DISCUSSION PAPER SERIES

DP14276

LICENSING AT THE PATENT CLIFF AND MARKET ENTRY

João Montez and Annabelle Marxen

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

LICENSING AT THE PATENT CLIFF AND MARKET ENTRY

João Montez and Annabelle Marxen

Discussion Paper DP14276
Published 06 January 2020
Submitted 02 January 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

LICENSING AT THE PATENT CLIFF AND MARKET ENTRY

Abstract

We study the incentives for a monopoly incumbent to reach an agreement allowing a generic to enter just before its patent expires, i.e., at the patent cliff, and its consumer and social welfare effects. In our model, entry by more than one entrant is unprofitable. Thus, in the absence of an agreement, the entry game has a “grab the dollar” structure, with each generic entering in each period with a low (high) probability if entry costs are high (low). In that case the incumbent can remain a monopolist for some time after patent expiry, until one or more generics finally enter. An early entry agreement guarantees a single generic enters the market immediately, and it allows the incumbent to extract the entrant's profit. It will be reached in equilibrium when entry costs are low or the entry process is short. In these instances, early entry agreements do however tend to hurt consumers. Yet, allowing for such agreements increases overall social welfare in a benchmark model of vertical differentiation, even if the expected competition on the market is reduced. The same holds in a benchmark model with captive consumers and shoppers, provided the share of captives is not too high.

JEL Classification: L1, L4, I1

Keywords: N/A

João Montez - joao.montez@unil.ch
Université de Lausanne and CEPR

Annabelle Marxen - Annabelle.Marxen@unil.ch
Swiss Competition Commission

Acknowledgements
The authors gratefully acknowledge financial support from the Swiss National Science Foundation through grant no. 00018 176448. We thank, for helpful comments and suggestions, Alexandre de Cornière, Bruno Julien, Jurgen Maurer, Juan-Pablo Montero, Damien Neven, Volker Nocke, Santiago Oliveros, Martin Peitz, Markus Reisinger, Patrick Rey, Katharine Rockett, Yossi Spiegel, Luis Vasconcelos, participants at seminars in Toulouse, Essex, Basel, Goethe University and Frankfurt School, and at Workshop on Advances in Industrial Organization Bergamo 2019, CEPR Conference on Applied IO 2018, CRESSE 2018, Swiss IO Day 2017, Swiss Society of Economics and Statistics 2017, and EARIE 2017. An earlier version circulated with the title "Generics and early entry agreements".

Licensing at the patent cliff and market entry *

Annabelle Marxen † João Montez ‡

December 2019

- WORKING PAPER -

Abstract

We study the incentives for a monopoly incumbent to reach an agreement allowing a generic to enter just before its patent expires, i.e., at the patent cliff, and its consumer and social welfare effects. In our model, entry by more than one entrant is unprofitable. Thus, in the absence of an agreement, the entry game has a "grab the dollar" structure, with each generic entering in each period with a low (high) probability if entry costs are high (low). In that case the incumbent can remain a monopolist for some time after patent expiry, until one or more generics finally enter. An early entry agreement guarantees a single generic enters the market immediately, and it allows the incumbent to extract the entrant's profit. It will be reached in equilibrium when entry costs are low or the entry process is short. In these instances, early entry agreements do however tend to hurt consumers. Yet, allowing for such agreements increases overall social welfare in a benchmark model of vertical differentiation, even if the expected competition on the market is reduced. The same holds in a benchmark model with captive consumers and shoppers, provided the share of captives is not too high.

Keywords: early entry agreements, market foreclosure

JEL classification: L13, L41, I11.

*An earlier version circulated with the title "Generics and early entry agreements". The authors gratefully acknowledge financial support from the Swiss National Science Foundation through grant no. 00018 176448. We thank, for helpful comments and suggestions, Alexandre de Cornière, Bruno Julien, Jurgen Maurer, Juan-Pablo Montero, Damien Neven, Volker Nocke, Santiago Oliveros, Martin Peitz, Markus Reisinger, Patrick Rey, Katharine Rockett, Yossi Spiegel, Luis Vasconcelos, participants at seminars in Toulouse, Essex, Basel, Goethe University and Frankfurt School, and at Workshop on Advances in Industrial Organization Bergamo 2019, CEPR Conference on Applied IO 2018, CRESSE 2018, Swiss IO Day 2017, Swiss Society of Economics and Statistics 2017, and EARIE 2017.

†Swiss Competition Commission. Disclaimer: This work has been carried out while being a PhD candidate at the University of Lausanne. The views and opinions expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Swiss Competition Commission.

‡Department of Economics, University of Lausanne. Corresponding author: joao.montez@unil.ch

1 Introduction

Agreements between potential rivals are common in the pharmaceutical sector. These have received great attention from the media and antitrust authorities, in a context of rising health expenditures and perceived anticompetitive behavior. The poster child are *pay for delay deals*, often perceived to be used by an incumbent to pay challenging entrants in order to keep them out of the market.

In this paper we focus on the more prevalent *early entry agreements*, whereby the incumbent instead receives a fee to grant entry to a generic just before an unchallenged patent expires, i.e., at the patent cliff. While our work is motivated by current practices and concerns in the pharmaceutical industry, our analysis and results are not industry specific: They apply more generally to all intellectual property and its licensing, and should inform related debates.

An example that received media attention occurred in August 2005, when Bristol-Myers Squibb's patent for cholesterol drug *Selipran* expired on the Swiss market. At that time, *Selipran* was a blockbuster drug, the second most sold drug with a turnover of 68 million Swiss francs. Just three months before patent expiry, in June 2005, the generic version *Pravalotin* started being commercialized by the generic producer Mepha. As reported by the Handelszeitung, Mepha is said to have paid Bristol-Myers Squibb a seven figure sum in an early entry agreement.[1]

Such agreements are typically shrouded in secrecy. The European Commission's pharmaceutical industry sector inquiry of 2009 does however provide us with a rare glimpse into such practices, finding that for instance there were twice as many early entry agreements as pay for delay deals in the period of observation.[2] In contrast to pay for delay deals, academic research on the former is almost nonexistent.

A naive perspective would qualify early entry agreements as welfare enhancing. As pointed out by a legal expert, *"the comparison with the counterfactual situation would generally lead to the conclusion that the agreement is pro-competitive: consumers have an earlier access to a cheaper alternative"* (Struys (2012)).

Such benefit is however likely to be insignificant, as the bulk of entry due to licensing takes place right at the patent cliff. Indeed the very late "early entry" described above is not atypical but the norm: in the EU, under these agreements, entry occurred *less than a year* before patent expiry in more than 80% of the markets that had remained insulated from generic competition (see § 817 and fig. 126 of EC (2009)).[3]

Instead, the European Commission expressed concerns that *"the early presence of a generic product limits the attractiveness of a market for other companies"* (see § 809 of EC (2009)). Our analysis substantiates and provides formal support to this concern. We find that early entry agreements lead to less entry in the long-run, and often hurt consumers. In spite of this, and perhaps surprisingly, we also find that they often increase social welfare.

We study an infinite period model with one incumbent, initially protected by a patent,

[1] Generikamarkt: Millionen für einen Vorsprung, http://www.handelszeitung.ch/unternehmen/generikamarkt-millionen-fuer-einen-vorsprung, accessed: October 7, 2019.

[2] Of the reported 285 non-litigation agreements, roughly 30% concerned early entry, while out of 207 settlement agreements, only 20% concerned pay for delay, EC (2009).

[3] In the few cases where the agreement resulted in entry more than a year before loss of exclusivity, § 820 concludes that typically this *"was not the first generic product on the market, which suggests that the originator companies were reacting to the presence on the market of one or more other generic companies"*. Moreover, there is evidence that in such cases non-linear contracts may dampen competition until patent expiry (see § 821 of EC (2009)), which further mitigates the potential entry benefits during the patent period.

and two potential entrants producing a generic version of the drug. In each period after patent expiry, generic firms decide whether to enter or not (if they have not done so before), and then all the firms that have entered choose prices. Each additional entry decreases others', and aggregate industry, revenues. Importantly, due to entry costs, entry by a single entrant is profitable but entry by both is not. Thus our analysis applies best to intermediate sized markets, that are sufficiently large to not remain a monopoly, but sufficiently small to not accommodate substantial entry.

In the absence of early entry, generics' incentives have a *grab the dollar game* structure. As long as no entry has taken place, each generic enters in each period with some probability, which is decreasing in the entry cost level.[4] As a result, the incumbent remains a monopolist for some time after patent expiry, before eventually one or both generics enter the market. This additional monopoly period increases with the entry cost and entry process length.

Before patent expiry, i.e., at the patent cliff, firms may also reach an early entry agreement: At the outset of the game, entrants make simultaneous offers to the incumbent for the exclusive right to enter just before patent expiry. If an offer is accepted, no further entry takes place after patent expiry (since entry by both is unprofitable). The incumbent therefore chooses to accept or reject an offer by comparing the acceptance profit with its profit in the alternative of uncoordinated entry.

The former exceeds the latter when either the entry costs are low or the entry process is short, as then, in the absence of an agreement, the monopoly period after patent expiry is likely to be short-lived and more competition is expected in the long-run. Therefore, consumers are also typically worse off as a result of such an agreement.

As consumer and industry incentives tend to be misaligned, to study welfare implications of such agreements we need additional structure on the economic fundamentals. We turn to features of over-the-counter (OTC) drug markets to propose two benchmark models. Even if from a bioequivalence point of view generics and originator drugs are identical, it is a well established fact that long after patent expiry the original drugs keep significant market shares despite charging large price premiums.[5]

Medical studies have documented that patients report on average significantly better effects when they know they are consuming an originator product. This brand effect is reinforced when consumers are regular users of the test brand. Such effects also hold for non-medical products, such as energy drinks (see e.g., Shiv et al. (2005)), and can be reinforced when consumers are informed about price differentials.[6] This suggests that the originator product may enjoy a higher (perceived) quality in consumption, which can be captured in the classic vertical differentiation framework of Mussa and Rosen (1978), which is also consistent with the mentioned price premium of originator drugs.

A second benchmark model builds on evidence that less informed consumers are willing to pay a premium for branded drugs (see Bronnenberg et al. (2015)). The existence of educational campaigns by national health regulators to encourage patients to opt for generics

[4] In a related game with multiple entrants but no incumbent, Elberfeld and Wolfstetter (1999) discuss similar behavior and dynamics of entry.

[5] For European evidence see e.g., the Sector Inquiry of the EC (2009) and for US evidence see e.g., U.S. Food & Drug Administration (2013) Statement on Pay-for-Delay.

[6] For instance, in a controlled clinical trial, where subjects were informed of the brand versus no-brand status of the given preparation (but not whether it contained a placebo), Branthwaite and Cooper (1981) find that mean pain relief of acetylsalicylic acid, commonly know as Aspirin, was significantly higher with a branded preparation. Waber et al. (2008) show that consumers report greater pain relief for an opioid analgesic when they are informed that the price is the regular price rather than a discounted one (the actual dispensed product to all participants was an inactive preparation).

further suggests that consumer information on the bioequivalence and availability of generics plays an important role on OTC markets. We capture these features in a model where a group of uninformed consumers only considers consuming the original product, and another informed group – knowing that both products are equivalent – shops for the best deal. This give us an asymmetric clearinghouse model with captive consumers and elastic demand (inspired by Varian (1980)), which also consistent with mentioned price premiums.

These two models allow us to make more precise statements on the welfare effects of early entry agreements in OTC markets. We find that in both cases they tend to be welfare increasing, even when entry is virtually free. This is perhaps unexpected since, when entry is virtually costless, entry by both generics is almost certain in the absence of an agreement and competition is then maximized. In other words, monopoly is inefficient, but too much competition can also have a negative welfare effect. These results contribute to the existing literature on welfare effects of entry, which in contrast to our work sees in the multiplication of entry costs the source of excessive entry. We qualify and explain this result next.

In the vertical differentiation model, the high quality incumbent responds to the entry of a single lower quality generic by reducing its price to an extent that increases the number of consumers it serves (thus it is said to fight single entry), and welfare increases relative to monopoly. With additional entry the generic price drops to marginal cost. The incumbent then becomes the monopoly seller of only a quality upgrade on a basic good that is provided at cost, and reverts to selling the same initial monopoly quantity (albeit at a lower price). It thus serves only those consumers that value quality the most, and on which it can still get a significant margin (thus it is said to accommodate additional entry). Additional generic entry reduces welfare (regardless of the entry cost) since the welfare loss associated with intermediate valuation consumers' switch from high-quality to low-quality consumption always outweighs the welfare gain that is obtained from the lowest valuation consumers' switch from no consumption to low-quality consumption. To our knowledge this counterintuitive result has not been documented in the literature.

