Three Essays in Industrial Organization: Pay-for-delay Settlements, Exclusive Contracts, and Price Discrimination

Inaugural-Dissertation
zur Erlangung des Grades
Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München

2011

vorgelegt von
Linda Gratz

Referent:	Prof. Dr. Klaus M. Schmidt
Korreferent:	Prof. Dr. Florian Englmaier
Promotionsabschlussberatung:	16. Mai 2012

Für meine Eltern

Dorothee und Klaus Gratz

Acknowledgements

First and foremost, I would like to thank Klaus M. Schmidt. He has been the best supervisor and mentor one could possibly imagine - offering invaluable guiding advice and inspiring encouragement. His support was truly exceptional.

I am also much indebted to Florian Englmaier, Fabian Herweg, and Markus Reisinger. Florian Englmaier and Fabian Herweg kindly agreed to serve as second and third reviewer on my dissertation committee, and with Florian Englmaier and Markus Reisinger I was writing the second and third chapter of this dissertation. They provided a great source of advice and support and it was always very joyful to work with them.

I was very fortunate to spent one semester as Research Fellow at Columbia University. I thank C. Scott Hemphill for supervising me during this time, giving me guiding advice from a legal perspective. I also thank Annette Kur and Klaus M. Schmidt for supporting me with my applications and the DAAD for financing the research stay.
I very much appreciate insightful discussions with friends, colleagues, and participants at conferences and seminars. Bernhard Ganglmaier and Johannes Maier deserve special mention for their invaluable advices and suggestions.

Likewise I thank numerous colleagues, who became very good friends and a great source of inspiration and motivation. I am especially grateful to Sara Corradi, Filipe Fischmann, Clemens König, Alice Krieger, Sisi Liang, and Carmen Thoma.

More than anything I would like to thank my family and friends for being so wonderful and supporting me all the time. Your constant support remains truly invaluable.

Linda Gratz

Contents

List of Tables

List of Figures

Preface

> *"We must have more competition and less red tape in pharmaceuticals. The sector is too important to the health and finances of Europe's citizens and governments to accept anything less than the best. The inquiry has told us what is wrong with the sector, and now it is time to act. When it comes to generic entry, every week and month of delay costs money to patients and taxpayers. We will not hesitate to apply the antitrust rules where such delays result from anticompetitive practices. The first antitrust investigations are already under way, and regulatory adjustments are expected to follow dealing with a range of problems in the sector."*

With these words, the former EC Competition Commissioner Nelie Kroes presented the Final Report on the Pharmaceutical Sector Inquiry in July 2009.[1] The report shows that originator and generic companies frequently settle patent-related disputes by agreeing that the alleged infringers *receive* payments from the patent holders. The question arises why originator companies holding ironclad patents would accept such unfavorable settlement terms. Through completion of litigation they could exclude their competitors, merely paying litigation costs. Clearly, originator companies only have reasons to settle on such seemingly unfavorable terms when holding *probabilistic* patents, i.e., patents that would be declared invalid or non-infringed by court with positive probability. In that case the expected generic entry date under litigation lies prior to the end of patent exclusivity. This implies that originator and generic companies can delay generic entry compared to litigation by settling out of court. And they have strong incentives to do so since a delay in generic entry increases their joint profits. Large payments from originator to generic companies are then rationalized as a reimbursement for the entry delay and as a share in the surplus that is generated through the entry delay.

[1] The Final Report on the Pharmaceutical Sector Inquiry can be downloaded at: http://ec.europa.eu/competition/sectors/pharmaceuticals/inquiry/communication_en.pdf. For more information on the Sector Inquiry see IP/09/1098 and MEMO/09/321.

