplus smaller with marketing exclusivity due to delayed entry by the second generic firm. On the entry cost side, as explained earlier, both firms incur the entry cost upfront with marketing exclusivity; without it, the second firm incurs the cost later and conditionally on a court's finding, which results in the savings of $\{1 - [\alpha p + (1 - \alpha)c]\}f > 0$ on interest.

Region 2: $f \in (\phi(\alpha)/2, \phi(\alpha)]$

In this regime,

$$W(2) - \hat{W}(1, 2, 3A) = (1 - \alpha)(c - cx)(S^D - S^T + f).$$

As in region 1, $(1 - \alpha)(c - cx)(S^D - S^T)$ measures the difference in social surplus, making marketing exclusivity inferior. However, unlike in region 1, the follower enters later with marketing exclusivity than without it, which generates the interest earnings of $(1 - \alpha)(c - cx)f$ under marketing exclusivity. Hence, social welfare is greater with marketing exclusivity if and only if

$$f > S^T - S^D.$$

In the preceding section we showed that if $f > \pi^T - \pi^D$ all firms are better off with marketing exclusivity (proposition 4C). Since $CS^T > CS^D$, $f > \pi^T - \pi^D$ implies $f > S^T - S^D$. Thus, $f > \pi^T - \pi^D$ is a sufficient condition for a welfare improvement under marketing exclusivity in region 2.

Region 3A: $f \in (\phi(\alpha), \psi(\alpha)]$

We have that

$$W(3) - \hat{W}(1, 2, 3A) = (1 - \alpha)[(c - t^*_1c)(S^M - S^T) + (t^*_1c - t^*_1c\,x)(S^D - S^T)]$$
$$+ [(1 - t^*_1) + (1 - \alpha)(c - t^*_1cx)]f.$$

The first term on the right is negative while the second is positive. The welfare impact of marketing exclusivity is in general ambiguous.

Region 3B: $f \in (\psi(\alpha), \Gamma(\alpha)]$

$$W(3) - \hat{W}(3B) = (1-\alpha)[(t^*_1 c - t^*_1 cx)(S^D - S^M) + (t^*_1 cx - p)(S^T - S^M)]$$
$$- [(t^*_1 - p) + (1-\alpha)(t^*_1 cx - p)]f.$$

Since both terms on the right are positive, the welfare effect of marketing exclusivity cannot in general be determined without ambiguity. This is despite the pro-competitive effect of marketing exclusivity that induces generic entry before patent expiration. Intuition is familiar; early entry yields a greater social surplus but also generates a higher discounted sum of entry costs.

We summarize the welfare implications of marketing exclusivity in the next proposition.

Proposition 6:

(A) In region 1, marketing exclusivity decreases social welfare.

(B) In region 2, marketing exclusivity increases social welfare if and only if $f > S^T - S^D$.

It is worth emphasizing that the preceding welfare calculus presupposes that the branded drug has already been discovered. For undiscovered drugs, however, whenever marketing exclusivity boosts the incumbent's profits, the incumbent has an incentive to devote more resources in development of new drugs. Since accelerated discoveries benefit consumers and generic firms alike, market exclusivity can potentially improve social welfare in the long run despite its negative "short-run" welfare impact.

7. Applications

In this section we apply our model to study the effect of marketing exclusivity in two special cases.[24] One concerns development of "orphan drugs" The other pertains to alleged misallocation of R&D resources to development of "me-too drugs." These are complex issues. We only touch on some of them about which we can say something using our model.

7.1. Orphan drugs

Branded drug companies have been criticized for not spending enough R&D resources to develop orphan drugs – drugs that treat small rare disease populations. To stimulate orphan drug research President Ronald Reagan signed the Orphan Drug Act into law in 1983. This Act contains a number of incentives for orphan drug development, including tax credits and grants for orphan drug research. In this paper we focus on the effect of seven-year market exclusivity for orphan drugs.[25] This exclusivity becomes effective immediately after FDA approval of orphan drugs, shielding manufacturers of orphan drugs from generic competition for seven years.

What this means in terms of our model is that the start of an entry competition game is postponed, as generic firms never enter during a seven year period of market exclusivity for orphan drugs. Thus, suppose that the game starts at $T_o > 0$ instead of at $T = 0$. In particular, we suppose that $T_o \geq P - (C + X)$, which implies that, if the leader challenges the patent at T_o, the follower cannot enter before patent expiration due to a 180 days of marketing exclusivity protecting the leader. Thus, the follower's net profit equals

$$F_2(t_1) = -pf.$$

Accordingly, the leader's profit functions are given by

$$L_2(t_1) = -t_1 f + (1-\alpha)(t_1 c - p)\pi^D \qquad (P - C - X \leq T_1 < P - C).$$

$$L_3(t_1) = -t_1 f, \qquad (P - C \leq T_1 \leq P)$$

where $t_1 \leq t_o = \exp(-rT_o) < 1$. Simultaneous entry yields these profit functions:

$$B_2(t_1) = -t_1 f + (1-\alpha)(t_1 c - p)\pi^D/2 \qquad (P - C - X \leq T_1 < P - C)$$

[24] We thank anonymous referees for suggesting these applications.
[25] See, e.g., Wellman-Labadie and Zhou (2010) for the incentives contained in the Orphan Drug Act.

$$B_3(t_1) \quad = -t_1 f \qquad\qquad (P - C \leq T_1 \leq P)$$

In figures 1 – 3, the game begins at $t_o < 1$ somewhere between p/cx and p/c. Assume that

$$L(t_o) = -t_o f + (1-\alpha)(t_o c - p)\pi^D > -pf,$$

so leading is more profitable than entering at P. Figure 1 suggests two possible equilibrium outcomes. In one, both generic firms enter at T_o; in the other, only one enters at T_o. The latter occurs if $B_2(t_o) = -t_o f + (1-\alpha)(t_o c - p)\pi^D/2 < -pf$, which rules out simultaneous patent challenges. In figure 2, only the leader challenges the patent as in the original game. The main difference is in figure 3; there is no longer a war of attrition in figure 3. Instead, the leader challenges the patent at t_o while the follower waits till the patent expires. Thus, with marketing exclusivity there is a patent challenge at t_o in all regions 1 – 3.

Now consider the counterfactual. As the game begins to the left of p/c in figure 5, the equilibrium has the leader challenging at t_o unless

(12) $$\hat{L}_1(t_o) \quad = -t_o f + (1-\alpha)(t_o c - p)\pi^T < -pf,$$

in which case there are no patent challenges. If (12) does not hold, the leader challenges the patent at T_o and the follower enters at P with certainty with or without marketing exclusivity. Thus, marketing exclusivity has no effect. If (12) holds, removing of marketing exclusivity deters patent challenging and benefits the branded drug company, increasing innovation incentives for orphan drugs.

7.2. Me-too drugs

Me-too drugs refer to reformulations of key products that use a different delivery system and therefore can be patented. Branded drug companies commonly introduce me-too drugs before the patent on the original drug expires. Introduction of me-too drugs intensifies competition within a given class of drugs, thereby reducing the profits to generics and possibly deterring generic entry.

To apply our model, suppose that introduction of a me-too drug does not deter generic entry but merely reduces generic profits by crowding the market. Assume that the incumbent has already developed a me-too drug and introduces it strategically; that is, it launches a me-too drug before the patent expires if there is a successful patent challenge; otherwise it introduces it after the patent covering the original drug expires. Finally, for the sake of economizing on space, we focus on the equilibrium region 1 only.

With marketing exclusivity, if the patent is found invalid, there is triopoly during the exclusivity period; the leader's profit equals $(c - cx)\pi^T$ and the incumbent's combined profits are $(c - cx)(\Pi^T + \Pi_m{}^T)$, where $\Pi_m{}^T$ denotes the (sum of) profits from a me-too drug.
In the post-exclusivity period, entry by the follower results in quadropoly. As a result, the leader and the follower receive the profits $(cx - p)\pi^Q$ apiece while the incumbent receives the combined profit of $(cx - p)(\Pi^Q + \Pi_m{}^Q)$, where $\Pi_m{}^Q$ denotes the profit attributable to a me-too drug under quadropoly.

By marketing a me-too drug, the incumbent increases its profit by $(c - cx)(\Pi^T + \Pi_m{}^T - \Pi^D) + (cx - p)(\Pi^Q + \Pi_m{}^Q - \Pi^T) > 0$. Implicit here is the assumption that introducing a me-too drug raises the incumbent's total profits when there are generic competitors. That is true in Cournot competition when marketing decisions for the original brand and a me-too drug are made independently of each other; see Baye, Crocker and Ju (1996). It also explains why a me-too drug is not introduced if the incumbent's patent is upheld; a monopoly always makes a greater profit than two duopolists.

Alternatively, without marketing exclusivity, introducing a me-too drug raises the incumbent's profit by $(c - p)(\Pi^Q + \Pi_m{}^Q - \Pi^T)$. The difference in gain is

$$(c - cx)[(\Pi^T + \Pi_m{}^T - \Pi^D) - (\Pi^Q + \Pi_m{}^Q - \Pi^T)].$$

In Cournot oligopoly with linear demand and constant and identical marginal costs, this difference is negative.[26] If so, removing of marketing exclusivity makes introduction of a me-too drug more profitable for the incumbent.

[26] With $n - 1$ (> 3) firms, $2/(n + 2)^2 - 1/n^2$ is increasing in n.

Although we focused on region 1, the above analysis is valid in other regions, as long as the equilibrium remains within the same region after the introduction of a me-too drug. However, introducing a me-too drug can cause the equilibrium to be in different regions for two reasons. First, diminished profits to the generic firms (the exclusivity profit π^D changes to π^T while the post-exclusivity profit π^T changes to π^Q) have the effect of rotating the graphs of $\phi(\alpha)$, $\psi(\alpha)$, and $\Gamma(\alpha)$ downward (counterclockwise) around the point $(\alpha, f) = (1, 0)$. As a result, the equilibrium tends to occur in a higher-numbered region for given (α, f). Second, it is possible that generic drugs infringe not only the patent on the original brand but the one covering a me-too drug, as well. If so, entry becomes riskier, implying that a given f is associated with a higher value of α. An increase in α may also cause the equilibrium to move out of the original region. However, as shown in section 5, a switch from region 1 to region 2 does not affect the incumbent's profits, with or without marketing exclusivity, and hence these indirect effects can be ignored. On the other hand, if a me-too drug causes drastic decreases in generic profits and/or in α, a new equilibrium can be in region 3 or beyond. In such cases the incumbent's profits increases, making an introduction of a me-too drug more profitable. An analysis of such cases is tedious but does not yield clear-cut results. Hence, we do not pursue it her. Instead, we conclude this subsection as follows: if an introduction of a me-too drug does not cause drastic indirect effects, removing of marketing exclusivity makes me-too drugs more profitable and diverts R&D resources away from development of truly innovative drugs.

8. Concluding remarks

In this paper we present a dynamic model of generic entry competition that captures many features of the Hatch-Waxman Act. Hatch-Waxman promotes generic entry in a two-pronged approach: streamlining of FDA approval procedures for generics, and granting of a 180-day marketing exclusivity right to the first challenger of the patent allegedly covering the original branded drug. Our analysis finds that these two measures have worked in tandem to restore incentives for new drug development while inducing generic entry.

We also find that marketing exclusivity harms consumers whenever it benefits the incumbent and that social welfare is unambiguously lower with marketing exclusivity when the entry cost is sufficiently low (region 1) and can be lower even if the entry costs are higher. Despite the dire consequences, however, our welfare results are derived conditionally on the innovation drug having already been discovered – the maintained assumption of our model. For innovation drugs yet to be discovered, the prospect of greater profits under marketing exclusivity can accelerate development of new drugs, benefiting both consumers and generic drug manufacturers. If these dynamic effects are taken into account, marketing exclusivity may have a welfare-enhancing effect in the long run. Thus, future research will more closely examine the long-run effect of Hatch Waxman on incentives, in particular, when there are competing innovation drug manufacturers.

Two additional extensions manifest themselves. First, because Hatch-Waxman promotes generic entry promotion in a two-pronged approach, it is unclear whether the actual generic penetrations we observed since 1984 were attributable to marketing exclusivity or to the entry cost reductions through easing of FDA approval procedures. Since the answer to this question holds important welfare implications in our analysis, separating the two effects will be a worthwhile extension in empirical research.

Second, legal and antitrust literatures have paid increasing attention to what are known as reveres payment settlements. Branded drug companies often arrange for such settlements with ANDA applicants in exchange for delayed entry. Reverse payment settlements are fairly common in practice but remains controversial.[27] With minor modifications to our model, we believe we can apply our analytical framework to examine the implications of reverse payment settlements on incentive restoration and generic entry competition.

[27] For example, see Bulow, (2004), Cotter (2004) and Jacobo-Rubio, Turner, and Williams (2016).

Appendix

Appendix A: In this appendix we first show that filing suit is the dominant strategy for the incumbent when the cost of doing so is negligibly small. We then examine the implications of non-negligible legal costs. Suppose what happens when the incumbent accommodates generic entry instead of filing suit. Then, the generic drug is marketed immediately. If the incumbent files suit at a later date $T_I > 0$, a court takes time to deliberate and announces its finding at $T_I + C$. Suppose that if the patent is found, the incumbent is fully compensated for the lost monopoly profits between $T = 0$ and $T_I + C$, during which its patent has been infringed. The expected profit to the incumbent from first accommodating the first challenger and then filing suit alter equals

(A1) $$(1 - ct_I)\Pi^D + \alpha(ct_I - p)\Pi^M + (1 - \alpha)(ct_I - p)\Pi^T + \alpha(1 - ct_I)(\Pi^M - \Pi^D).$$

The last turn represents the compensation of the lost profits while the incumbent's patent was being infringed. The expression in (A1) is maximized when $t_I = \exp(-rT_1)$ is 1, that is, if the incumbent fils suit as soon as possible. Evaluating (A1) at $t_I = 1$ yields

(A2) $$(1 - c)\Pi^D + \alpha(c - p)\Pi^M + (1 - \alpha)(c - p)\Pi^T + \alpha(1 - c)(\Pi^M - \Pi^D).$$

Recall that, if the incumbent sued the first generic firm without accommodation, its profit would equal

(A3) $$(1 - c)\Pi^M + \alpha(c - p)\Pi^M + (1 - \alpha)[\, (c - cx)\Pi^D + (cx - p)\Pi^T\,].$$

Subtracting (A2) from (A3) yields

$$(1 - \alpha)\, [\, (1 - c)(\Pi^M - \Pi^D) + (c - cx)(\Pi^D - \Pi^T)\,] > 0.$$

This inequality demonstrates that filing suit immediately is the dominant strategy for the incumbent.