In the captive consumer model, if there is double generic entry, all informed consumers purchase the generic at cost and the incumbent sells to its captive uninformed consumers at the monopoly price. Thus, relative to monopoly, double entry benefits informed consumers but not the uninformed ones. On the other hand, with a single generic entrant, the incumbent sometimes targets only the uninformed consumers with the monopoly price, and other times holds a sale trying to capture the informed consumer segment from the entrant (to sell a larger volume but at a lower price). The frequency and extent of these discounts is increasing in the volume at stake (the number of informed consumers). Thus, while both groups benefit from single entry relative to monopoly (as the incumbent fights single entry), a second entrant benefits informed consumers but hurts uninformed ones (as the incumbent accommodates additional entry).

When most of the consumers are informed, a single entrant is all that it takes to have steep discounts frequently, and in equilibrium both prices to be close to marginal cost most of the time (as in the limiting standard Bertrand duopoly case where all consumers are informed). There is then little benefit to consumers (in aggregate) from a second entrant, but a significant reduction in the industry profit. Thus, if the share of uninformed consumers is not too large, allowing for early entry agreements also increases welfare.

These results highlight that antitrust authorities face a trade-off. An authority that is mostly concerned about social welfare may want to take a lenient approach with respect to early entry agreements, while an authority that mostly cares about consumers will want to take a strict stance, in particular if entry costs are low or the entry process is short (instances where early entry agreements lead, on average, to less entry and higher overall prices).

In our model, early entry agreements can be seen as a form of entry deterrence through licensing. In this sense, the work closest to ours is Rockett (1990). Like in our model, entry by a single entrant is profitable but entry by both is not. The incumbent then uses a license as an instrument to select which of two entrants with heterogeneous marginal costs will become its competitor. The incumbent faces a trade-off of licensing to a low cost competitor, thereby being confronted with softer competition but receiving a lower licensing fee, or choosing to grant a license to a stronger competitor, with potentially a higher licensing fee but tougher market competition post entry.

An important modeling difference is that Rockett (1990) focuses on a sequential (thus coordinated) entry game with asymmetric entrants, while we study an entry game with symmetric entrants and uncoordinated entry. In her work, if entrants are symmetric, then licenses are used as a rent extraction tool only, corresponding to the asymmetric equilibria of our game. With uncoordinated entry, our focus shifts from the identity of the selected entrant (*which*), to the number of entrants gaining access to the market (*how many*). Moreover, with sequential entry licensing must always occur in equilibrium. This is however not the case with simultaneous uncoordinated entry, which allows us to provide a theory of when early entry agreements may or not to be reached.

Moreover, as Rockett (1990) studies a reduced form model, only the strategic incentives of the firms are considered. We characterize not only the industry incentives, but also consumer surplus effects, and provide a more complete understanding of welfare effects in two benchmark models.

Exploring a related tradeoff, in Duchêne et al. (2015) an incumbent can offer a low licensing fee and royalty to a rival incumbent, giving it access to a cost reducing technology, such that a now more efficient rival dissuades further entry. This entry deterrence strategy can increase consumer surplus if the welfare gains of lower prices through lower royalty costs (price distortion effect) outweigh the losses of softer competition (duopoly instead of triopoly), which can occur if fixed costs are low.[7] Instead of complete entry deterrence, Gallini (1984) discusses how an incumbent may use licensing in order to deter an entrant from R&D into a potentially better technology.

Most of the existing literature on generic entry agreements focus on *pay for delay* agreements. A key element of that literature is that, as a result of litigation, a court may invalidate the patent. A longheld view is that settlements are welfare enhancing since they allow firms to avoid wasteful litigation costs. However, they can also be a tool for the incumbent to delay generic entry beyond the expected date in case of litigation, as a payment larger than the saved costs of litigation can be used to compensate the generic producer for its lost profit. This concern was formalized by Shapiro (2003), whose results have been extensively used in the pay for delay case law. In the presence of asymmetric information, where the incumbent has more information about the strength of the patent, timing and cash transfers allow for an agreement as they reveal private information (Willig and Bigelow (2004) and Dickey et al. (2010)). Gratz (2012) shows that the possibility of erroneous results by the antitrust enforcer and costly transaction costs makes a rule of reason less appealing,

[7] Yi (1998) extends Rockett's work by allowing for a two-parts tariff, composed of a fixed licensing and a per-unit royalty. Eswaran (1994) generalizes Rockett's work by considering multiple weak potential entrants and endogenizing the royalty rate.

and favors the legality for patent settlements with a threshold limit for cash payments.[8]

Two key differences between early entry agreements and pay for delay deals are that the former typically occur in the absence of any patent litigation and in practice are reached only towards the very end of a patent which remained unchallenged. This observation suggests that pay for delay deals seem to be used to manage competition during the life of a weak patent, and early entry agreements instead used to manage competition after the life of a strong patent.

The interaction between litigation costs and licensing also raises interesting strategic considerations in the presence of multiple entrants. For instance, Choi (1998) includes the possibility of patent litigation in a setting where multiple entrants play a sequential entry game. Contrary to our work, the market can accommodate all entrants. He analyses the effects of information revelation in patent litigation, where the court decision is seen to eliminate the uncertainty on a patent's strength. If the incumbent chooses to litigate the first entrant, this enables informational free riding by the second entrant, resulting in a waiting game between the entrants. If patent strength is intermediate, the incumbent may however choose not to litigate the first entrant as the negative revelation effect outweighs the deterrence effect on the second entrant. In that case, a preemptive game takes place, where each entrant wants to be the first one to imitate, knowing he will be accommodated but further entry deterred.

Combining pay for delay and strategic deterrence, Palikot and Pietola (2019) focus on the externalities created by settlements that postpone entry until patent expiry, and analyze why settlements and licenses may coexist. In their model, each additional settlement places a positive externality on the remaining entrants (as this increases the expected profit of challenging the patent), which increases the cost of further settlements. In equilibrium, the incumbent may then choose to license some entrants and to litigate others. This divide-and-conquer strategy reduces the cost of each settlement that is reached, at the cost of accepting some entry during the patent period. Bokhari et al. (2017) analyze the stability of equilibria in pay for delay. By allowing the originator to introduce its own generic version (often referred to as an authorized generic), either in-house or via the first challenger, they conclude that if the first mover advantages are large enough then the threat of launching an authorized generic is a credible deterrent to additional entry.

Classical vertical differentiation models à la Shaked and Sutton (1982) (see e.g., Wauthy (1996) for a presentation) have been used to study entry deterrence and its impact on welfare. Most papers study duopoly settings with endogenous quality choice, and use the differentiation level choice as a mechanism for entry deterrence.[9] In a three firm model like ours, Donnenfeld and Weber (1992), Donnenfeld and Weber (1995), and Peitz (2002) show that the two first movers may both choose relatively high quality, and thus low quality differentiation, to intensify market competition and deter entry by the third firm. Total industry profits may then fall, but the quality adjusted price may decrease sufficiently as

[8] The 2013 landmark *Actavis* case established in the US the principle under which pay for delay agreements are likely to have anticompetitive effects if the difference between the reverse payment and litigation costs is positive. Europe has opted for a similar quantitative approach. While some authors argue in favor of a per se illegality for patent settlement that contain considerable cash transfers (Bulow (2004), Hovenkamp et al. (2003), Bokhari (2013) or Davis (2009)), most legal scholars however advocate the use of a rule of reason (e.g., Crane (2002), Cotter (2004) and Schildkraut (2004)).

[9] In addition to using vertical differentiation as an instrument for entry deterrence, Davis et al. (2004) model the monopolist's pricing decision as a further deterrence tool. Entry deterrence will then become welfare increasing if the effects from product design improvement outweigh the effects from higher pricing. For a study on the effect of quality-dependent marginal costs, see Lutz (1997) and Noh and Moschini (2006).

to increase consumer surplus and total welfare. A sufficiently high level of entry costs is however key for entry deterrence, and such a positive welfare effect.

In contrast we find that, in the same classic vertical differentiation framework, if instead entrants have exogenous and sufficiently homogeneous qualities, then limiting entry will increase welfare regardless of the entry cost. In that case agreements with deterrence effects, like early entry agreements, can also increase welfare.

This result is particularly interesting also in light of the extensive literature that studies the occurrence of excessive entry in the presence of entry costs. For example, in Mankiw and Whinston (1986), each entrant fails to internalize a business stealing effect, which leads to excessive entry from a welfare perspective due to the multiplication of entry costs.[10]

We also study the welfare of entry in a clearinghouse model, as pioneered by Varian (1980) (see e.g., Baye et al. (2006) and Shelegia and Wilson (2016) for thoughtful reviews of the literature). Unlike most of that literature, in our paper demand is elastic and firms are asymmetric, as only the incumbent has captive consumers.[11] Methods to analyze such models have only been developed recently (see e.g., Shelegia and Wilson (2016) and Montez and Schutz (2018)). The present paper is a first in making use of these progresses to study the welfare effects of entry and entry deterrence in an asymmetric clearinghouse framework.

The paper is organized as follows. In Section 2 we introduce the general model, and characterize when early entry agreements are reached in equilibrium and identify those instances where agreements harm consumers—as a function of the entry costs and length of the entry process. In Section 3 we study the welfare implications of early entry agreements in two benchmark models of price competition. Section 4 concludes. All proofs that are not treated in the main text can be found in the Appendix.

2 Early entry agreements and entry deterrence

2.1 The model

There are three firms. Firm I is an incumbent, and firms 1 and 2 are potential entrants. We consider a game where the incumbent and either entrant can reach an *early entry agreement* prior to a dynamic Bertrand game with costly entry, which takes place after patent expiry. We study the nontrivial case where the entry costs are such that entry by one entrant is profitable but entry by both is not. These entry costs $F > 0$ capture not only actual production costs but also, for example, the cost of establishing generic's bioequivalence, authorization, and marketing costs. Our analysis thus applies best to intermediate sized markets: sufficiently large to not become a natural monopoly, but sufficiently small to not accommodate substantial entry.

The early entry agreement period, taking place just before patent expiry, has two stages. In the first stage, each entrant i simultaneously makes a monetary offer b_i to I for the exclusive right to enter immediately before the patent expires. In the second stage, the

[10]In their work the tension between the business stealing effect (output contracts for all if there is entry deterrence) and the product diversity effect (each new entrant increases product variety) determines whether free entry leads to insufficient or excessive entry. In the former scenario, entry deterrence may then be welfare increasing.

[11]Welfare is typically not discussed in this literature since it works with inelastic demand. In a model that departs from the clearinghouse setup, but is related to it, Anderson et al. (2015) study the effect on welfare when firms are heterogenous and all consumers shoppers. Instead of holding sales, firms advertise to gain positive profits.

incumbent either accepts one offer or it rejects both—and it always accepts one offer when indifferent. The focus on exclusive entry just prior to patent expiry is without loss of generality: neither earlier entry nor non-exclusive offers would be used in the equilibrium of a richer bidding game. The acceptance of offer b_i from entrant i is denoted by $a_i^I = 1$ and its refusal by $a_i^I = 0$. If entrant i's offer has been accepted then i enters at the end of the early entry period, and it pays b_i to the incumbent plus the entry costs F.

Once the patent expires, at $t = 0$, a dynamic infinite-horizon Bertrand game with potential entry takes place. Firms then make decisions at each discrete period $t \geq 0$. Each period has two stages. In the first stage firms 1 and 2 decide simultaneously whether to pay F and enter the market at t (if they have not done so before), or to wait and potentially enter at a later date. An entry at t by entrant $i \in \{1, 2\}$ is denoted by $a_i^t = 1$, with $a_i^t = 0$ denoting that i does not enter at t. Entrant i is said to be active at t if $a_i^{t'} = 1$ for some $t' \leq t$ or $a_i^I = 1$, and inactive at t if $a_i^{t'} = 0$ for all $t' \leq t$ and $a_i^I = 0$. For example, if the early entry offer of firm i has been accepted, then i is active at every $t \geq 0$, in which case $a_i^t = 0$ for all t. In the second stage of each period $t \geq 0$, active firms compete in the market by simultaneously choosing prices. With abuse of notation we denote the price of a non-active firm i by $p_i = \{\emptyset\}$. The actions in each period $t \geq 0$ therefore are: a price $p_I^t \in \mathbb{R}_0^+$ for the incumbent I, a pair (a_i^t, p_i^t) with $a_i^t \in \{0, 1\}$ and $p_i^t \in \mathbb{R}_0^+ \cup \{\emptyset\}$ for each entrant $i \in \{1, 2\}$.