As a consequence of the findings within the Sector Inquiry into Pharmaceuticals the European Commission (EC) paid particularly close attention to *pay-for-delay* settlements. On 21^{st} of October 2011 it opened a fourth proceeding to assess whether a contractual arrangement between Johnson & Johnson and Novartis may have had the effect of hindering generic entry, possibly violating competition law and causing significant consumer harm. As indicated by Nelie Kroes, the EC wants to apply the antitrust rules where delays in generic entry result from anticompetitive practices. It basically has two choices for its legal standard: the application of the *rule of reason* under which courts would evaluate pro- against anticompetitive settlement effects in order to decide whether to approve settlements or not, or the application of the *rule of per se illegality* under which courts would rule settlements that comprise value transfers from originator to generic companies per se illegal. It is not clear from the outset which of these two legal standards would be preferable from a consumer welfare perspective. Neither is it clear from the outset whether the application of antitrust rules to pay-for-delay settlements would generally be desirable. US courts have ruled pay-for-delay settlements *per se legal*. They acted on the presumption that patents underlying pay-for-delay settlements are valid since at the time of settlement there has not been a declaratory judgment to the contrary. Based on this presumption it was argued that any anticompetitive effects of settlement agreements are within the exclusionary zone of the patent and thus cannot be redressed by antitrust law.

The first chapter of this dissertation consists of an *Economic Analysis of Pay-for-delay Settlements and Their Legal Ruling*. Within a theoretical framework we compare the welfare effects of the different rules that can be applied toward pay-for-delay settlements with the goal to give appropriate policy recommendations.

We analyze a marketplace for pharmaceuticals, which originator companies have entered with patented products. Because the originator companies' patents are probabilistic, generic companies contemplate market entry with generic products prior to the patents' expiration. In case of generic entry patent disputes are triggered, resulting in either litigations or settlement agreements. When settling, the companies decide on generic entry dates and if permitted on value transfers. Settlements are ruled by courts. We apply the consumer welfare standard to assess which rule is favorable in practice.

We find that the rule of per se legality induces maximal collusion among settling companies and therefore yields the lowest consumer welfare compared to the alternative rules. In this context, collusion means that settling companies delay generic entry compared to the litigation alternative. While under the rule of per se illegality settling companies are entirely prevented from colluding, under the rule of reason they collude to a limited degree when antitrust enforcement is subject to error.

Intuitively, these results speak for the application of the rule of per se illegality, since only under this rule collusion can be prevented. However, as a main result we find that, contrary to intuition, limited collusion can be welfare enhancing as it increases settling companies' profits and thus fosters generic entry. The general trade-off that arises is that the more settling companies collude, the more competition is restrained under each concluded settlement, but the higher is the number of concluded settlements. If generic companies' incentives to enter and challenge probabilistic patents are rather weak, the rule of reason will outperform the rule of per se legality as the cost adhered to collusion, i.e., the cost that under each concluded settlement competition is restrained, will be lower than the benefit from collusion, i.e., the benefit of additional settlement agreements.

The benefit adhered to collusion does not arise under the rule of per se legality. Under this rule consumer welfare does not increase through additional settlements as settling companies collude maximally, meaning they agree to sustain monopoly for the whole patent duration.

We critically question our first result by asking whether there exist alternative incentive devices to foster generic entry that are more efficient than indirectly permitting collusion between originator and generic companies. Such an alternative incentive device could be the provision of an exclusivity right to generic entrants that first challenge originators' patents, as implemented within the Hatch-Waxman Act (HWA) of 1984 in the US. Under the HWA first generic entrants obtain 180 days of marketing exclusivity during which no subsequent generic company may enter. We show that this prominent incentive device, which restricts competition between generic companies, is in fact ineffective in fostering generic entry. It even impairs generic entry. This suggests that the regulation in which first generic entrants obtain an exclusivity right should be abolished.

The second chapter—a joint work with Prof. Dr. Markus Reisinger—asks *Can Naked Exclusion Be Procompetitive?* Often, incumbent upstream firms make use of exclusive contracts with downstream firms. The incumbent upstream firms offer the downstream firms a payment and, in return, the downstream firms commit themselves to purchase exclusively from the incumbent. Clearly, these exclusive contracts may have procompetitive effects and create efficiency gains within the vertical production chain. At the same time, however, they may have anticompetitive effects as potential upstream entrants may recognize that they will not be able to sell to downstream firms in case of entry. Hence, potential entrants may be foreclosed, even if they are more efficient than the incumbent upstream firms. In antitrust cases on exclusive contracts courts therefore balance potential procompetitive effects through efficiency gains against potential anticompetitive effects through entry deterrence or increased wholesale prices in order to decide whether to allow or prohibit the corresponding deals.[2] They presume that exclusive contracts will have anticompetitive or at best neutral effects if no efficiencies are generated.