We next consider the case there are non-negligible legal fees in infringement litigation. Since the literature is not clear on who ends up paying to cover such costs, we make some specific assumptions to avoid the number of innumerable taxonomical subcases to consider. Thus, we assume that the incumbent pay the legal fee $L > 0$, while generic firms bear no part of court fees regardless of a court's finding. Assume also that if the incumbent accommodates the first entrant it also accommodates the second when the latter enters. With these assumptions, the incumbent's profit from accommodation equals $(1 – p)\Pi^T$. Then in regions 1 and 2 with marketing exclusivity there is the unique α_1 that equates $(1 – p)\Pi^T$ to the profit from filing suit less L:

$$(1 – p)\Pi^T = (1 – c)\Pi^M + \alpha_1(c – p)\Pi^M + (1 – \alpha_1)[(c – cx)\Pi^D + (cx – p)\Pi^T] – L,$$

where the left-hand side equals the incumbent's profit (6) less L. Provided that $0 < \alpha_1 < 1$, the incumbent chooses to not to file suit but accommodate entry for all patents weaker (smaller α) than α_1. Similarly, in region 3 there is the unique α_3 that fulfills the condition

$$(1 – p)\Pi^T$$
$$= (1 – ct^*_1)\Pi^M + \alpha_3(ct^*_1 – p)\Pi^M + (1 – \alpha_3)[t^*_1 (c – cx)\Pi^D + (t^*_1cx – p)\Pi^T] – L.$$

where the right-hand side is the incumbent's profit (10) less L. The incumbent accommodates entry for all patents weaker than α_3. It is easy to check that $\alpha_3 < \alpha_1$.

Next, if there is no marketing exclusivity, there is $\hat{\alpha}$ in regions 1 through 3A such that

$$(1 – p)\Pi^T = (1 – c)\Pi^M + \hat{\alpha} (c – p)\Pi^M + (1 – \hat{\alpha})(c – p)\Pi^T – L,$$

where the right-hand side equals the profit in (7) less L. Again, entry is accommodated for any patent with $\alpha < \hat{\alpha}$. Since the incumbent earns more profit with marketing exclusivity, it follows that $\alpha_3 < \alpha_1 < \hat{\alpha}$. This implies that in regions 1 – 3A without marketing exclusivity entry is more likely to be accommodated, harming the incumbent but benefiting generic firms and consumers. In region 3B, there was no entry without marketing exclusivity. However, this result occurred because the generic firms believed that the incumbent would file suit if they entered. Now, since filing suit is not a credible threat in this region, generic firms enter, knowing they will be accommodated for $\alpha < \hat{\alpha}$. Then the result is exactly the same as in region 3A. By the same argument we can show that, even in the region beyond the boundary function $\Gamma(\alpha)$, where there is no generic entry originally, now generics will enter to get accommodated. To sum up the findings, removing marketing exclusivity induces entry that gets accommodated, harming the incumbent and benefiting generics and consumers. Since the incumbent is harmed by accommodated entry, it has less of an incentive to develop new drugs. Thus, higher legal fees militate against Hatch-Waxman achieving the incentive restoration objective.

Appendix B: In this appendix we prove that $\Gamma(\alpha) > \psi(\alpha) > \phi(\alpha)/2$. The first inequality obtains form a straightforward calculation:

$$\Gamma(\alpha) - \psi(\alpha) = \frac{(1-\alpha)(c-cx)(\pi^D - \pi^T)}{1-p} > 0.$$

For the second we have that

$$\psi(\alpha) - \phi(\alpha)/2 = \frac{(1-\alpha)(c-p)\pi^T}{1-p} - \frac{(1-\alpha)(c-cx)\pi^D/2}{1-\alpha p-(1-\alpha)cx}$$

$$> \left(\frac{c-p}{1-p} - \frac{c-cx}{1-\alpha p-(1-\alpha)cx}\right)(1-\alpha)\pi^D/2 .$$

The inequality holds because $2\pi^{T^o} > \pi^D$ as in assumption 1D. We can rewrite the last expression as

$$\frac{[1-c+\alpha(cx-p)](c-cx)}{(1-p)[1-\alpha p-(1-\alpha)cx]}(1-\alpha)\pi^D/2 > 0.$$

Thus, $\psi(\alpha) > \phi(\alpha)/2$.

Appendix C. In this appendix we discuss the case in which $\phi(\alpha) > \psi(\alpha) > \phi(\alpha)/2$. In this case, since $\psi(\alpha)$ is linear and $\phi(\alpha)$ is convex, it is possible to have the results depicted in figure A1, where we see the new region enclosed by the graphs of $\psi(\alpha)$ and $\phi(\alpha)$ emerging. Call it subregion 2B, as it is part of region 2, where with marketing exclusivity the leader enters at T = 0 and the follower enters at the end of the exclusivity period, conditionally. Without marketing exclusivity by contrast there are no patent challenges in subregion 2B. Thus the effect of marketing exclusivity is similar to that in subregion 3B. Marketing exclusivity promotes generic entry, benefiting the generic firms but harming the incumbent. It also improves social welfare.

References

Adams, C. P, and Brantner, V.V., 2006, Estimating the cost of new drug development: is it really $802 million? *Health Affairs* 25, 420-428.

Baye, M. R., Crocker, K. J., and Ju, J., 1996, Divisionalization, franchising, and divestiture incentives in oligopoly, *American Economic Review* 86, 223-236.

Bellefmalle, P., and Peitz, M., 2015, *Industrial organization: markets and strategies*. Cambridge University Press: Cambridge

Berndt, E. R., and Aitken, M. L., 2010, Band loyalty, generic entry and price competition in pharmaceuticals in the quarter century after the 1984 Waxman-Hatch Legislation, *NBER Working Paper* 16431.

Bulow, J., 2004, The gaming of pharmaceutical patents, in Jaffe, A. B., Lerner, J., and Stern, S., eds., *Innovation policy and the economy*. The MIT Press: Cambridge.

Choi, J. P., 1998, Patent litigation as an information-transmission mechanism, *American Economic Review* 88, 1249-1263.

Cotter, T. F., 2004, Antitrust implications of patent settlements involving reverse payments: defending a rebuttable presumption of illegality in light of some recent scholarship, *Antitrust Law Journal* 71, 1069-1097.

DiMasi, J.A., R. W. Hansen, H G. Grabowski, and L. Lasagna, 1991, Cost of innovation in the pharmaceutical industry, *Journal of Health Economics* 10, 107-142.

DiMasi, J.A., R. W. Hansen, H G. Grabowski, 2003, The price of innovation: new estimates of drug development costs, *Journal of Health Economics* 22, 151-185.

Frank, R.G., and D. S. Salkever, 1997, Generic entry and the pricing of pharmaceuticals, *Journal of Economics and Management Science* 6, 75-90.

Fubenberg, D., and Tirole, J., 1985, Preemption and rent equalization in the adoption of new technology, *Review of Economics Studies* 52, 383-401.

Fudenberg, D. and Tirole, J., 1990, *Game theory*. Cambridge, MA: MIT Press.

Grabowski, H., and Kyle, M., 2007, Generic competition and market exclusivity periods in pharmaceuticals. *Managerial and Decision Economics*, 27, 1 – 12.

Hemphill, C. S., and Lemley, M. A., 2011, Earning exclusivity: generic drug incentives and the Hatch-Waxman Act, *Antitrust Law Journal* 77, 947-989.

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

Hendricks, K., Weiss, A., and Wilson, C., 1988, The war of attrition in continuous time with complete information, *International Economic Review* 29, 663-680.

Iizuka, T., 2012, Physician agency and adoption of generic pharmaceuticals, *American Economic Review* 102, 2826-2853.

Jacobo-Rubio, R., Turner, J. L., and Williams, J. W., 2016, Generic entry, pay-for-delay settlements, and the distribution of surplus in the US pharmaceutical industry, unpublished manuscript.

Katz, M. and Shapiro, C., 1987, R&D rivalry with licensing or imitation, *American Economic Review* 77, 402-420.

Lemley, M. A., and C. Shapiro, (2005), Probabilistic patents, *The Journal of Economic Perspectives* 19, 75-98.

Morton, F. S., and Kyle, M., 2012, Markets for pharmaceutical products, chapter 12, in Pauly, M. V., McGuire, T. G., and Barros, P. P., eds, *Health economics*, North-Holland: Amsterdam.

Mossinghoff, G. J., 1999, Overview of the Hatch-Waxman Act and its impact on drug development process, *Food Drug Law Journal* 54, 187-194.

Reiffen, D., and Ward, M. R., 2005, Generic drug industry dynamics, *Review of Economics and Statistics* 87, 37-49.

Rosenthal, J., 2002, Hatch-Waxman use or abuse – collusive settlements between brand-name and generic drug manufacturers, *Berkeley Technology Law Journal* 17, 317-335.

Salant, S., Switzer, S., and Reynolds, R., 1983, Losses from horizontal merger: the effects of an exogenous change in industry structure on Cournot-Nash equilibrium. *Quarterly Journal of Economics* 95, 185-199.

Simon, L., and Stinchcombe, M., 1989, Extensive form games in continuous time: pure strategies, *Econometrica* 57, 1171-1214.

Wellman-Labadie, O., and Zhou, Y., 2010, The US Orphan Drug Act: Rare disease research stimulator or commercial opportunity? *Health Policy* 95, 216-228.

Figure 1

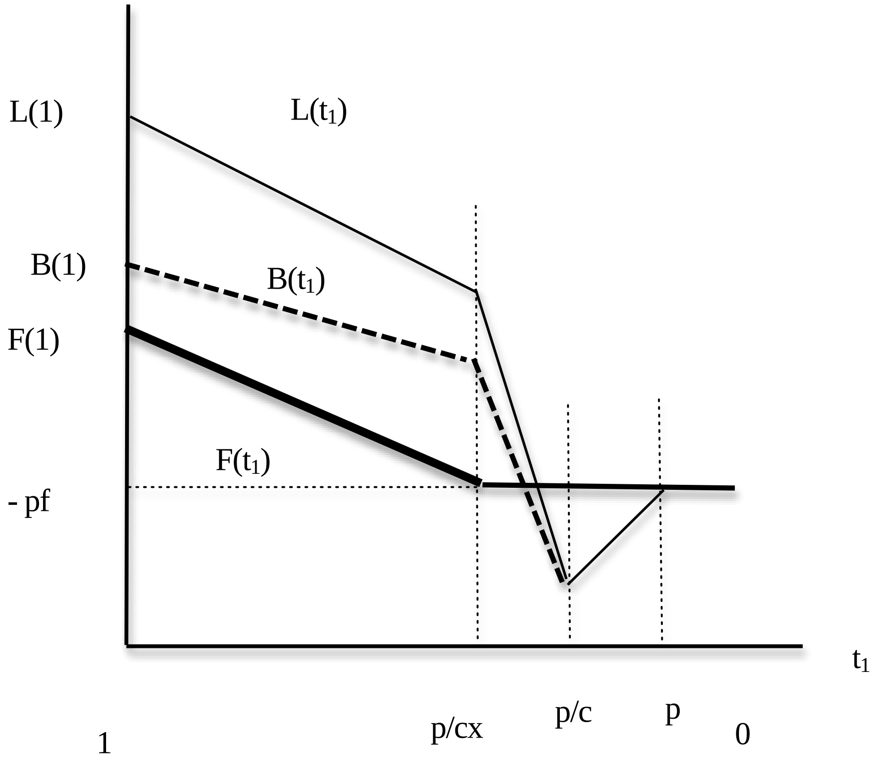

Figure 2

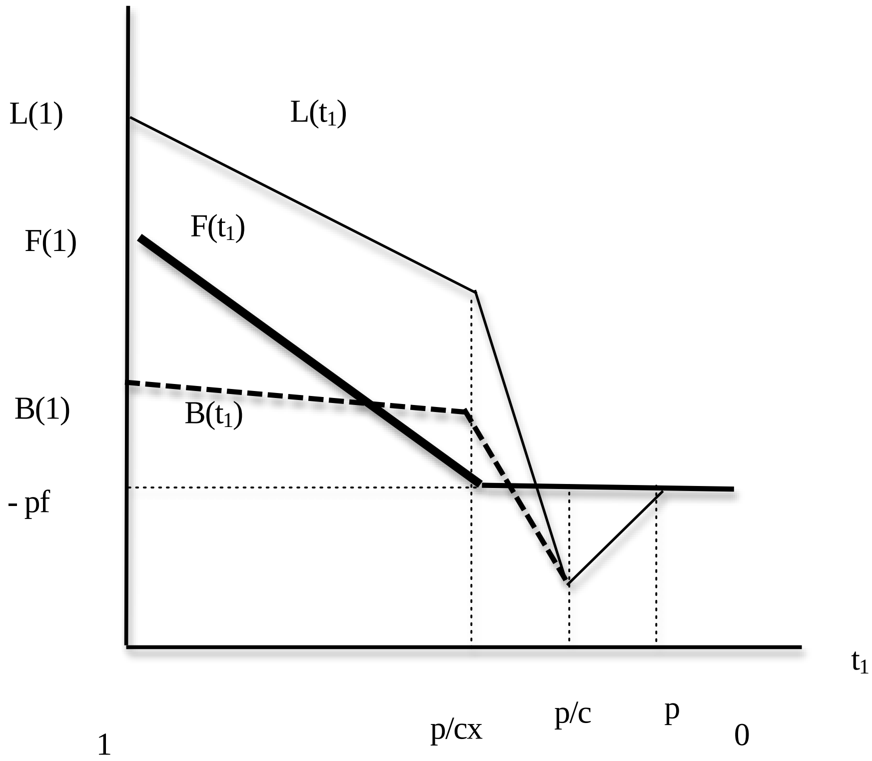

Figure 3

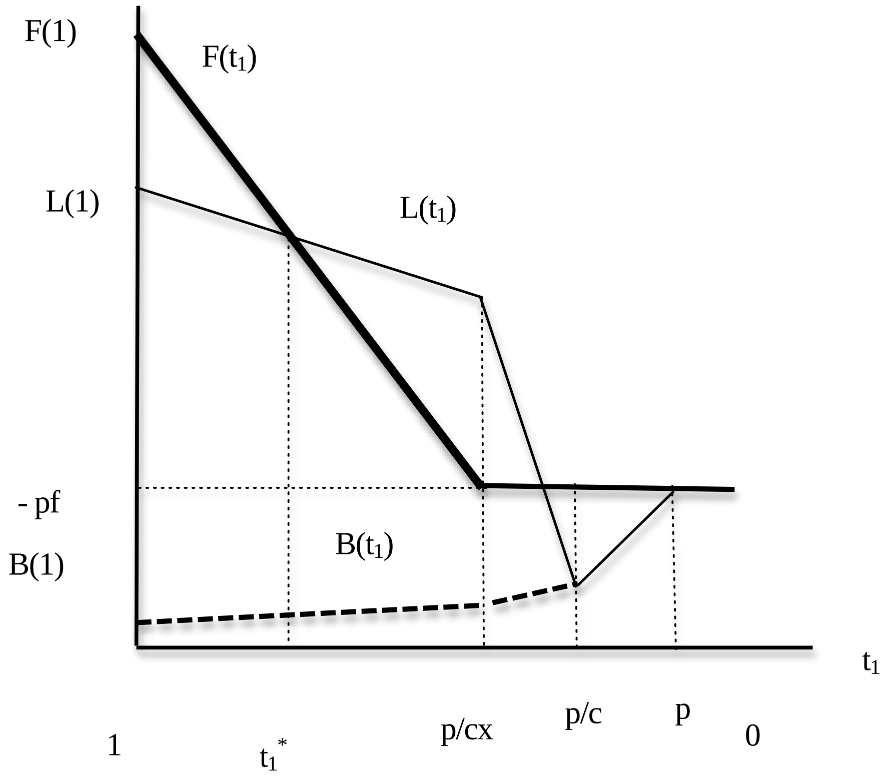

Figure 4

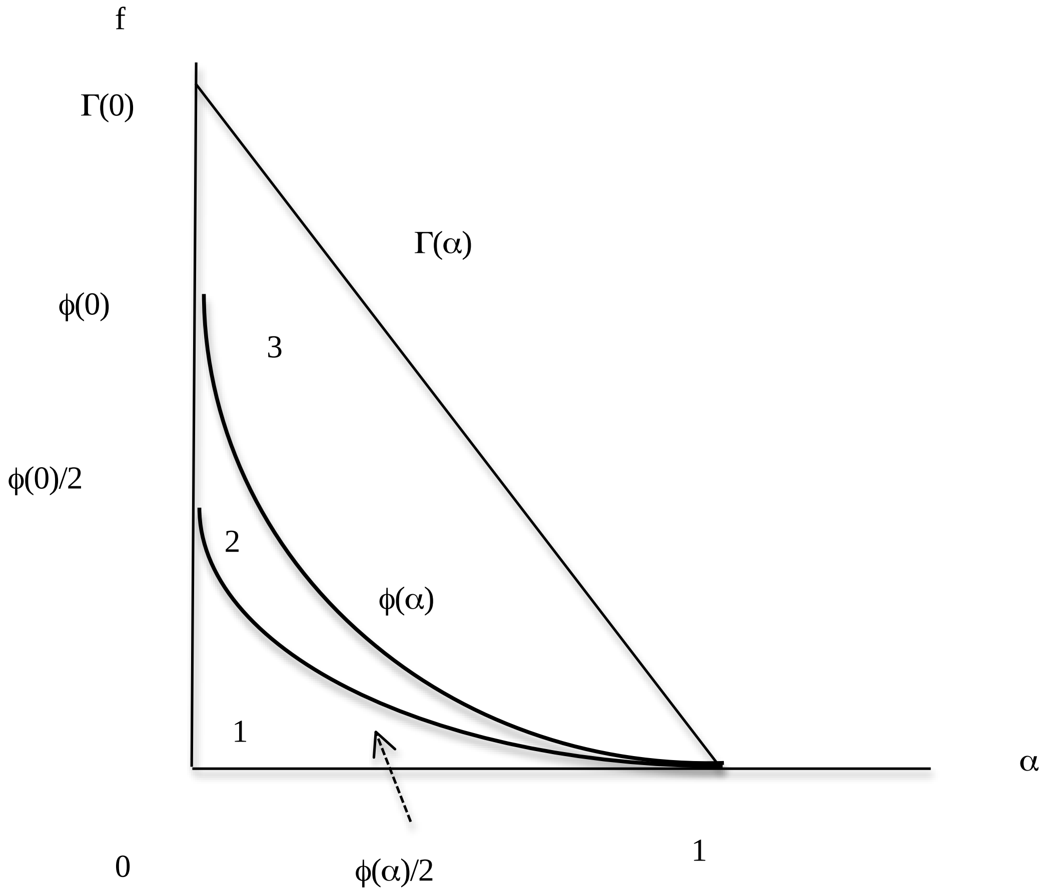

Figure 5

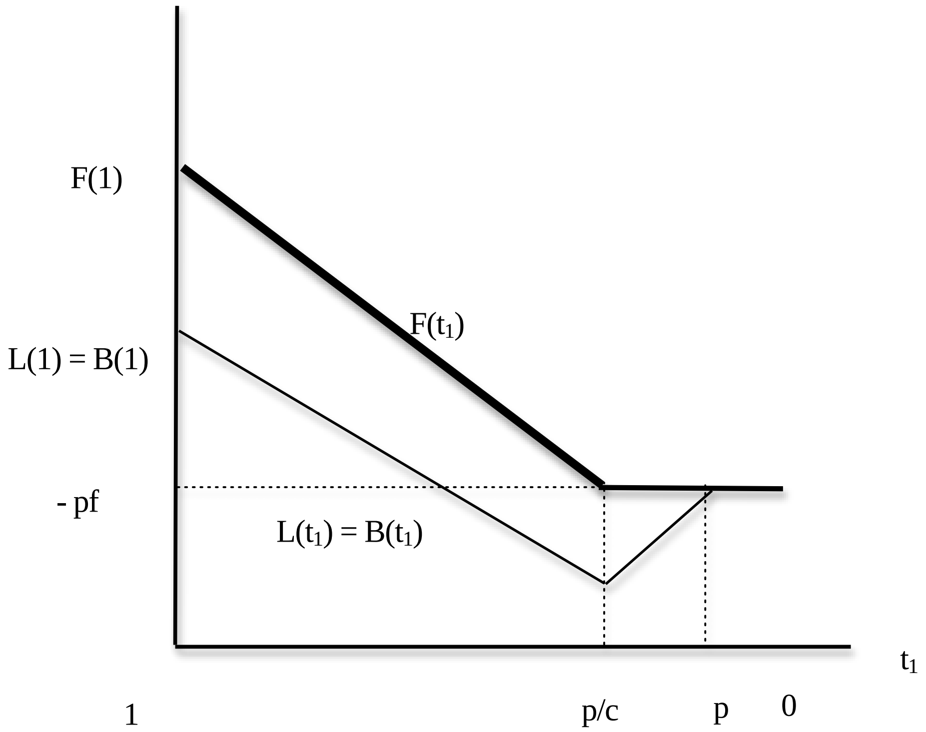

Figure 6

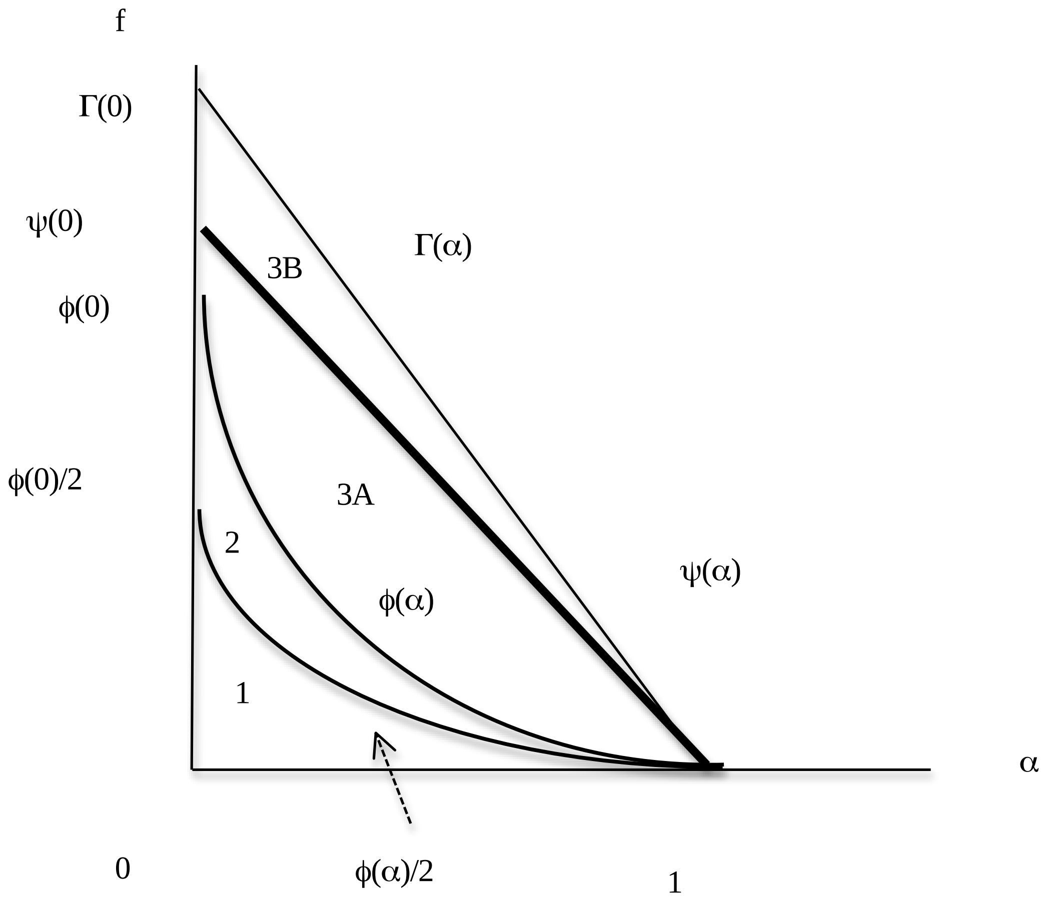

Figure A1