Firms make decisions at discrete time periods but the market operates in continuous time, and thus payoffs accrue to firms in real-time. Next we make assumptions on the instantaneous date τ market revenue of each firm $i \in \{I, 1, 2\}$, denoted by $\pi_i(P(\tau))$, and $P(\tau)$ is the vector of prices at τ. Consider first the case where both entrants are inactive. Then $p_1 = p_2 = \emptyset$. The incumbent's *monopoly* price p^m is unique, maximizing $\pi_I(P)$. The resulting price vector $(p^m, \emptyset, \emptyset)$ is P^m. Consider next that only firm j is inactive. The simultaneous one-shot *duopoly* pricing game between I and i, with respectively payoffs $\pi_I(P)$ and $\pi_i(P)$, has a unique Nash equilibrium (possibly in mixed strategies) which, with abuse of notation, is denoted by P_i^d. Finally, suppose both entrants are active. The simultaneous one-shot price *competition* game between the three firms has a unique equilibrium (possibly in mixed strategies) which, also with abuse of notation, is denoted by P^c.

The incumbent's revenue in the one-shot pricing game is (weakly) decreasing in the number of active rivals. Moreover, reflecting that generic versions are perceived to be imperfect substitutes to the original, generic competition falls short of driving the incumbent's revenue to zero. Thus

$$\pi_I(P^m) > \pi_I(P_i^d) \geq \pi_I(P^c) > 0.$$

Inactive firms' revenues are zero. The duopoly revenue of the active generic firm i is positive but competition by two generics drive entrants' revenues to zero.[12] Thus

$$\pi_i(P_i^d) > \pi_i(P^c) = \pi_j(P_i^d) = 0 \text{ for } i \in \{1, 2\} \text{ and } j \neq i.$$

Each additional entry strictly reduces the *aggregate* industry revenues, i.e.,

$$\pi_I(P^m) > \pi_I(P_i^d) + \pi_i(P_i^d) > \pi_I(P^c).$$

The last assumption also implies that, even if the date of entry were to be endogenously determined, in equilibrium early entry would not occur until the very last moment before

[12]It is not essential that revenues become zero, just that entry by both entrants becomes unprofitable.

patent expiry (as was assumed here). As discussed in the introduction, this seems consistent with observed timing of contracts.

The perpetuity value of the revenue accruing to player i, when prices are P and r is the discount rate, and the fraction of an entrant's duopoly perpetuity consumed by entry costs are respectively

$$\Pi_i(P) = \frac{\pi_i(P)}{r} \text{ and } \kappa = \frac{F}{\Pi_i(P_i^d)}.$$

In non-trivial cases, *single entry is profitable* after accounting for costs, thus $\kappa \in (0,1)$. If a single generic firm enters, its net profit is

$$\Pi_i(P_i^d) - F = (1-\kappa)\Pi_i(P_i^d).$$

The real-time between each period is $\Delta > 0$. In our model Δ captures the real-time length of the marketing authorization and entry process. In Europe, for example, the marketing authorization process alone is close to 8 months.[13] The incumbent's payoffs is

$$U_I = \int_0^\infty \pi_I(P(\tau))e^{-r\tau}d\tau + \sum_{i\in\{1,2\}} b_i a_i^I,$$

the payoff of each entrant $i \in \{1,2\}$ is

$$U_i = \int_0^\infty \pi_i(P(\tau))e^{-r\tau}d\tau - F\sum_{t=0}^{\infty} a_t^i e^{-r\Delta t} - (b_i + F)a_i^I.$$

The consumer surplus when prices are P in a one shot game is $cs(P)$, and its perpetuity value is $CS(P)$. We assume that each subsequent entry increases consumer surplus, thus

$$cs(P^m) < cs(P_i^d) < cs(P^c).$$

The consumer surplus in the overall game is

$$CS = \int_0^\infty cs(P(\tau))e^{-r\tau}d\tau,$$

and social welfare is

$$W = \sum_{i\in\{I,1,2\}} U_i + CS.$$

The game has complete information. A strategy for an entrant is an early agreement bid, an entry decision, and price pair for every possible history of the game. A strategy for the incumbent is a rule deciding on the acceptance and refusal of bids, and a price for every possible history of the game.

A game's history comprises the early entry bids, the acceptance or refusal of those bids, and the sequence of entry and pricing decisions up to a period's stage. The payoff relevant history is simpler: at the early entry stage it is the entrants' bid pair, and for all subsequent stages it is the set of entrants that have become active prior to the stage under consideration.

[13]The most common form to apply for a marketing authorization in the European Union is through the centralized procedure, with applicants filing their documents with the European Medicines Agency (EMA), which drafts a report within 210 days, and the European Commission mostly following the EMA's recommendation on the marketing authorization a month later.

We denote the payoff relevant history by $h(t, s)$, where t is the period and $s \in \{1, 2\}$ the period's stage.

A Markovian strategy for the incumbent is a rule specifying acceptance and refusal of offers as a function of the bid pair (b_1, b_2), and subsequent price choices that depend only on the set of active entrants. A Markovian strategy for an entrant is an initial bid and an entry and price decision that only depends on the set of active entrants.

We focus on Markov Perfect Equilibria (MPE), specifically symmetric (SMPE). We refer to SMPE whenever we simply refer to *the equilibrium* of a (sub)game. The Markovian assumption avoids collusion issues that can occur in infinite horizon games. The focus on symmetry does not represent an interest on symmetry per se (as the qualitative features of the symmetric equilibrium survive the introduction of economic asymmetries), nor a focus on its mixed strategies (as the equilibrium which we focus on can be purified).[14] Instead, it captures the strategic uncertainty that should arise in a regulatory framework that keeps marketing authorization applications secret.[15] This uncertainty results in uncoordinated entry choices that are captured by such equilibria, and which seem more plausible than the alternative of coordinated entry decisions (which are captured by the asymmetric equilibria, or the outcome equivalent case of sequential entry choices).[16]

2.2 Preliminary results

As mentioned above, the focus on Markovian strategies is made so that, even if the game's horizon is infinite, multiple entry is always followed by intense price competition, as this is the only equilibrium of the repeated stage game:

Lemma 1. In any MPE, if both entrants have entered by the pricing stage of t, i.e., if $h(t, 2) = \{1, 2\}$, then firms choose prices P^c at every $t' \geq t$.

Therefore, once a single generic entry has taken place, the market becomes unattractive to an additional entrant, as that entrant would be unable to recover its entry cost. Thus:

Lemma 2. In any MPE, if a single generic firm i has entered by the entry stage of some period t, i.e. if $h(t, 1) = \{i\}$ with $i \in \{1, 2\}$, then at every $t' \geq t$ firm $j \neq i$ chooses $a_j^{t'} = 0$ and prices are P_i^d.

Thus both single and double entry are absorbing states. This is not the case for states where no entrant is active, discussed next. For the remainder, *equilibrium* refers to SMPE.

Lemma 3. In the equilibrium of a no entry subgame, i.e. when $h(t, 1) = \{\emptyset\}$, each firm $i \in \{1, 2\}$ chooses $a_i^t = 1$ with probability $(1 - \kappa)$ and its expected payoff is zero.

[14] In the former case, if the entrants have slightly different entry costs then there would still be a mixed strategy with similar features to the symmetric one we study here. In the latter case, if there exists at least some amount of uncertainty concerning the fixed entry costs, then an argument along the lines of Harsanyi (1973) can be used to purify the mixed strategy equilibrium, with firms deciding on an entry date (conditional on their rival having not entered before) based on their idiosyncratic entry cost realization. The distinction between pure and mixed strategy equilibria is then a rather artificial one.

[15] Importantly, in the EU, all market applications remain confidential until the point where the EMA emits its recommendation to the Commission.

[16] For completeness, we derive asymmetric equilibria, and discuss them in the conclusion. The main take away is that early entry agreements then have no effect on consumers nor welfare. The agreements would then still be used by the incumbent, but only as a tool that extracts entrants' profits.

The proof follows steps similar to e.g., Elberfeld and Wolfstetter (1999), and is omitted. If no entry has occurred before the first stage of period t, with probabilities $2\kappa(1-\kappa)$ and $(1-\kappa)^2$ the game respectively transitions to absorbing states with single or with double entry, and with probability κ^2 there is still no entry at t. The incumbent can therefore expect to keep its monopoly position beyond t for $(1-(1-\kappa)^2)^{-1}$ periods, before eventually one or both generic firms enter. Once this happens, all active firms choose the prices of the associated one-shot price game forever after, and no additional entry takes place. The expected payoff of the incumbent at the outset of that period t is then

$$A = (1-\kappa)^2\Pi_I(P^c) + 2\kappa(1-\kappa)\Pi_I(P_i^d) + \kappa^2(\int_0^\Delta \pi_I(P^m)e^{-r\tau}d\tau + Ae^{-r\Delta}),$$

which solving for A gives

$$A = \frac{(1-\kappa)^2\Pi_I(P^c) + 2\kappa(1-\kappa)\Pi_I(P_i^d) + \kappa^2(1-e^{-r\Delta})\Pi_I(P^m)}{1-\kappa^2 e^{-r\Delta}}.$$

The next lemma summarizes what happens if an early entry agreement is not reached:

Lemma 4. The subgame that follows no early entry agreement (i.e., $a_1^I = a_2^I = 0$) has a unique equilibrium: conditional on the game's history, firms play according to Lemmas 1-3. The incumbent remains a monopoly for an expected real-time length of $\Delta(1-(1-\kappa)^2)^{-1}$, its expected payoff is $U_I^{NA} = A$, and the entrants' expected profit is zero.

It is instructive to consider the limiting cases where $\Delta \to \infty$ and $\Delta \to 0$. The first case coincides with the outcome of a "one shot" entry opportunity game, and

$$\lim_{\Delta\to\infty} U_I^{NA} = (1-\kappa)^2\Pi_I(P^c) + 2\kappa(1-\kappa)\Pi_I(P_i^d) + \kappa^2\Pi_I(P^m).$$

The second case is when entry takes a "twinkle of an eye": the payoff of the incumbent becomes a weighted average of its payoffs in the two absorbing states of duopoly and competition, with weights that reflect the relative per period frequency of single and double entry, i.e.,

$$\lim_{\Delta\to 0} U_I^{NA} = \gamma\Pi_I(P^c) + (1-\gamma)\Pi_I(P_i^d) \text{ where } \gamma = \frac{(1-\kappa)^2}{1-\kappa^2}.$$

In general, we can conveniently write the incumbent's payoff as a weighted average of its monopoly profit and the latter limiting profit, i.e.,

$$A = \frac{(1-\kappa)^2(\gamma\Pi_I(P^c) + ((1-\gamma))\Pi_I(P_i^d)) + \kappa^2(1-e^{-r\Delta})\Pi_I(P^m)}{1-\kappa^2 e^{-r\Delta}}.$$

As the weight on the monopoly profit increases in Δ, A also increases in Δ (proof in the Appendix). Thus, even if it is possible for either generic firm to enter immediately at the patent expiry date (by applying for authorization beforehand), the incumbent still benefits from a lengthier authorization and entry process as this leads to an increase in the real-time period over which it can remain a monopoly post-patent expiry. A also increases in κ (proof in the Appendix). Thus, conditional on not reaching an early entry agreement, the incumbent's payoff increases with F: the short-run monopoly period is then likely to last longer, and in the long-run it becomes less likely that both generics will enter.

2.3 Early entry outcomes

By refusing both early entry offers, the incumbent extends its monopoly position beyond patent expiry in the short-run. However it also takes the risk of facing intense competition with low long-run profits—as there is a strictly positive probability that both generics may enter. By reaching an early entry agreement, the incumbent guarantees that the industry becomes a duopoly at patent expiry, and remains so forever after. The conditions under which an agreement is reached in equilibrium will be determined next.

Since entrants are symmetric, conditional on accepting one of the offers, the incumbent always selects the highest offer. The profit of entrant i is zero when its offer is refused.[17] Therefore, in any equilibrium where an agreement is reached, competition between the two entrants ensures that the bid of each entrant i is equal to i's duopoly profit, i.e.[18]

$$b_i = \Pi_i(P_i^d) - F = (1-\kappa)\Pi_i(P_i^d) \text{ for } i \in \{1, 2\}.$$

The next corollary summarizes equilibrium behavior:

Lemma 5. If an agreement is to be reached in equilibrium, then $b_1 = b_2 = (1-\kappa)\Pi_i(P_i^d)$ and no additional entry takes place after patent expiry—thus prices are P_i^d at every $t \geq 0$. If both offers are refused, then after patent expiry all firms play according to Lemma 4.

Therefore, if an offer is accepted, the incumbent fully appropriates the industry duopoly profit and its profit is

$$U_I^A = \Pi_I(P_i^d) + (1-\kappa)\Pi_i(P_i^d).$$

If the incumbent refuses both offers, both entrants make zero profit and the incumbent's profit is the aggregate industry profit. Thus the incumbent accepts an offer if and only if immediate entry increases the industry profit relative to uncoordinated entry, i.e., when

$$U_I^A \geq U_I^{NA} \Leftrightarrow \Pi_I(P_i^d) + (1-\kappa)\Pi_i(P_i^d) \geq A.$$

The next result characterizes the conditions under which this condition holds:

Proposition 1. The incumbent accepts an early entry offer (in the unique equilibrium) if and only if entry costs are not too high, or the authorization and entry process is not too long: There exists a $\overline{\kappa} \in (0, 1)$ such that $U_I^A \geq U_I^{NA}$ for all Δ if $\kappa \leq \overline{\kappa}$, and for all $\Delta < \overline{\Delta}(\kappa)$ when $\kappa \in (\overline{\kappa}, 1)$ ($\overline{\Delta}(\cdot)$ being strictly positive and decreasing in its range).

To get an intuitive understanding, it is useful to consider limiting cases. In the absence of an agreement, entry by both entrants immediately after patent expiry is almost certain when F (and thus κ) is arbitrarily close to zero. Then, the right-hand side of the previous inequality is close to $\Pi_I(P^c)$ regardless of Δ, and so the incumbent always accepts an offer as $\Pi_I(P_i^d) + \Pi_i(P_i^d) > \Pi_I(P^c)$.

On the other hand, the entry probability $(1-\kappa)$ is close to zero when F is close to $\Pi_i(P_i^d)$ (and thus κ is close to 1). The entrants' lack of coordination then acts as a barrier

[17] Its expected profit when both offers are refused, and also its profit when j's offer is accepted.

[18] In analogy with the Bertrand game, this is the unique outcome of the game with mixing of bids (see Harrington (1989)). Klemperer (2003) popularized the idea that the perfectly competitive outcome is not the only equilibrium of the standard Bertrand game, but this is true only if monopoly profits are unbounded (see Kaplan and Wettstein (2000)). Here such equilibria cannot exist since this would require the incumbent to accept arbitrarily negative bids—which are rationally refused.

to entry, keeping the incumbent a monopoly for a significant time beyond the patent expiry date. The incumbent should then reject both offers, to essentially capture an amount close to the monopoly perpetuity instead of the aggregate duopoly profit.

The latter heuristic argument does however fail for Δ sufficiently small. In that case, in the absence of an agreement, entry by either one or both firms always takes place in the "twinkle of an eye" (even for κ close to 1), and therefore the incumbent makes no more than $\Pi_I(P_i^d)$. The incumbent should then always accept an offer to instead capture the aggregate industry duopoly profit while extracting the single entry profit.

A testable prediction of this model is that faster application procedures should be associated with a higher incidence of early entry agreements, as these become more prevalent when either Δ is low, situations in which the incumbent is likely to loose its monopoly position soon after the patent expiry date.

As explained next, once agreements become possible it is no longer the case that the incumbent always benefits from a higher entry cost. Recall that, when no agreement is reached (which is the case when F is sufficiently high), the incumbent's profit increases with F. We now learned that, if an agreement is reached (which is the case when F is sufficiently low), the rent extracted from entrants decreases with F, and thus the incumbent's profit decreases with F. Thus the relationship between the incumbent's profit and F becomes U-shaped once early agreements are allowed.

We consider next the effect on consumers. If no entry occurred before the first stage of period t, the expected consumer surplus from that period t onwards is

$$B = (1-\kappa)^2 CS(P^c) + 2\kappa(1-\kappa)CS(P_i^d) + \kappa^2 (\int_0^\Delta cs(P^m)e^{-r\tau}d\tau + Be^{-r\Delta}).$$

Solving for B gives

$$B = \frac{(1-\kappa^2)(\gamma CS(P^c) + (1-\gamma)CS(P_i^d)) + \kappa^2(1-e^{-r\Delta})CS(P^m)}{1-\kappa^2 e^{-r\Delta}}.$$

This is a weighted average of the consumer surplus under monopoly, and the expected consumer surplus in the limiting case where Δ tends to zero. As the weight on the monopoly outcome increases with Δ, B is decreasing in Δ. Thus, in the absence of an agreement, consumers are hurt by a lengthier entry process: this increases the real-time period over which they face monopoly prices in the short-run, even if it does not change the relative probability that either one or two generics will enter in the long-run.

An early entry agreement hurts consumers if and only if $B \geq CS(P_i^d)$. We obtain:

Proposition 2. Early entry agreements hurt consumers if either the entry costs are not too high, or the entry process is not too long (and it benefits them otherwise): There exists a $\widehat{k} \in (0,1)$ such that $B \geq CS(P_i^d)$ for all Δ if $\kappa \leq \widehat{k}$, and for all $\Delta < \widehat{\Delta}(\kappa)$ when $\kappa \in (\widehat{\kappa}, 1)$ ($\widehat{\Delta}(\cdot)$ being strictly positive and decreasing in its range).

The findings so far can be summarized as follows: An early entry agreement is reached if and only if it increases industry profits relative to the alternative of no agreement. This is the case if either entry costs are sufficiently small or the entry process is sufficiently fast. In those instances, in the absence of an agreement, some entry (single or double) occurs very fast. The agreement then replaces with a sure duopoly what would likely have been a more competitive outcome. Thus, in those same instances agreements hurt consumers. Industry and consumer interests therefore tend to be misaligned.[19]

Taking a welfare perspective, a reasoning similar to the case of industry profits and consumer surplus shows that the welfare in the absence of an agreement is a weighted average of monopoly welfare and (since the entrants' profits are dissipated by uncoordinated entry) the sum of the consumer surplus and the incumbent's profit in the limiting case where entry takes the "twinkle of an eye". We obtain:

Proposition 3. Allowing for early entry agreements (weakly) increases welfare for all delta if, in terms of welfare, i) duopoly is preferred to monopoly, and ii) an agreement is preferred to no agreement for arbitrarily small delta.

If these two conditions are satisfied, an agreement would increase welfare but may not be reached in equilibrium.[20] The first condition is reasonably satisfied in most models. The second condition is not standard and thus, a more complete understanding of the welfare effect of agreements requires introducing additional structure on the economic fundamentals (and in particular on demand). We take this task in the next section, in the context of two benchmark models of price competition.

3 Early entry agreements and welfare

Generic drugs are typically sold at a lower price than the incumbent's. Despite this, incumbents often maintain a significant market share. This suggests that, while branded and generic drugs may seem similar from a bioequivalence point of view, consumers do not perceive the original and generic products to be perfect substitutes. As presented in the introduction, in controlled trials consumers report greater effects with branded products. Consumer brand loyalty is also well documented. Indeed, some demographic segments seem reluctant to switch to generics, and other segments do not seem to hold generics in their consideration set.[21]

We study two benchmark models of price competition that are consistent with these facts. The first one is Bertrand with vertical differentiation of the Mussa and Rosen (1978) type: The incumbent's product is perceived to be of higher quality than generics, but generics are perceived to be perfect substitutes to each other. The second is a Bertrand model with captive consumers of the Varian (1980) type: The incumbent has a fraction of loyal or captive consumers, who never consider switching to a lower priced generic, but the remainder of the consumers perceive all products to be perfect substitutes—and thus, if they buy, they always choose the lowest priced product.

All firms have the same marginal cost of production, which is assumed to be zero without loss of generality. An implication of this is that in both models the incumbent does not wish to exploit consumer heterogeneity during the patent period by launching its own generic version during the early patent period—which also seems consistent with the general industry behavior (see e.g., EC (2009)).

[19] See online appendix for an example with benefits to both consumers and the industry.

[20] As this choice is determined by the incumbent's interests, rather than welfare.

[21] For example, Bronnenberg et al. (2015) find that better informed buyers, such as medical staff but also kitchen chefs, are less likely to buy branded medicines. The average consumer chooses national brands 26 percent of the time, while pharmacists would do so only 9 percent of the time.

3.1 Early entry with vertical differentiation

The incumbent produces a product of quality $s_I = 1$ (normalized without loss of generality), and entrants produce a generic of lower quality $s_L < 1$. With the Mussa and Rosen (1978) linear-multiplicative utility, the net utility of a consumer with taste for quality θ from buying quality s at price p is

$$U = \theta s - p.$$

There is a continuum of consumers, and their tastes are uniformly distributed in $(0, 1]$. We refer to a consumer with taste θ as consumer θ.

Take some vector of prices P. A consumer θ^* is indifferent between buying the generic or not buying any good when $\theta^* s_L - \min\{p_1, p_2\} = 0$, a consumer θ^{**} is indifferent between buying the generic and buying from the incumbent when $\theta^{**} s_L - \min\{p_1, p_2\} = \theta^{**} - p_I$, and a consumer θ^{***} is indifferent between buying from the incumbent or not buying when $\theta^{***} - p_I = 0$. These consumers are respectively

$$\theta^* = \frac{\min\{p_1, p_2\}}{s_L},\ \theta^{**} = \frac{p_I - \min\{p_1, p_2\}}{1 - s_L}, \text{ and } \theta^{***} = p_I.$$

The demands of the incumbent and of a generic firm i, provided they are non-negative, are

$$D_I(P) = \begin{cases} 1 - \theta^{**}(P) \text{ if } p_I > \frac{\min\{p_1,p_2\}}{s_L} \\ 1 - \theta^{***}(P) \text{ otherwise.} \end{cases}$$

and

$$D_i(P) = \begin{cases} \theta^{**}(P) - \theta^*(P) \text{ if } p_I > \frac{p_i}{s_L} \text{ and } p_i < p_j \\ \frac{1}{2}(\theta^{**}(P) - \theta^*(P)) \text{ if } p_I > \frac{p_i}{s_L} \text{ and } p_i = p_j \\ 0 \text{ otherwise.} \end{cases}$$

Under monopoly

$$P^m = \left(\frac{1}{2}, \emptyset, \emptyset\right),\ D_I(P^m) = \frac{1}{2} \text{ and } \pi_I(P^m) = \frac{1}{4},$$

under a duopoly with entrant i prices are

$$P_i^d = \left(\frac{2(1 - s_L)}{4 - s_L}, \frac{s_L(1 - s_L)}{4 - s_L}, \emptyset\right),$$

demands are

$$D_I(P_i^d) = \frac{2}{4 - s_L} \text{ and } D_i(P_i^d) = \frac{1}{4 - s_L},$$

and per period revenues are

$$\pi_I(P_i^d) = \frac{4(1 - s_L)}{(4 - s_L)^2} \text{ and } \pi_i(P_i^d) = \frac{s_L(1 - s_L)}{(4 - s_L)^2}.$$

When all three firms are in competition, prices are

$$P^c = \left(\frac{1 - s_L}{2}, 0, 0\right),$$

demands are

$$D_I(P^c) = \frac{1}{2} \text{ and } D_1(P^c) = D_2(P^c) = \frac{1}{4}$$

and revenues are

$$\pi_I(P^c) = \frac{1-s_L}{4} \text{ and } \pi_1(P^c) = \pi_2(P^c) = 0.$$

Simple computations show that each additional entry reduces the *aggregate* revenue, i.e.,

$$\pi_I(P^m) > \pi_I(P_i^d) + \pi_i(P_i^d) > \pi_I(P^c).$$

As variety weakly increases and prices strictly decrease with each additional entry, by revealed preferences the consumer surplus is strictly increasing in generic entry. The conditions of the model of the previous section are thus met.

Surprisingly, in the context of this benchmark model, we find that agreements are always welfare improving, even when entry is virtually costless and perfect competition by generics is essentially reached for free in the absence of an agreement—with lower prices for both the original and generic drugs. Indeed, we have:

Proposition 4. In the game with vertical differentiation, allowing for early entry agreements always increases welfare.

Studying welfare typically involves considering all three potential states (no entry, single entry and double entry) with the associated equilibrium probabilities. In the present model we avoid such complications as in the next section we prove a novel result in the vertical differentiation literature: Whenever single entry is profitable, the entry of the first generic increases welfare, but any additional generic entry decreases welfare. In other words, the social optimal number of generics is one.[22]

3.1.1 The optimal number of generics is one

Entry costs typically lead to socially inefficient entry (see e.g. Spence (1976), Perry (1984) and Mankiw and Whinston (1986)). We look for the welfare-maximizing number of firms, taking as given their non-cooperative pricing behavior after entry. In contrast with the extant literature, we find that a social planner finds it desirable to limit entry even in the absence of entry costs—and regardless of the generic's quality:

Proposition 5. In a textbook vertical differentiation model, if entry is profitable then the social optimal number of generic entrants is one.

In a nutshell, the result can be explained as follows. Entry lowers all prices. Lower generic prices lead some (low valuation) consumers who were not purchasing any good to start purchasing the generic good. However, with the entry of the second generic, a sharp decline in all prices results in the increase of the *relative* price of high quality. This relative price change induces (medium valuation) consumers, who were already purchasing the high quality good, to switch to the relatively cheaper generic. The welfare gain made with the new buyers of low quality, which is a fortiori bound to be small (since those new consumers do not value the good much in the first place), is always outweighed by the welfare loss made by the switch by medium valuation buyers from high to low quality. This is true even when entry costs are zero. Due to cost duplication, accounting for entry costs further reinforces this result.

[22] Thus an agreement is always welfare improving, as it guarantees the immediate entry of a single generic.

The proof is instructive, and therefore presented in the main text next. For completeness, we start by comparing the single entry outcome to the no entry outcome. We then compare the single entry outcome to the double entry one.

In the no entry situation, the high quality monopoly serves half of the market, i.e., consumers $\theta \in \left[\frac{1}{2}, 1\right]$. Upon entry of one generic, one more product is available and the high-quality product becomes available at a lower price. By revealed preferences, consumers are better off and the incumbent is worse off. Moreover, the incumbent fights the first entrant: it lowers its price to the point of expanding the segment of the market it serves, as then $(1-\theta^{**}) > \frac{1}{2}$. Thus the incumbent moves from a monopoly niche strategy (serving only the highest valuation consumers) to a mass market strategy (serving both high and medium valuation consumers).

Welfare is given by the total economic value created by consumption minus entry costs. Single entry increases the economic value of consumption, as all consumers who used to purchase the high quality good still purchase that good, and some consumers who did not previously purchase any good now buy some good. Moreover, if entry is profitable, the economic value created by additional consumption always exceeds the entry cost: the entrant captures as revenues only a fraction of the value created by the consumption of its good, which must exceed the entry cost. Thus single entry increases welfare, if profitable.

Next we consider the effect of a second generic. The argument that, even before accounting for fixed costs, single generic entry is better than double generic entry can be understood geometrically in three steps using Figure 1.[23] On the x-axis we have the identity of each consumer θ, and on the y-axis the value created by the consumption of that consumer (which takes the value θs_I, θs_L, or 0 depending on whether θ respectively consumes the original, the generic, or no good). A x-axis segment represents the market segment served by a good, and the corresponding integral the value generated by the consumption of that good. For instance, the value created by the original good under duopoly corresponds to the area below the s_I line in the segment $(1-\theta^{**})$.

First, with two generics, intense competition brings the generic price to zero. Therefore all consumers purchase either the low or the high quality good. The segment of consumers who did not purchase before and now do, ranges from 0 to θ^* in duopoly, and the value created by such additional consumption (captured by the area of the triangle ABC) is

$$\frac{\theta^* s_L}{2} = \frac{s_L}{2}\left(\frac{1-s_L}{4-s_L}\right)^2.$$

Second, given that the generic becomes available for free, the incumbent faces a demand that reflects the consumers' willingness to pay to upgrade from the free generic to the high quality good, which is given by $1 - \frac{p_I}{1-s_L}$. As this demand is linear, the incumbent finds it optimal to again serve only half of the consumer segment—as it did under monopoly. Thus those consumers with valuations $\theta \in \left[\frac{2-s_L}{4-s_L}, \frac{1}{2}\right)$, which under single entry were buying high quality, switch to the low quality good. The segment that downgrades has size

$$D_I(P_1^d) - D_I(P^c) = \frac{s_L}{2(4-s_L)},$$

Third, the consumption value lost by each consumer that downgrades is $(1-s_L)\theta$. The loss due to this quality downgrade is thus captured by the area of the polygon EFJH. In turn, this economic loss exceeds the loss of consumption of the least eager of those consumers

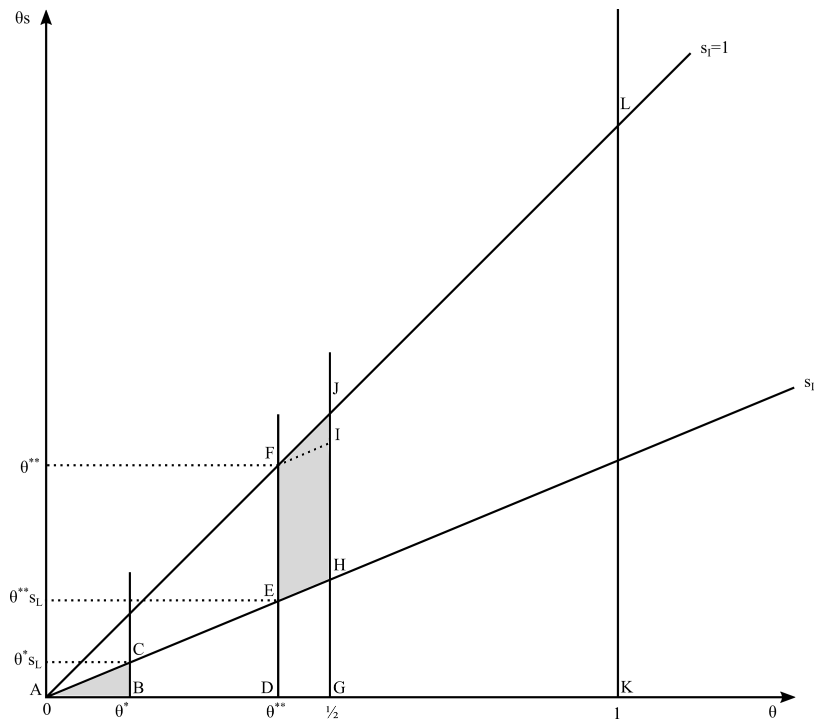
[23] In the plot of figure 1, s_I is normalized to one and $s_L = 0.5$

multiplied by the segment of those consumers, which is captured by the area of the polygon EFIH. The latter exceeds the gain from new low quality consumers that we saw above is captured by the triangle ABC area, i.e.,

$$(1-s_L)(\frac{2-s_L}{(4-s_L)})\frac{s_L}{2(4-s_L)} > \frac{s_L}{2}(\frac{1-s_L}{4-s_L})^2 \Leftrightarrow 2-s_L > 1-s_L.$$

This loss is amplified by replication of entry costs, which completes the proof.

Figure 1: Welfare



3.2 Early entry with loyal consumers

In our next model we consider that overall market demand is given by $1-p$, but a fraction μ of the consumers only considers purchasing from the incumbent (its captive segment). The remaining fraction $1-\mu$ of consumers, the shoppers, purchases from the lowest priced firm when buying (to simplify the exposition, we assume that, at the same price, shoppers purchase a generic). For a given price vector P, the demand of the incumbent I and of a generic firm i, provided they are non-negative, are therefore

$$D_I(P) = \begin{cases} 1-p \text{ if } p_I < \min\{p_1, p_2\} \\ \mu(1-p) \text{ otherwise.} \end{cases}$$

and

$$D_i(P) = \begin{cases} (1-\mu)(1-p) \text{ if } p_i < p_j \text{ and } p_i < p_I \\ \frac{1}{2}(1-\mu)(1-p) \text{ if } p_i < p_I \text{ and } p_i = p_j \\ 0 \text{ otherwise.} \end{cases}$$

The overall market monopoly price is thus $p^m = \frac{1}{2}$, and with a monopolist incumbent

$$P^m = \left(\frac{1}{2}, \emptyset, \emptyset\right), \ D_I(P^m) = \frac{1}{2}, \ \pi_I(P^m) = \frac{1}{4}, \text{ and } cs(P^m) = \frac{1}{8}.$$

When all three firms are in the market, competition among generics brings their price down to zero. The incumbent then finds it optimal to focus on its captive consumers only, and to charge them the monopoly price. Thus prices and demands are

$$P^c = (\frac{1}{2}, 0, 0), \ D_I(P^c) = \frac{1}{2}\mu, \text{ and } D_i(P^c) = \frac{1}{2}(1-\mu),$$

and the revenues and consumer surplus are

$$\pi_I(P^c) = \frac{1}{4}\mu, \ \pi_i(P^c) = 0, \text{ and } cs(P^c) = \frac{1}{2}(1-\mu) + \frac{1}{8}\mu.$$

The less obvious outcome is duopoly, which is presented next. The duopoly model is equivalent to an asymmetric version of Varian's model of sales with elastic demand. There is a unique equilibrium, and it involves mixed strategies that capture the firms' incentives to remain unpredictable. In equilibrium the generic firm charges on average lower prices than the incumbent, and therefore most of the time the incumbent sells only to its captive segment. However, generic prices remain sufficiently high to attract the incumbent to occasionally hold sales, as it tries to capture the shopper segment in addition to its captive segment. The next result characterizes the equilibrium behavior:

Proposition 6. In the unique duopoly equilibrium of the captive consumer model, each firm randomizes its price over a common support $\left[\frac{1}{2}(1-\sqrt{1-\mu}), p^m\right)$, using the CDFs

$$G_I(p) = 1 - \frac{\mu}{4p(1-p)} = (1-\mu)G_i(p),$$

and the incumbent sets p^m with the remaining probability $1 - G_I(p^m) = \mu$. The *expected* revenues are

$$\pi_I(P_1^d) = \frac{1}{4}\mu \text{ and } \pi_1(P_1^d) = \frac{1}{4}(1-\mu)\mu.$$

The entrant has a more aggressive pricing strategy than the incumbent, and therefore it has a higher probability of charging lower prices, since $G_I(p) = (1-\mu)G_i(p)$. Both CDFs decrease with μ, so both firms become less aggressive in their pricing as the fraction of captive consumers increases. The consumer welfare in both segments is therefore strictly decreasing in the fraction of captive consumers.

We have that

$$\pi_i(P_i^d) > \pi_i(P^c) = \pi_i(P_j^d) = 0 \text{ for } i \in \{1,2\} \text{ and } j \neq i,$$

and that

$$\pi_I(P^m) > \pi_I(P_i^d) = \pi_I(P^c) > 0.$$

Note that the incumbent makes the same revenue with one or two entrants. The reason is that under duopoly the incumbent is essentially deciding whether to set a low price to steal the shopper segment form the entrant, or to focus on its captive segment only (and charge them the monopoly price). As the unique equilibrium is in mixed strategies, it must be indifferent between these two options. The latter one is however the option the incumbent takes if there are two entrants. Therefore it must make the same profit in both cases.

All revenue assumptions of the general model of the previous section are satisfied, as each additional entry also reduces the *aggregate* revenue

$$\pi_I(P^m) > \pi_I(P_i^d) + \pi_i(P_i^d) > \pi_I(P^c).$$

Next we focus on consumers. In duopoly the captive consumers face a price of p^m with probability μ, and lower prices with the complementary probability. In monopoly and competition these consumers face with certainty the monopoly price. Thus captive consumers are better off in duopoly than under competition or monopoly. On the other hand, under duopoly the shoppers can purchase at the lowest of the two prices, which always exceeds the competitive level but remains below the monopoly price. Shoppers are therefore better off in duopoly than monopoly, but they are even better off under competition. Therefore additional entry always benefits shoppers but not necessarily the captive consumers.

Different consumer segments thus have different preferences over early entry agreements, as they always benefit captive consumers but they may hurt shoppers. It can be shown that (on aggregate) the assumption of the previous section on the consumer surplus is satisfied:

Lemma 5. In the captive consumer model, $cs(P^c) > cs(P_i^d) > cs(P^m)$.

Given the conditions above, all the results of the general model of the previous section apply to the captive consumer model as well.

Next we consider the welfare effect of allowing for early entry agreements. Unfortunately, in the present model we need to take the exact state transition probabilities into account. Given the issuing additional complexity, we use Proposition 3 to derive some conclusions.

Since welfare is higher for any realization of prices in duopoly than under monopoly, in terms of welfare an early entry agreement is preferred to monopoly. To derive sufficient conditions under which allowing for early entry agreements increases welfare, we therefore only need to study the limiting case where entry takes the "twinkle of an eye", i.e., where $\Delta \to 0$. In that case an agreement would always be reached in equilibrium with some entrant i, and the welfare is deterministic and given by

$$\lim_{\Delta \to 0} W^A = \Pi_I(P_i^d) + (1-\kappa)\Pi_i(P_i^d) + CS(P_i^d).$$

Since the expected profits of the entrants are zero in the absence of an agreement and $\Pi_I(P^c) = \Pi_I(P_i^d)$, the limiting expected welfare if an agreement is not allowed is

$$\lim_{\Delta \to 0} W^{NA} = \Pi_I(P^d) + \gamma CS(P^c) + (1-\gamma)CS(P_i^d).$$

Recalling that $\gamma = (1-\kappa)^2/(1-\kappa^2)$, in the limiting case where entry takes the "twinkle of an eye", welfare is reduced by allowing for agreements if and only if

$$CS(P^c) - CS(P_i^d) > \frac{1-\kappa^2}{1-\kappa}\Pi_i(P_i^d).$$

Using this condition we show:

Proposition 7. In the captive consumer model, in the limiting case where $\Delta \to 0$, an early entry agreement increases welfare if either the share of captives is not too high (i.e., for all κ if $\mu \leq \mu^* \simeq 0.76$) or if the entry cost is not too low (i.e., for every $\mu > \mu^*$ there exists a unique $\kappa(\mu) \in (0, 1)$ such that the agreement increases welfare if and only if $\kappa \geq \kappa(\mu)$, and $\kappa(\cdot)$ is a continuous and decreasing function). Under the same conditions, for every delta, allowing for early entry agreements increases welfare.

The intuition is the following. When most of the consumers are shoppers, a single entrant is all that it takes to have frequent and steep discounts, and both prices to be close to marginal cost almost all the time (as expected in the standard Bertrand duopoly case with no captive consumers). In that case there is little benefit to consumers (in aggregate) from additional entry, but there is a reduction in the industry profit. This profit reduction also reflects a replication of entry costs. Thus early entry agreements tend to increase welfare when there are few captive consumers or entry costs are not too low. The last statement, that under the same conditions, for every delta, allowing for early entry agreements increases welfare, follows from Proposition 3.

4 Discussion and conclusion

We studied the strategic incentives of an originator to reach an early entry agreement with a generic firm, and showed that they can have anticompetitive effects: While early generic entry has a short-run procompetitive effect, it also reduces the incentives for other firms to enter and can lead to less competition in the long-run. When agreements are reached in equilibrium, the tend to harm consumers, depriving them from more entry and lower prices.

Reaching welfare conclusions required additional structure on fundamentals, and we studied two benchmarks: a vertically differentiated goods model, and a homogeneous goods model with loyal consumers and shoppers.

In the standard vertical differentiation setup we found, surprisingly, that having a single entrant is always socially optimal—regardless of the entry cost. When there is a single entrant, the incumbent chooses a quality adjusted price that leads medium valuation consumers, which do not purchase its good under monopoly, to buy its product. With a second entrant, the generics' price falls to marginal cost and the generic product becomes relatively less expensive than the product of the incumbent. This has two effects: Consumers with the lowest valuations, who were not consuming before, now purchase the generic. Consumers with medium valuations, who would have purchased the original following the entry of a single generic, instead buy the relatively cheaper but less valuable generic. It turns out that the first positive welfare effect of additional entry is always outweighed by the second negative effect.

In the standard vertical differentiation framework of Mussa and Rosen (1978) we use, the consumption value of each consumer is a multiplicative factor of its taste for quality by the good's quality. One may have expected that multiple entry would become more desirable if a good also has an intrinsic value, regardless of its quality. However the opposite is true, as we explain next. Once the intrinsic value becomes sufficiently high, a single entrant is sufficient for the market to be covered (i.e., for all consumers to purchase in equilibrium). Then additional entry no longer increases the number of consumers served, but it still leads less consumers to purchase high quality and to a duplication of the fixed costs.

This result also assumes a uniform distribution of consumers' tastes. We conjecture that if relatively more consumers have a taste for intermediate or high quality then the negative welfare effect of entry will be exacerbated, but it is also possible that the result could be reversed if instead the taste distribution becomes sufficiently skewed towards low quality. Unfortunately the vast vertical differentiation literature does not offer a clear characterization of duopoly outcomes for other taste distributions, and so this must remain an open issue.

If, in this vertical differentiation model, there is quantity instead of price competition then the level of fixed costs becomes crucial—as shown in an online appendix. The welfare conclusions become more nuanced, and in line with the extant literature on excessive entry. When single entry is profitable, a monopoly can be socially optimal when entry costs are sufficiently high. When entry costs are low, whether single or multiple generic entry is the socially desirable outcome depends on the the generic's quality: the former is preferred to the latter if and only if that quality is sufficiently low.

In the second benchmark model with loyal consumers, early entry agreements are welfare increasing, provided the number of loyal consumers is not too high. This may seem surprising, since an agreement always benefits loyal consumers: these consumers benefit from occasional sales an originator holds to attract shoppers in duopoly, but charges them the monopoly price if instead multiple generics enter the market. However the frequency and extent of such sales, and therefore their social benefit, also decreases as the share of loyal consumers increases—which finally drives the result.

We focus on symmetric equilibria, which capture situations where the generic entry is uncoordinated. As discussed in the introduction, this seems to describe best the incentives and features of the authorization process of generic drugs. The game-theoretic literature on entry games also suggests that coordination may be hard to achieve, even when firms are able to revise entry decision ex-post or pre-play communication is possible (see e.g., Dixit and Shapiro (1986) and Farrell (1987)). Should such coordinated generic entry still be achieved, in the absence of an agreement, our model predicts that an asymmetric equilibrium with immediate entry by a single generic is always reached. It then becomes a dominant strategy for the originator to accept the highest early entry offer, which becomes a pure rent extraction tool that transfers industry rents from the entrant to the incumbent. In that case, consumer surplus and welfare are unaffected by agreements. In light of this, even if one is agnostic about which type of equilibrium is played, it remains that in this model early entry agreements tend to be (weakly) welfare increasing.

Asymmetric equilibria in our game also correspond to the outcomes of a sequential entry game as in Rockett (1990). In her work, an early entry agreement becomes a tool for rival selection, as the entrants' asymmetry creates a tradeoff for the originator between the licensing fee amount and how tough a rival entrant will be in the market. Such selection concerns are absent in our work with homogeneous rivals, but could be incorporated in the analysis. For instance, if the only source of heterogeneity are entry costs, the lowest cost entrant wins the auction for early entry. In this case, early entry agreements lead to a positive selection, as, if agreements are ruled out, it is possible that the least efficient generic instead enters (this in addition to the other inefficiencies discussed in our work).

A remark on our pessimistic prospect for consumer welfare of early entry agreements is that they still allow generic drugs to become available before patent expiry, and thus consumers should benefit from increased competition and lower prices at an earlier date with the agreements than without. As discussed in the introduction, in practice this benefit seems to be very short lived. A policy that is however likely make early entry agreements more beneficial to consumers is that authorized entry prior to patent expiry can only take place before some critical date, so that "early means early". The benefit from the temporary increase in competition can then at least mitigate the long-term effects from diminished competition after patent expiry. Such policies can thus be used to implement a socially desirable outcome, while transferring surplus from the industry to consumers.

A remark on our optimistic prospect for social welfare of early entry agreements is that we assumed the entrant is unconstrained in competition. In practice originators use a combination of license, supply, and distribution agreements to implement early entry rights.[24] Such practices could limit competition beyond patent expiry, and thus nuance our results. It may therefore be important to keep policies that limit the extent to which such clauses may limit competition after patent expiry—as is often done in practice, as after patent expiry antitrust authorities will challenge certain practices as collusive horizontal agreements.

Especially in the pharmaceutical industry, dynamic effects also seem relevant. An empirical literature, starting with works by e.g., Pakes (1985) and Hall et al. (1986), suggests that there is a strong positive relationship between R&D expenditures and patent strength and value. In our model, the incumbent can extract an entrant's profit with an early entry agreement. One may argue that this will also increase incentives to invest in R&D for new and better drugs, potentially benefiting consumers down the line. This link would strengthen the finding that early entry agreements increase total welfare, while challenging our result that they tend to harm consumers. Such dynamic considerations are typically not taken into account in the existing case law on entry agreements, and have been left out of our model as well. A more complete understanding of dynamic considerations, and their social welfare and consumer surplus effects, could be an interesting topic for future research.

[24] See for example paragraph 821 of EC (2009). Another common industry practice are authorized generics, whereby the incumbent produces the drug itself and markets it as a generic product. However, in the context of our model, an early entry by an independent generic would lead the originator to withdraw its authorized generic.

Appendix

Proof Proposition 1. The incumbent accepts an offer if and only if

$$U_I^A \geq U_I^{NA} \Leftrightarrow \Pi_I(P_i^d) + (1-\kappa)\Pi_i(P_i^d) \geq A.$$

When $\Delta \to \infty$, A converges to the result of the one-shot game, i.e., to

$$(1-k^2)(\gamma\Pi_I(P^c) + (1-\gamma)\Pi_I(P_i^d)) + \kappa^2\Pi_I(P^m).$$

Taking a second limit, the above converges to $\Pi_I(P_i^c)$ as $\kappa \to 0$ and to $\Pi_I(P_i^m)$ as $\kappa \to 1$. Thus the inequality $U_I^A \geq U_I^{NA}$ holds in the former but not in the latter case. When $\Delta \to 0$, A converges to the weighted average of the absorbing states

$$\gamma\Pi_I(P^c) + (1-\gamma)\Pi_I(P_i^d).$$

Taking a second limit, the above converges to $\Pi_I(P_i^c)$ when $\kappa \to 0$ and to $\Pi_I(P_i^d)$ when $\kappa \to 1$. Thus the inequality $U_I^A \geq U_I^{NA}$ holds for every κ. By continuity, $U_I^A \geq U_I^{NA}$ holds if κ is sufficiently small (irrespective of Δ) and when κ is sufficiently large it holds if and only if Δ is sufficiently small.

Given the above, and since A is strictly increasing in Δ and in κ, there further exists a $\overline{\kappa} \in (0,1)$ such that $U_I^A \geq U_I^{NA}$ for all Δ if and only if $\kappa \leq \overline{\kappa}$, and for all $\Delta < \overline{\Delta}(\kappa)$ if and only if $\kappa \in (\overline{\kappa}, 1)$ (being $\overline{\Delta}(\cdot)$ strictly positive and decreasing in its range). We next prove the monotonicity results.

$$\begin{aligned}\frac{\partial A}{\partial \Delta} > 0 \Leftrightarrow (r\kappa^2\Pi_I(P^m)e^{-r\Delta})(1-\kappa^2 e^{-r\Delta}) > \\ \kappa^2 r e^{-r\Delta}[(1-\kappa^2)(\gamma\Pi_I(P^c) + (1-\gamma)\Pi_I(P_i^d)) + \kappa^2\Pi_I(P^m)(1-e^{-r\Delta})] \\ \Leftrightarrow \Pi_I(P^m) > \gamma\Pi_I(P^c) + (1-\gamma)\Pi_I(P_i^d)\end{aligned}$$

Since the incumbent's profit is decreasing in additional entry, the last inequality holds.

$$\begin{aligned}\frac{\partial A}{\partial \kappa} &= \frac{(1-\kappa^2 e^{-r\Delta})[2(\kappa-1)\Pi_I(P^c) + (2-4\kappa)\Pi_I(P_i^d) + 2\kappa(1-e^{-r\Delta})\Pi_I(P^m)]}{(1-\kappa^2 e^{-r\Delta})^2} \\ &+ \frac{2\kappa e^{-r\Delta}[(1-\kappa)^2\Pi_I(P^c) + 2\kappa(1-\kappa)\Pi_I(P_i^d) + \kappa^2\Pi_I(P^m)(1-e^{-r\Delta})]}{(1-\kappa^2 e^{-r\Delta})^2}.\end{aligned}$$

Further note that

$$\frac{\partial A}{\partial \kappa \partial \Pi_I(P^c)} = \frac{-2\,(1-\kappa)\left(1-\kappa e^{-r\Delta}\right)}{(1-\kappa^2 e^{-r\Delta})^2} < 0$$

Replace $\Pi_I(P^c)$ by $\Pi_I(P_i^d) > \Pi_I(P^c)$ to obtain

$$\frac{\partial A}{\partial \kappa} > \frac{2\kappa\left(\Pi_I(P^m) - \Pi_I(P_i^d)\right)\left(1-e^{-r\Delta}\right)}{(1-\kappa^2 e^{-r\Delta})^2} > 0,$$

where the last inequality holds since $\Pi_I(P^m) - \Pi_I(P_i^d) > 0$.

Proof Proposition 2. The proof follows a similar structure to Proposition 1. Agreements hurt consumers if and only if $B \geq CS(P_i^d)$, or equivalently

$$\frac{(1-\kappa^2)(\gamma CS(P^c)+(1-\gamma)CS(P_i^d))+\kappa^2(1-e^{-r\Delta})CS(P^m)}{1-\kappa^2 e^{-r\Delta}} \geq CS(P_i^d).$$

When $\Delta \to \infty$, B converges to the result of the one-shot game, i.e., to

$$(1-k^2)(\gamma CS(P^c)+(1-\gamma)CS(P_i^d))+\kappa^2 CS(P^m).$$

Taking a second limit, the above converges to $CS(P^c)$ as $\kappa \to 0$ and to $CS(P^m)$ as $\kappa \to 1$. Thus the inequality $B \geq CS(P_i^d)$ holds in the latter but not in the former case. When $\Delta \to 0$, B converges to the weighted average of the absorbing states,

$$\gamma CS(P^c)+(1-\gamma)CS(P_i^d).$$

Taking a second limit, the above converges to $CS(P^c)$ when $\kappa \to 0$ and to $CS(P_i^d)$ when $\kappa \to 1$. Thus the inequality $B \geq CS(P_i^d)$ holds for every κ. By continuity, $B \geq CS(P_i^d)$ holds if Δ is sufficiently small (irrespective of κ) and when Δ is sufficiently large it holds if and only if κ is sufficiently small.

Given the above, and since B is strictly decreasing in Δ and in κ, there exists a $\widehat{k} \in (0,1)$ such that $B \geq CS(P_i^d)$ is satisfied for all Δ if $\kappa \leq \widehat{k}$, and for all $\Delta < \widehat{\Delta}(\kappa)$ when $\kappa \in (\widehat{\kappa},1)$ ($\widehat{\Delta}(\cdot)$ being strictly positive and decreasing in its range). We next prove the monotonicity results.

$$\begin{aligned}\frac{\partial B}{\partial \Delta} < 0 &\Leftrightarrow (r\kappa^2 CS(P^m)e^{-r\Delta})(1-\kappa^2 e^{-r\Delta}) > \\ &\kappa^2 r e^{-r\Delta}[(1-\kappa^2)(\gamma CS(P^c)+(1-\gamma)CS(P_i^d))+\kappa^2 CS(P^m)(1-e^{-r\Delta})] \\ &\Leftrightarrow CS(P^m) < \gamma CS(P^c)+(1-\gamma)CS(P_i^d)\end{aligned}$$

Consumer surplus is increasing in additional entry, so the last inequality holds.

$$\begin{aligned}\frac{\partial B}{\partial \kappa} &= \frac{(1-\kappa^2 e^{-r\Delta})[2(\kappa-1)CS(P^c)+(2-4\kappa)CS(P_i^d)+2\kappa(1-e^{-r\Delta})CS(P^m)]}{(1-\kappa^2 e^{-r\Delta})^2} \\ &+ \frac{2\kappa e^{-r\Delta}[(1-\kappa)^2 CS(P^c)+2\kappa(1-\kappa)CS(P_i^d)+\kappa^2 CS(P^m)(1-e^{-r\Delta})]}{(1-\kappa^2 e^{-r\Delta})^2}.\end{aligned}$$

Moreover

$$\frac{\partial B}{\partial \kappa \partial CS(P^m)} = \frac{2\kappa\left(1-e^{-r\Delta}\right)}{(1-\kappa^2 e^{-r\Delta})^2} > 0$$

Replace $CS(P^m)$ by $CS(P_i^d) > CS(P^m)$ in the derivative above, to obtain

$$\frac{\partial B}{\partial \kappa} < \frac{2\left(1-\kappa\right)\left(1-\kappa e^{-r\Delta}\right)\left(CS(P_i^d)-CS(P^c)\right)}{(1-\kappa^2 e^{-r\Delta})^2} < 0.$$

The last inequality holds since $CS(P_i^d) - CS(P^c) < 0$.

Proof Proposition 3. The expected welfare in the absence of an agreement is a weighted average between the welfare under monopoly and

$$\gamma(\Pi_I(P^c)+CS(P^c))+(1-\gamma)(\Pi_I(P_i^d)+CS(P_i^d)),$$

where the latter is the expected welfare in the limiting case where $\Delta \to 0$. The weight put on the monopoly outcome ranges from zero when $\Delta \to 0$ to κ^2 as $\Delta \to \infty$. Thus, if both previously mentioned welfares are smaller than the welfare under agreement, which is invariant in Δ and given by

$$\Pi_I(P_i^d) + (1-\kappa)\Pi_i(P_i^d) + CS(P_i^d),$$

then an agreement is preferred to no agreement for all Δ. Thus either such an agreement is not reached, or it increases welfare when reached. It follows that allowing for early entry agreements weakly increases welfare.

Proof Proposition 6. By focusing exclusively on its captive consumers, the incumbent can always guarantee itself a payoff of $\frac{1}{4}\mu$. Therefore it will never charge a price that would give him a lower payoff when serving the whole market at that price. This establishes the existence of a lowest critical $\underline{p}$, such that all lower prices are dominated, and which solves

$$\frac{1}{4}\mu = \underline{p}(1-\underline{p}) \Leftrightarrow \underline{p} = \frac{1}{2}(1-\sqrt{1-\mu}).$$

For this reason, the entrant is guaranteed a payoff of at least

$$(1-\mu)\underline{p}(1-\underline{p}).$$

Let G_i denote the cumulative distribution function (CDF) of prices of the entrant. If the incumbent charges some price $p > \underline{p}$, then with probability $G_i(p)$ its price exceeds the price of the entrant and it sells only to its captive segment and generates a revenue of $\mu p(1-p)$, while with probability $1 - G_i(p)$ its price is lower than the entrant's and therefore the incumbent captures both segments to generate a revenue of $p(1-p)$. The expected revenue is then

$$(1-G_i(p))p(1-p) + G_i(p)\mu p(1-p) = (1-G_i(p)(1-\mu))p(1-p).$$

(We have avoided the case of ties, as it is never optimal to set a price at which a rival has a mass point.) Likewise, let G_I denote the CDF of prices of the incumbent. If the entrant charges some price $p > \underline{p}$, then with probability $G_I(p)$ its price exceeds the price of the incumbent and it sells nothing, while with probability $1 - G_I(p)$ its price is lower than the entrant's price and therefore the entrant captures the shoppers' segment to generates a revenue of $(1-\mu)p(1-p)$. The expected revenue is then

$$(1-G_I(p))(1-\mu)p(1-p).$$

Since $p(1-p)$ peaks at $p = p^m$, neither firm uses a price $p > p^m$ in its strategy support. It follows that the lowest price in the strategy support of both firms must be the same, as otherwise the firm with the lowest price could increase its profit by increasing its price.

The price $\underline{p}$ is the lowest price in the strategy support of both firms: Suppose not, then the revenue of the entrant must exceed $(1-\mu))\underline{p}(1-\underline{p})$ and that of the incumbent must exceed $\underline{p}(1-\underline{p}) = \frac{1}{4}\mu$. The highest price in the strategy support of the entrant cannot exceed the highest price in the support of the incumbent (since it would then make zero with that price), and vice versa (since the incumbent would then make at most $\frac{1}{4}\mu$, which is less than the revenue in the conjectured equilibrium). Finally, if that highest price is the same for both firms, then at that price the incumbent only sells to its captive segment and its revenue is at most $\frac{1}{4}\mu$, again less than the its revenue in the conjectured equilibrium.

Using the fact that $\underline{p}$ is the lowest price in the strategy support of both firms, standard arguments show that there must be a common support of prices, with no gaps or mass points in its interior. We then obtain the CDF $G_i(p)$ by solving the indifference condition

$$\frac{1}{4}\mu = (1 - G_i(p))p(1-p) + G_i(p)\mu p(1-p) \Leftrightarrow G_i(p) = \frac{1}{1-\mu}(1 - \frac{\mu}{4p(1-p)}),$$

and the CDF $G_I(p)$ by solving the indifference condition

$$(1-\mu)\underline{p}(1-\underline{p}) = (1 - G_I(p))(1-\mu)p(1-p) \Leftrightarrow G_I(p) = 1 - \frac{\mu}{4p(1-p)}.$$

Both CDFs reach a maximum at p^m. Moreover $G_i(p^m) = 1$ and $G_I(p^m) = 1 - \mu$. The entrant thus exhausts its CDF at p^m, and the incumbent places a mass point with the remaining density μ at p^m.

To summarize, in the unique duopoly equilibrium: Both firms compete for the shopper segment and randomize their prices over a common support $\left[\frac{1}{2}(1-\sqrt{1-\mu}), p^m\right]$, using the respective CDFs $G_i(p) = \frac{G_I(p)}{1-\mu}$ and $G_I(p) = 1 - \frac{\mu}{4p(1-p)}$ and with the remainder probability μ the incumbent decides to instead exclusively serve the captive segment at the monopoly price. The *expected* revenues are

$$\pi_I(P_1^d) = \frac{1}{4}\mu \text{ and } \pi_1(P_1^d) = \frac{1}{4}(1-\mu)\mu.$$

Proof Lemma 5. The consumer surplus in monopoly is $cs(P^m) = \frac{1}{8}$. The consumer surplus in competition is $cs(P^c) = \mu\frac{1}{8} + (1-\mu)\frac{1}{2}$. Thus $cs(P^c) > cs(P^m)$ for all $\mu \in (0,1)$.

Under duopoly we must compute the surplus of each group. The surplus of captive the consumer segment (normalized by its size) is given by

$$\int_{\underline{p}}^{p^m} g_I(p)\frac{(1-p)^2}{2}dp + \frac{\mu}{8},$$

where $g_i(p) = G_i'(p)$ and the consumer surplus when facing a price of p is $\frac{(1-p)^2}{2}$. Next we consider the shopper segment. The probability that p is the lowest price is

$$g_I(p)(1 - G_i(p)) + g_i(p)(1 - G_I(p)) = \frac{g_I(p)(2(1 - G_I(p)) - \mu)}{1-\mu}.$$

Thus, the expected consumer surplus in the shopper segment (normalized by its size) is

$$\int_{\underline{p}}^{p^m} \frac{g_I(p)(2(1 - G_I(p)) - \mu)}{1-\mu}\frac{(1-p)^2}{2}dp.$$

Therefore the consumer surplus under duopoly, $cs(P_i^d)$, is

$$\mu(\int_{\underline{p}}^{p^m} g_I(p)\frac{(1-p)^2}{2}dp + \frac{\mu}{8}) + (1-\mu)\int_{\underline{p}}^{p^m} \frac{g_I(p)(2(1 - G_I(p)) - \mu)}{1-\mu}\frac{(1-p)^2}{2}dp,$$

or

$$\left[\frac{1}{32p^2}\mu^2\left(6p+6p^2\ln p+2p^2\ln(p-1)-1\right)\right]_{p=\frac{1}{2}}$$
$$-\left[\frac{1}{32p^2}\mu^2\left(6p+6p^2\ln p+2p^2\ln(p-1)-1\right)\right]_{p=\frac{1}{2}(1-\sqrt{1-\mu})}$$
$$+\frac{1}{8}\mu\left(2\sqrt{1-\mu}-\mu+2\mu\ln 2+2\mu\ln\left(\frac{1}{2}-\frac{1}{2}\sqrt{1-\mu}\right)+2\right).$$

$cs(P_i^d)$ is decreasing in μ, and converges to $cs(P^m)$ as $\mu \to 1$. $cs(P_i^d) > cs(P^m)$. Algebric computations show that $cs(P^c) > cs(P_i^d)$ for all $\mu \in (0,1)$.

Proof Proposition 7. Reorganizing the condition in the main text, in limiting case $\Delta \to 0$, an agreement reduces welfare if and only if

$$\frac{\Pi_i(P_i^d)}{CS(P^c)-CS(P_i^d)} < \frac{1-\kappa}{1-\kappa^2}.$$

The left hand side is strictly decreasing in μ, is equal to 1 for $\mu^* \simeq 0.76$ and converges to $\frac{2}{3}$ as $\mu \to 1$. The right hand side is decreasing in k, is equal to 1 for $\kappa = 0$, and converges to 0.5 as $\kappa \to 1$. Thus the agreement increases welfare for all κ if $\mu \leq \mu^*$ and for every $\mu > \mu^*$ there exists a unique $\kappa(\mu) \in (0,1)$ such that the agreement increases welfare if and only if $\kappa \geq \kappa(\mu)$ (and $\kappa(\cdot)$ is continuous and decreasing). The last statement follows from Proposition 3, and the fact that single entry is welfare increasing when profitable.

References

S. Anderson, A. Baik, and N. Larson. Personalized pricing and advertising: An asymmetric equilibrium analysis. *Games and Economic Behavior*, 92:53 – 73, 2015.

M. R. Baye, D. Kovenock, and C. G. de Vries. It takes two to tango: Equilibria in a model of sales. *Games and Economic Behavior*, 4(4):493 – 510, 1992.

M. R. Baye, J. Morgan, and P. Scholten. Information, Search, and Price Dispersion. Working Papers 2006-11, Indiana University, Kelley School of Business, Department of Business Economics and Public Policy, Dec. 2006.

F. A. Bokhari. What Is the price of pay-to-delay deals. *Journal of Competition Law & Economics*, 9(3):739–753, 06 2013.

F. A. Bokhari, F. Mariuzzo, and A. Polanski. Entry limiting agreements for pharmaceuticals: pay-to-delay and authorized generic deals. Working paper, 2017. URL https://ssrn.com/abstract=2626508. Accessed: October 7, 2019.

A. Branthwaite and P. Cooper. Analgesic Effects Of Branding In Treatment Of Headaches. *British Medical Journal (Clinical Research Edition)*, 282(6276):1576–1578, 1981.

B. J. Bronnenberg, J.-P. Dubé, M. Gentzkow, and J. M. Shapiro. Do Pharmacists Buy Bayer? Informed Shoppers and the Brand Premium. *The Quarterly Journal of Economics*, 130(4):1669–1726, 07 2015.

J. Bulow. *The Gaming of Pharmaceutical Patents*, pages 145–187. The MIT Press, February 2004.

J. Choi. Patent Litigation as an Information-Transmission Mechanism. *American Economic Review*, 88(5):1249–63, 1998.

T. F. Cotter. Antitrust implications of patent settlements involving reverse payments: defending a rebuttable presumption of illegality in light of some recent scholarship. *Antitrust Law Journal*, 71(3):1069–1097, 2004.

D. A. Crane. Exit Payments in Settlement of Patent Infringement Lawsuits: Antitrust Rules and Economic Implications. *Florida Law Review*, 54:747, 2002.

J. P. Davis. Applying Litigation Economics to Patent Settlements: Why Reverse Payments Should Be Per Se Illegal Symposium: Antitrust Enforcement in the Pharmaceutical Industry. *Rutgers Law Journal*, 41:255, 2009.

S. Davis, K. Murphy, and R. Topel. Entry, Pricing, and Product Design in an Initially Monopolized Market. *Journal of Political Economy*, 112(S1):S188–S225, 2004.

B. Dickey, J. Orszag, and L. Tyson. An Economic Assessment of Patent Settlements in the Pharmaceutical Industry. *Annals Health Law*, 19:367, 2010.

A. K. Dixit and C. Shapiro. *Entry Dynamics with Mixed Strategies.* L.G. Thomas, ed., Strategic Planning, Lexington: Lexington Books, 1986.

S. Donnenfeld and S. Weber. Vertical product differentiation with entry. *International Journal of Industrial Organization*, 10(3):449–472, 1992.

S. Donnenfeld and S. Weber. Limit Qualities and Entry Deterrence. *RAND Journal of Economics*, 26(1):113–130, 1995.

A. Duchêne, D. Sen, and K. Serfes. Patent Licensing and Entry Deterrence: The Role of Low Royalties. *Economica*, 82:1324–1348, December 2015.

EC. European Commission Pharmaceutical Sector Inquiry Final Report Part 1, July 2009. URL http://ec.europa.eu/competition/sectors/pharmaceuticals/inquiry/staff_working_paper_part1.pdf. Accessed: October 7, 2019.

W. Elberfeld and E. Wolfstetter. A dynamic model of Bertrand competition with entry. *International Journal of Industrial Organization*, 17(4):513 – 525, 1999.

M. Eswaran. Licensees as Entry Barriers. *The Canadian Journal of Economics / Revue canadienne d'Economique*, 27(3):673–688, 1994.

J. Farrell. Cheap Talk, Coordination, and Entry. *The RAND Journal of Economics*, 18(1): 34–39, 1987.

D. Fudenberg and J. Tirole. *Game Theory*. MIT Press, 1991.

J. Gabszewicz and J. Thisse. Price competition, quality and income disparities. *Journal of Economic Theory*, 20(3):340–359, 1979.

S. P. Gallasch. *The Anticompetitive Misuse of Intellectual Property Rights in the European Pharmaceutical Sector*. PhD thesis, University of East Anglia, 2014.

N. Gallini. Deterrence by Market Sharing: A Strategic Incentive for Licensing. *American Economic Review*, 74(5):931–41, 1984.

L. Gratz. Economic Analysis of Pay-for-delay Settlements and Their Legal Ruling. Munich Discussion Paper 2012-6, Ludwig-Maximilians-Universität München, Volkswirtschaftliche Fakultät, 2012.

B. H. Hall, Z. Griliches, and J. A. Hausman. Patents and R and D: Is There a Lag? *International Economic Review*, 27(2):265–283, 1986.

J. Harrington. Collusion among asymmetric firms: The case of different discount factors. *International Journal of Industrial Organization*, 7(2):289–307, 1989.

J. C. Harsanyi. Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points. *International Journal of Game Theory*, 2(1), Dec 1973.

H. Hovenkamp, M. Janis, and M. Lemley. Anticompetitive Settlement of Intellectual Property Disputes. *Minnesota Law Review*, 87, 02 2003.

N. M. Hung and N. Schmitt. Quality competition and threat of entry in duopoly. *Economics Letters*, 27(3):287–292, 1988.

T. Kaplan and D. Wettstein. The Possibility of Mixed-Strategy Equilibria with Constant-Returns-to-Scale Technology under Bertrand Competition. *Spanish Economic Review*, 2: 65–71, 04 2000.

P. Klemperer. Why Every Economist Should Learn Some Auction Theory Advances in Economics and Econometrics. Invited Lectures to 8th World Congress of the Econometric Society, M. Dewatripont, L. Hansen and S. Turnovsky (eds.), Cambridge University Press, 2003.

M. Kyle and A. Marxen. Entry agreements and generic drug competition in Europe. Working Paper, 2019.

S. Lutz. Vertical product differentiation and entry deterrence. *Journal of Economics*, 65 (1):79–102, 1997.

N. Mankiw and M. Whinston. Free Entry and Social Inefficiency. *Rand Journal of Economics*, 17(Spring):48–58, 1986.

J. Montez and N. Schutz. All-Pay Oligopolies: Price Competition with Unobservable Inventory Choices. CEPR Discussion Paper No. DP12963, 2018.

M. Mussa and S. Rosen. Monopoly and product quality. *Journal of Economic Theory*, 18 (2):301–317, August 1978.

Y.-H. Noh and G. Moschini. Vertical Product Differentiation, Entry-Deterrence Strategies, and Entry Qualities. *Review of Industrial Organization*, 29(3):227–252, 2006.

A. Pakes. On Patents, R&D, and the Stock Market Rate of Return. *Journal of Political Economy*, 93(2):390–409, April 1985.

E. Palikot and M. Pietola. Pay-for-Delay with Settlement Externalities. Working paper, 2019. URL https://emilpalikot.files.wordpress.com/2019/09/patents2019_manuscript.pdf. Accessed: October 7, 2019.

M. Peitz. The pro-competitive effect of higher entry costs. *International Journal of Industrial Organization*, 20(3):353–364, March 2002.

M. Perry. Vertical equilibrium in a competitive input market. *International Journal of Industrial Organization*, 2(2):159–170, 1 1984.

K. E. Rockett. Choosing the Competition and Patent Licensing. *The RAND Journal of Economics*, 21(1):161–171, 1990.

S. Salop and J. Stiglitz. Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion. *The Review of Economic Studies*, 44(3):493–510, 1977.

S. C. Salop. Question: What Is the Real and Proper Antitrust Welfare Standard – Answer: The True Consumer Welfare Standard. *Loyola Consumer Law Review*, 22:336, 2009-2010.

M. G. Schildkraut. Patent-splitting Settlement and the Reverse Payment Fallacy. *Antitrust Law Journal*, 71(3):1033–1068, 2004.

R. Schlatter. Generikamarkt: Millionen für einen Vorsprung. in: Handelszeitung, July 2005. URL http://www.handelszeitung.ch/unternehmen/generikamarkt-millionen-fuer-einen-vorsprung. Accessed: October 7, 2019.

A. Shaked and J. Sutton. Relaxing Price Competition Through Product Differentiation. *The Review of Economic Studies*, 49(1):3–13, 1982.

C. Shapiro. Antitrust Limits to Patent Settlements. *The RAND Journal of Economics*, 34 (2):391–411, 2003.

S. Shelegia and C. M. Wilson. A Generalized Model of Sales. Working papers, Department of Economics and Business, Universitat Pompeu Fabra, 2016.

Y. Shilony. Mixed Pricing in Oligopoly. *Journal of Economic Theory*, 14(2):373 – 388, 1977.

B. Shiv, Z. Carmon, and D. Ariely. Placebo Effects of Marketing Actions: Consumers May Get What They Pay For. *Journal of Marketing Research*, 42(4):383–393, 2005.

A. Spence. Product Differentiation and Welfare. *American Economic Review*, 66(2):407–14, 1976.

M. Struys. Intellectual Property and Competition Law in the Health Sector. AIPPI Conference, Seoul, 2012. URL https://aippi.org/download/seoul12/presentations/Pres_Pharma_1_MStruys.pdf. Accessed: October 7, 2019.

L. Theiler. Generika in der Kritik: Swissmedic findet vermehrt Teile von krebserregenden Stoffen, Nov. 2019. URL https://www.srf.ch/news/wirtschaft/generika-in-der-kritik-swissmedic-findet-vermehrt-teile-von-krebserregenden-stoffen. Accessed: November 4, 2019.

J. Tirole. *The theory of industrial organization.* MIT Press, 1988.

U.S. Food & Drug Administration. Prepared Statement of the Federal Trade Commission Before the United States Senate Committee on the Judiciary Subcommittee on Antitrust, Competition Policy and Consumer Rights on Pay-for-Delay Deals: Limiting Competition and Costing Consumers, 2013. URL https://www.ftc.gov/sites/default/files/documents/public_statements/prepared-statement-federal-trade-commission-pay-delay-deals-limiting-competition-and-costing/130723payfordelay.pdf. Accessed: October 7, 2019.

H. R. Varian. A Model of Sales. *The American Economic Review*, 70(4):651–659, 1980.

R. Waber, B. Shiv, Z. Carmon, and D. Ariely. Commercial features of placebo and therapeutic efficacy. *JAMA*, 299(9):1016–1017, 2008.

X. Wauthy. Quality Choice in Models of Vertical Differentiation. *The Journal of Industrial Economics*, pages 345–353, 1996.

R. D. Willig and J. P. Bigelow. Antitrust Policy toward Agreements That Settle Patent Litigation Symposium: Legal, Economic, and Policy Reviews of FTC Nonmerger Antitrust Enforcement. *Antitrust Bulletin*, 49:655, 2004.

S.-S. Yi. Whom to License Patented Technology. *Managerial and Decision Economics*, 19 (3):189–195, 1998.