This presumption is consistent with results of the previous literature. As is well known, *Chicago School* scholars (e.g., Posner, 1976, and Bork, 1978) contend that, given downstream buyers are independent monopolists or final consumers, the effect of exclusive contracts will be neutral if no efficiencies are generated as the incumbent cannot compensate the downstream buyers for signing such deals. Several authors challenge this argument, pointing out instances in which exclusive contracts can nevertheless be profitable and lead to entry deterrence. For example, Rasmusen, Ramseyer, and Wiley (1991) and Segal and Whinston (2000) show that entrants can be foreclosed if they cannot reach minimum efficient scale when selling only to a fraction of buyers. Fumagalli and Motta (2006), Abito and Wright (2008), and Wright (2009) assess that entrants can also be foreclosed when buyers are downstream competitors, in which case exclusive contracts are a device for contracting parties to protect their profits from competition.

In a recent paper, Simpson and Wickelgren (2007) provide an insightful analysis of the welfare effects of exclusive contracts, incorporating the possibility of contract breach. They study the cases in which downstream firms are either independent monopolists or (almost) perfect Bertrand competitors and find that only in the latter case the incumbent is able to profitably induce downstream firms to sign exclusive contracts. In their model, however, signing does not lead to entry deterrence due to the possibility of contract breach. Nevertheless, they conclude that the effects of exclusive contracts are anticompetitive under perfect Bertrand competition, because the entrant only induces a single downstream firm to breach and this firm monopolizes the downstream market, which results in higher final consumer prices. In accordance with the other authors, Simpson and Wickelgren (2007) therefore argue that in the absence of efficiency gains the effect of exclusive contracts is anticompetitive or at best neutral.

In Chapter 2, we extend the analysis by Simpson and Wickelgren (2007) to account for general degrees of downstream competition. As a central result we find that for moderate degrees of downstream competition exclusive contracts can have *procompeti-*

[2] Recent examples are *United States vs. Transitions Optical*, *United States vs. Dentsply*, *Pernod Ricard and Campbell Distillers vs. Bacardi-Martini*, *Langnese-Iglo vs. European Commission*, and *United States vs. Microsoft*.

tive effects, even if no efficiencies are generated. Thus, our result reverses the results of the *Chicago School* and of subsequent theoretical models. Intuitively, when downstream competition is moderate the entrant wants to induce not just a single downstream firm but *both* downstream firms to breach as it can receive sizable profits from both firms. Because breaching downstream firms have to pay expectation damages to the incumbent, they only breach when they can obtain sufficiently large profits. In order to render breaching profitable the entrant must therefore sell its input at a relatively low wholesale price. In particular, it needs to set a lower wholesale price than it would set in the absence of exclusive contracts. As a consequence, final consumer prices fall and total welfare rises.

This result shows that despite the fact that exclusive contracts may be intended as an anticompetitive device, they can have procompetitive effects. Thus, similar to our result in the first chapter, we reach the conclusion that a seemingly anticompetitive practice can have procompetitive effects. This is important for antitrust authorities to consider. When assessing apparently anticompetitive practices a critical view seems to be highly warranted.

In the third chapter on *Price Discrimination and Fairness Concerns*—a joint work with Prof. Dr. Florian Englmaier and Prof. Dr. Markus Reisinger—we change the perspective. While in the first two chapters we take the view of an antitrust authority asking how to rule potentially anticompetitive practices in order to maximize consumer welfare, in the third chapter we take the view of a firm asking how to price discriminate in order to maximize profits.

Price discrimination is an important strategic instrument for firms in many product markets. According to standard theory firms can increase their profits substantially by selling the same good or service to different consumer segments at varying prices. Standard theory, however, does not take into account that consumers might perceive price discrimination as unfair, especially when they have to pay higher prices than other consumers. In reaction to perceived price unfairness consumers may punish firms by reducing their demand or by buying from other firms altogether. The adverse effects on profitability that may arise could offset the gains from market segmentation.

A large literature addresses the issue of fair pricing but focuses on the question of how the profitability of *price increases* is affected by consumers' fairness concerns (e.g., Kahneman, Knetsch, and Thaler, 1986a,b). Within this literature it is argued that consumers compare the payoffs of firms with their own payoffs. Accordingly, if price increases are not justified by increased costs and lead to an increase in firms' reference payoffs, consumers will react unfavorably.

Surprisingly little research has been devoted to the question how the profitability of *third degree price discrimination* is affected by consumers' fairness concerns. In Chapter 3, we analyze this question within a laboratory experiment and provide a theoretical explanation for the results, which is based on a framework developed by Falk and Fischbacher (2006).

Besides the optimal price discriminating tariff, we are interested in whether the provided contextual information matters. In particular, we want to find out whether firms obtain higher profits when charging poorer consumers lower prices compared to when the wealth of the different consumer groups is unknown. There is strong indication that consumers perceive price discrimination as less unfair when it is justified by income differences. For instance, consumers seem to object student discounts at cinemas less than price discrimination on the internet based on consumers' purchasing history or search behavior.[3]

Our experimental results show that the profitability of third degree price discrimination is negatively affected by consumers' fairness concerns. The higher the price differential that firms charge, the stronger are negative reactions by disadvantaged consumers compared to positive reactions by advantaged consumers. As a consequence, firms obtain higher profits by charging a weaker price differential than the one predicted to be optimal under standard theory.[4]

Moreover, we find that price discriminating firms obtain higher profits when they inform consumers that those consumers who are charged a lower price also have a lower income. This is because the disadvantaged consumers react less negative when they know that they have a higher income and the advantaged consumers react less positive when they know that they have a lower income. Overall, the negative reactions attenuate compared to the positive reactions. Related to practice, this means that firms can increase profitability by charging lower prices to consumers who are generally regarded to be poorer, e.g., to students or the elderly.

We explain our experimental results within a theoretical framework that builds upon concepts developed in an extensive literature on *social preferences* of economic agents. In this literature, Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) argue that agents evaluate the fairness of an action by whether payoffs are equitable or not. In their models agents reciprocate to reduce inequity. Rabin (1993) and Dufwenberg and Kirchsteiger (2004) argue that agents are mainly concerned with procedural

[3]According to a study by Huang, Chang, and Chen (2005) 80,2% of adult users strongly object price discrimination on the internet.

[4]Under standard theory consumers do not exhibit fairness preferences, that is, they solely maximize their material payoff.

justice and evaluate the fairness of an action by its underlying intention. Falk and Fischbacher (2006) incorporate both arguments within a concept of kindness, emphasizing that agents might be concerned with outcomes as well as with procedural justice.

In contrast to the previous models, our model involves not two but three players, a firm and two consumers. A consumer judges the *intention* behind a firm's pricing decision by its *outcome*, which is the difference between the material payoff she can obtain by purchasing from the firm and the material payoff she believes the *other consumer* can obtain. Consequently, consumers regard price discrimination as unfair when they are charged higher prices than other consumers. Their perception of price unfairness intensifies when the price difference gets larger, and it diminishes when they know that they have a higher income. The model stipulates that in reaction to perceived price unfairness consumers punish firms by reducing their demand. Accordingly, consumers who are charged lower prices and regard firms' pricing decisions as fair reward firms by increasing their demand. To the extent that negative consumer reactions are stronger than positive ones, the model predicts that the profitability of third degree price discrimination will be adversely affected when the price differential increases, and the adverse effect will be weaker when consumers know that those consumers with the higher income are charged higher prices.

Intuitively, when those consumers who are charged higher prices know that they also have a higher income, they perceive a higher price as less unfair and therefore punish less. On the other hand, consumers who are charged lower prices feel entitled to a lower price when they know that they have a lower income and thus reward less. Since negative consumer reactions to price discrimination are stronger than positive ones, the positive acceptance effect on the side of the disadvantaged consumers has stronger profit implications than the negative entitlement effect on the side of the advantaged consumers.

Each of the following chapters is a self-contained paper with an own introduction and appendix. Hence, each chapter can be read independently of the other two. A joint bibliography of all papers can be found at the end of the dissertation.

Chapter 1

Economic Analysis of Pay-for-delay Settlements and Their Legal Ruling

1.1 Introduction

The generic company Barr Pharmaceuticals recently received $398.1 million from the originator company Bayer AG for giving up an invalidity claim and halting the production of a generic version of Bayer's antibiotic *Cipro* until the end of patent exclusivity. The question arises why originator companies would want to settle on such unfavorable terms. If they held ironclad patents, they could exclude their competitors through litigation, merely paying litigation costs. There is strong indication that in cases like this, in which the patent holders make large payments to the alleged infringers, the patents at issue are in fact not ironclad but *probabilistic*, meaning with positive probability the patents would be declared invalid or non-infringed by court. In these cases, the expected generic entry date under litigation lies prior to the end of patent exclusivity, which implies that originator and generic companies can use settlements to delay generic entry compared to litigation. Large payments from originator to generic companies are then rationalized as a reimbursement for the entry delay and as a share in the surplus generated through the entry delay.

So, originator and generic companies potentially restrict competition through *pay-for-delay* settlements, delaying generic entry compared to the litigation alternative.[1] Under antitrust law, agreements that restrict competition are *per se illegal* as they lead to static inefficiency. The present legal standard in the US, however, stipulates that pay-for-delay settlements are *per se legal*. Since the originator companies still

[1] For a survey on pay-for-delay settlements see Hemphill (2009).

hold valid patents at the time of settlement, it has been argued that the application of antitrust rules would be inadmissible. In Bayer vs. Barr, for instance, the District Court reasoned that any anticompetitive effects were "within the exclusionary zone of the patent", and thus could not be redressed by antitrust law. So currently, US courts act on the presumption that granted patents are ironclad as long as there has not been a declaratory judgment to the contrary. In light of the probabilistic nature of the patents at issue the question arises, whether this is desirable. *Should courts continue to rule pay-for-delay settlements per se legal or apply an alternative rule, either the rule of per se illegality or the rule of reason?* Under both these alternative rules antitrust law would be applied. Under the rule of per se illegality courts would rule settlements that comprise value transfers from originator to generic companies per se illegal. In contrast, under the rule of reason courts would inquire into the market conditions more comprehensively, balancing pro- against anticompetitive settlement effects. The question, which rule toward pay-for-delay settlements yields the highest consumer welfare, is important also in light of the fact that in the EU regulation of pay-for-delay settlements is still in its infancy.[2] We try to answer this question within a theoretical analysis.

In our framework, originator companies have entered a marketplace for pharmaceuticals with patented products. Because the patents are probabilistic, generic companies contemplate market entry with generic products prior to the patents' expiration. In case of generic entry patent disputes are triggered, resulting in either litigations or settlement agreements. Both parties view settlements superior to litigations as the settlement profits they can obtain are at least as high as the expected litigation profits. When settling, the companies decide on generic entry dates and if permitted on value transfers. Settlements are ruled by courts. We apply the consumer welfare standard to assess which rule is favorable in practice.

Since under the rule of per se legality value transfers within settlements are legal, companies can maximize their joint profits by colluding maximally. That is, they can maximize their joint profits by delaying generic entry until the end of patent duration (as in the case of Bayer vs. Barr). Under the rule of per se illegality value transfers from originator to generic companies within settlements are illegal, so that originator companies cannot compensate generic companies for a delay in entry compared to litigation. Thus, settling companies agree upon entry terms that would in expectation result under

[2]Since 2008 the European Commission (EC) publishes monitoring reports on pay-for-delay settlements (see EC 2008, 2009, 2010). Since 2009, it opened first formal antitrust investigations in a number of pay-for-delay settlements for suspected breaches of Articles 101 and 102 of the Treaty on the Functioning of the European Union (TFEU) (see MEMO/09/322, IP/10/8, IP/11/511 and IP/11/1228). Generally, it advocates restrictions on pay-for-delay settlements.

litigation and are entirely prevented from colluding. Under the rule of reason settling companies are allowed to transfer values but courts prohibit those settlements that they regard anticompetitive. We take into consideration that under the rule of reason courts might make errors when evaluating settlements. They might approve anticompetitive settlements and prohibit procompetitive settlements. Our analysis reveals that imprecise evaluations induce settling companies to collude. The reason is that the likelihood that anticompetitive settlements get approved increases when courts' evaluations become less precise. It then pays more for the companies to choose particularly late generic entry dates. