Pay-for-Delay

BY FIONA SCOTT MORTON[1]

This article lays out the economics of competition between branded and generic pharmaceuticals and its welfare consequences. I explain the logic behind so-called "pay-for-delay" or "reverse payments" in the context of the current IP environment where weak (probabilistic) patents are frequently granted by the PTO. The article goes on to relate the Supreme Court decision in Activis to these concepts. I argue that the "scope of a patent" is closely related to its probability of being valid. The Supreme Court dissenting opinion states that IP owners should be allowed to operate within the scope of the patent. For a very weak patent, that might be a very limited scope and bring the dissent into agreement with the majority opinion that a weak patent owner should not be allowed to create market power where the patent did not grant it. However, the dissenting opinion closes with a rejection of using the concept of probabilistic patents in legal analysis.

I. ROLE OF THE PHARMACEUTICAL BRAND

The brand receives patent protection in order to incentivize innovation. As is well known, and does not bear covering in detail here, society developed the patent system in order to give innovators property rights that create incentives for innovation. This initial period of patent protection is likely to be important for encouraging innovation in pharmaceutical treatments. There is considerable economics literature that asks whether the patent system is a net drain or stimulant to innovation, but it is widely believed that the pharmaceutical industry is one of the most positive cases.

A successful brand has invested in clinical trials and other activities to obtain regulatory approval. In addition to basic research and early clinical trials, the regulator in the US or Europe requires extensive clinical trials to determine safety and efficacy. These often involve thousands of patients and take many years. The financial investment in bringing a new pharmaceutical product to market is therefore considerable (recent US estimates approach $1 billion).

At the end of the approval process the innovator typically has a patent on the original molecule as well as additional patents on other aspects of the brand, such a pill shape or extended release formulation. The innovator also may have market exclusivity, which in the US is a guaranteed minimum number of years before generics may enter. Lastly, the innovator may have data exclusivity on the results of its clinical trials for a period of time. Data exclusivity means that no other firm can use the trial data showing the effectiveness of the drug during the time of exclusivity.

II. SOCIETY WANTS COMPETITION FOR THE BRAND ONCE THE PATENT EXPIRES

The idea of a patent is to reward the innovator with a limited-time exclusivity on its invention after which the innovation is available without cost to society. A generic entrant that neither has to pay the fixed cost of inventing the drug, nor the fixed cost of determining that it works in real people, faces only the low marginal cost

of manufacture. Such a generic can cover its costs by selling the drug for a price close to this marginal cost, and in that way benefit society by greatly expanding access to the product, and, in many modern economies, reducing the tax burden needed to finance the public healthcare system. Generic entry into pharmaceutical markets where patents have expired is therefore desirable (assuming optimal patent life).

In the US, legislation explicitly designed to achieve this goal, the Hatch Waxman Act (HW), was passed in 1984. It established an exclusivity period for innovators, thereby guaranteeing the brand a minimum period when it could sell without facing competition no matter how slow the FDA approval process. It also reduced the fixed cost of entry for generics by requiring only that they demonstrate a product is bioequivalent to (almost the same as) the brand, rather than running long safety and efficacy trials. This shorter procedure was called an 'Abbreviated' New Drug Application, or ANDA. Prior to HW, if a generic challenged a brand's patent and the patent was found invalid, the market would immediately be open to all generic entrants. Because competition is fierce in the generic industry, profits would be low and provide no return on the cost of the initial litigation for the challenger.

HW fixed this problem by providing that the first generic to defeat a brand patent be given six months of generic exclusivity. These generic entrants are known as "Paragraph IV" entrants because upon entry, they certify to the FDA that they are infringing a brand's patent but assert that the patent is not valid or not infringed. The first paragraph IV filer that succeeds in proving its case gains the right to be the only generic on the market for six months (the FDA will not approve others), during which time it competes only with the brand. These six months are lucrative, and therefore provide the financial incentive to litigate a weak brand patent. After six months, other generics may enter and society gets the benefit of the fact that the patent has been shown to be invalid. The critical feature of the legislation is that should the initial litigant not enter, the six month period does not begin and no other generic can enter, either. Thus without the first generic entrant, the brand faces no generic competition at all. Moreover, only one Paragraph IV exclusivity period is granted under the law; no subsequent ANDA-filer may earn it, and therefore no subsequent generic entrant has a strong financial incentive to litigate. Settling with the first Paragraph IV generic entry therefore strongly reduces the likelihood of subsequent generic entry until all the brand's patents have expired.

THE CRITICAL FEATURE OF THE LEGISLATION IS THAT SHOULD THE INITIAL LITIGANT NOT ENTER, THE SIX MONTH PERIOD DOES NOT BEGIN AND NO OTHER GENERIC CAN ENTER, EITHER. THUS WITHOUT THE FIRST GENERIC ENTRANT, THE BRAND FACES NO GENERIC COMPETITION AT ALL. MOREOVER, ONLY ONE PARAGRAPH IV EXCLUSIVITY PERIOD IS GRANTED UNDER THE LAW; NO SUBSEQUENT ANDA-FILER MAY EARN IT, AND THEREFORE NO SUBSEQUENT GENERIC ENTRANT HAS A STRONG FINANCIAL INCENTIVE TO LITIGATE. SETTLING WITH THE FIRST PARAGRAPH IV GENERIC ENTRY THEREFORE STRONGLY REDUCES THE LIKELIHOOD OF SUBSEQUENT GENERIC ENTRY UNTIL ALL THE BRAND'S PATENTS HAVE EXPIRED.

After HW took effect the number of generic entrants in US pharmaceutical markets increased, generic entry immediately after patent expiration became frequent (often measured in days), and generic penetration

grew steadily to the point where generics now routinely capture more than 90 percent of all prescriptions for a given molecule. Research shows that generic prices decline with the number of entrants, typically beginning at about 80 percent of the branded price and falling to as little as 10 percent of the brand's price in large markets with a dozen or more entrants. The statistics in Europe are a little less favorable (the generic often being the local firm, and pharmacies often successfully lobbying for high margins). However, generic entry after patent expiration in Europe is similarly beneficial for consumers and healthcare budgets.

III. TEMPTATION

The cost structure of the pharmaceutical industry is very important in shaping legal strategies. As noted above, the brand's price is typically five-to-ten times the generic price. The brand's gross margin is therefore very high, often around 90 percent. Generic entry takes sales away from the brand, often in a dramatic and significant manner. The generic, due to competing with a homogeneous product in an industry with low barriers to entry, does not earn large economic profits. Indeed, price is often close to marginal cost and therefore the generic saves the consumer much more in consumer surplus than the generic collects in profits. To fix ideas, imagine the brand's price is 100, the generic price is 20, and manufacturing cost is 10. On each unit sold the generic earns only 10 while saving the consumer 80 and costing the brand 90.

Even the first generic into the market does not earn the high economic profits typical of a brand, and thus it is easy for the brand to make the generic an attractive offer while leaving ample profit for itself. The brand offers *X* percent of its remaining brand profit to the successful Paragraph IV generic entrant–*if* it agrees not to enter the market. Since other generics may not enter until the initial paragraph IV entrant has enjoyed its six months of exclusivity, and have little incentive since they will not receive any exclusivity protection, entry is effectively blocked. This leaves 100-*X* percent of profits for the brand, rather than almost zero, which is what it would receive should generics enter. The parties in this way can divide up the monopoly profits in the market. The parties prefer to settle patent litigation in this manner because it preserves the brand's monopoly profit for division between the two firms; the missing element, of course, is that consumer surplus is not considered in the arrangement. The profit accruing to both parties would be shared with consumers to a significant degree if the usual entry process took place.

TO FIX IDEAS, IMAGINE THE BRAND'S PRICE IS 100, THE GENERIC PRICE IS 20, AND MANUFACTURING COST IS 10. ON EACH UNIT SOLD THE GENERIC EARNS ONLY 10 WHILE SAVING THE CONSUMER 80 AND COSTING THE BRAND 90.

The key to the strategy is found in the HW regulation: FDA approval of subsequent generics requires the Paragraph IV generic entrant to sell its product for six months, and therefore the first generic entrant can block subsequent entry by failing to enter the market. In a market without regulated entry, subsequent entrants would face the same costs and rewards as the first 'entrant' -- who chose to stay out. In the classic pay-for-delay scenario, the brand and the generic settle their patent litigation by agreeing that the generic will stay out of the market for a certain number of years while the brand pays the generic a lump sum. If the generic had infringed a valid patent, and owed damages, one might expect settlement to involve a payment in the reverse direction, from the generic to the brand; this is why these settlements are called "reverse payments." The delay in the arrival of generic entry that results from such a settlement is the source of the other common term "pay for delay." Notice also that this tactic allows the brand to leverage what might be quite weak patents into monopoly profit.

IV. PROBABILISTIC PATENTS AND THE SUPREME COURT

What is a weak patent? It is one that has a low probability of being found valid by a court. The notion of "probabilistic patents" was introduced into the law and economics literature by Carl Shapiro in various articles over the last decade, and has significantly changed economic thinking concerning the way intellectual property works.[2] The 2013 Supreme Court pay-for-delay decision (*FTC v Activis*) makes it clear that the important insights in those pieces have only begun to be integrated into legal reasoning.

A patent, when issued by the PTO, may or may not be valid. It has some probability $p<=1$ of being found valid in litigation. The statistics on patent validity in the US unfortunately suggest that p is, on average, quite small. The PTO issues 15,000 patents each month and each application receives 15-to-20 hours of patent examiner time on average.[3] Of patents that are litigated to trial (0.1 percent), approximately half are found invalid.[4] Between 55 and 67 percent of patents are not renewed with user fees, suggesting that their owners do not believe they are worth retaining, perhaps because of invalidity or perhaps because of commercial irrelevance.[5]

The Supreme Court pay-for-delay decision revisits a case where the Federal Trade Commission sued Activis under the theory of harm described above. A district court dismissed the case, and the Eleventh Circuit affirmed, relying on the public policy favoring settlement of disputes and stating that the patent owner was contracting within the scope of the patent.

THE PARTIES PREFER TO SETTLE PATENT LITIGATION IN THIS MANNER BECAUSE IT PRESERVES THE BRAND'S MONOPOLY PROFIT FOR DIVISION BETWEEN THE TWO FIRMS; THE MISSING ELEMENT, OF COURSE, IS THAT CONSUMER SURPLUS IS NOT CONSIDERED IN THE ARRANGEMENT. THE PROFIT ACCRUING TO BOTH PARTIES WOULD BE SHARED WITH CONSUMERS TO A SIGNIFICANT DEGREE IF THE USUAL ENTRY PROCESS TOOK PLACE.

First consider the competitive situation in the event the brand has a known-to-be valid patent. Such a patent would not attract a generic challenge. In the event of a challenge, the most the brand would be willing to pay to settle would be exactly the cost of its litigation because both the brand and the generic know that its patent is valid. Leaving aside litigation costs, in such a setting a reverse payment will never be used. Thus we see that the Eleventh Circuit's reasoning contains an internal contradiction: if the patent is valid, no reverse payment should exist. Indeed, the presence of a reverse payment itself suggests the patent may not be valid.

The majority opinion of the Supreme Court contains an important subtlety:

> "…and we are willing to take this fact as evidence that the agreement's anticompetitive effects fall within the scope of the exclusionary *potential* of the patent."[6]

The key difference in the conclusions of the two courts, in my view, comes from the interpretation of the word "potential." We know from the empirical evidence that, if litigated, many patents do not have much exclusionary power. Thus the "exclusionary potential" of a patent might be quite small. Implicitly, the reasoning in the majority opinion is that the agreement's anticompetitive effects (the elimination of generic competition) may not fall within the proper scope of a weak patent. The majority observes that the reverse payment allows

the patent holder effectively to buy the remaining validity it has not earned through its own innovation:

> "The payment in effect amounts to a purchase by the patentee of the exclusive right to sell its product, a right it already claims, but would lose if the patent litigation were to continue and the patent were held invalid or not infringed by the generic product…"[7]

The majority is also quite clear on the large competitive harm from this behavior.

> "But settlement on the terms said by the FTC to be at issue here–payment in return for staying out of the market–simply keeps prices at patentee-set levels, potentially producing the full patent-related $500 million monopoly return while dividing that return between the challenged patentee and the patent challenger. The patentee and the challenger gain; the consumer loses."[8]

The dissent by Chief Justice John Roberts relies heavily on the scope of the patent reasoning; it is the main argument he uses to find in favor of Solvay. The dissent's argument assumes that the patent is valid and infringed with $p=1$. Of course, in that environment, excluding a competitor is appropriate. However, we know that patents with $p=1$ are rare, and that the average patent has a significant chance of being found invalid. This fact is never raised in the dissent. Rather, it returns again and again to the issue of the scope of the patent without appreciating its probabilistic nature. However, the arguments the dissent makes are exactly right once the probabilistic nature of patents is taken into account. This small change substantially reconciles the two opinions.

> "The correct approach should therefore be to ask whether the settlement gives Solvay monopoly power beyond what the patent already gave it."[9]

Interestingly, though it is stated as a criticism, this statement is exactly what the majority is doing. The difference is that in the US in the modern era, we cannot assume an unlitigated patent is valid. That is not a position that is empirically correct . The only way to view patents that is consistent with the data is as probabilistic rights.[10] The majority is acting consistently with the premise that the *expected* validity of any given patent is less than one, and that the patent at issue in this case might have a very low probability of being valid. If the patent has a low probability of being valid, the settlement may give the patent owner monopoly power well beyond what the patent provides.

The dissent goes on to discuss the notion of the zone within which the patent holder may operate as being critical to the application of antitrust law.

> "The point of patent law is to grant limited monopolies as a way of encouraging innovation… In doing so it provides an exception to antitrust law, and the scope of the patent–*i.e.*, the rights conferred by the patent – forms the zone within which the patent holder may operate without facing antitrust liability."[11]

> "The key, of course, is that the patent holder–when doing anything, including settling–must act within the scope of the patent. If its actions go beyond the monopoly powers conferred by the patent, we have held that such actions are subject to antitrust scrutiny."[12]

> "…that when a patent holder acts outside the scope of its patent, it is no longer protected from

antitrust scrutiny by the patent."[13]

These quotations from the dissent are exactly right–*and* consistent with the majority opinion. For example, if the patent is valid with $p=.5$, its scope is limited; the rights holder could settle for an entry date that is half the patent life to achieve all the profit to which it is entitled. Settling with the generic to achieve the full patent life would be stepping outside the "zone" of the rights conferred by the patent and, by these arguments, exposes the patentee to antitrust liability. In a world of probabilistic patents, antitrust enforcement requires that the patent holder not leverage his $p<1$ into $p=1$.

However, it is clear that the dissent rejects the idea that patents could be probabilistic and their scope thereby limited. Below, I reproduce part of the majority opinion that Chief Justice Roberts quotes in disagreement, and then the dissent's own conclusion:

> "First, the majority explains that 'the patent here may or may not be valid, and may or may not be infringed' Because there is 'uncertainty' about whether the patent is actually valid, the Court says that any questions regarding the legality of the settlement should be 'measure[ed]' by 'procompetitive antitrust policies' rather than 'patent law policy.'"[14]
>
> "And the scope of the patent – *i.e.* what rights are conferred by the patent – should be determined by reference to patent law."[15]

From this it appears that Chief Justice Roberts thinks that determining legality using "patent law policy" means that the judge or enforcement agency should assume all patents are valid, in contrast to the empirical evidence. Notice that such a policy endows a party that files what could be a completely useless patent with tremendous market power. Indeed, such a policy encourages the filing of trivial patents because there is a chance they can be used to prolong and maintain monopoly profits.

THE ONLY WAY TO VIEW PATENTS THAT IS CONSISTENT WITH THE DATA IS AS PROBABILISTIC RIGHTS.[10] THE MAJORITY IS ACTING CONSISTENTLY WITH THE PREMISE THAT THE *EXPECTED* VALIDITY OF ANY GIVEN PATENT IS LESS THAN ONE, AND THAT THE PATENT AT ISSUE IN THIS CASE MIGHT HAVE A VERY LOW PROBABILITY OF BEING VALID. IF THE PATENT HAS A LOW PROBABILITY OF BEING VALID, THE SETTLEMENT MAY GIVE THE PATENT OWNER MONOPOLY POWER WELL BEYOND WHAT THE PATENT PROVIDES.

Finally, I will address the issue of why such reverse payments have been rare in the past. ("The majority points to no case where a patent settlement was subject to antitrust scrutiny merely because the validity of the patent was uncertain. Not one.")[16] Hatch-Waxman sets up unique incentives that cause reverse payments to be worth making. In a standard market where entry is not regulated by the government, there is no way to stop attempted entry by every possible entrant. Normally, as the majority opinion points out, a rights holder would have to pay off a never-ending line of patent challengers. In this setting a patent holder would not find reverse payments to be efficacious in maintaining its monopoly. By contrast, HW endows that first generic Paragraph IV challenger with the ability to block others from entering. The regulation itself creates only one partner with this power and therefore one party with whom it is

HATCH-WAXMAN SETS UP UNIQUE INCENTIVES THAT CAUSE REVERSE PAYMENTS TO BE WORTH MAKING. IN A STANDARD MARKET WHERE ENTRY IS NOT REGULATED BY THE GOVERNMENT, THERE IS NO WAY TO STOP ATTEMPTED ENTRY BY EVERY POSSIBLE ENTRANT.... BY CONTRAST, HW ENDOWS THAT FIRST GENERIC PARAGRAPH IV CHALLENGER WITH THE ABILITY TO BLOCK OTHERS FROM ENTERING.

worth colluding. It is not surprising that we have not seen many patent owners try this tactic before HW, nor that quite a few have tried it since.

VI. POLICY GOALS

The Supreme Court decision finding that reverse payments should be subject to the rule of reason allows antitrust enforcers to pursue many admirable policy goals. One important policy goal is to create an environment where innovation is rewarded with patent rights. Secondly, it seems desirable to encourage procompetitive and legitimate market interactions among firms, such as JVs, supply agreements, co-marketing, etc., even if the parties are a firm that makes brand and a firm that makes generics. However, policy makers have strong reasons to promote generic entry into a market as soon as intellectual property rights permit. Competition enforcement can help prevent the brand from blocking entry with legal strategies that bolster its market power beyond the value of its innovation.

The FTC proposal to encourage settlement negotiations over entry date, but not money, has good incentive features. The idea is that instead of allowing the brand to compensate the generic to stay off the market, the parties may only bargain over the date the generic may enter. In that setting, the generic wants to enter early (because that is the only way it can earn profits) while the brand wants the generic to enter late (to preserve its own profits). The strength of the patent will be key in determining the final negotiated date that settles the litigation. In this way the monopoly power that results from the settlement is exactly consistent with the patent's scope, and consumers benefit from competition when that scope ends. ▲

REFERENCES:

Farrell, Joseph and Carl Shapiro (2008) "How Strong are Weak Patents" *American Economic Review*: 98:4:1347-1369.

Hemphill, Scott (2006) "Paying for Delay: Pharmaceutical Patent Settlements as a Regulatory Design Problem" *NYU Law Review*, Vol. 81, p1553.

Lemley, Mark and Carl Shapiro (2005) "Probabilistic Patents" *Journal of Economic Perspectives*: 19:2:75-98.

Shapiro, Carl (2003) "Antitrust Limits to Patent Settlements" *RAND Journal of Economics:* 34:2:391-411

1. Professor of Economics, Yale University
2. See, for example, Farrell and Shapiro (2008) Lemley and Shapiro (2005), and Shapiro (2003)
3. Farrell and Shapiro (2008) p 1347
4. Lemley and Shapiro (2005) p75-6
5. Lemley and Shapiro (2005) p80
6. Fed. Trade Comm'n v. Actavis, Inc., 133 S. Ct. 2223, 2230 (2013) (emphasis added).
7. Id. at 2234.
8. Id. at 2234-35.
9. Id. at 2238 (Roberts, C.J., dissenting).
10. Lemley and Shapiro (2005)
11. Id.
12. Id. at 2239.
13. Id. at 2241.
14. Id. at 2239-40.
15. Id. at 2240.
16. Id. at 2242.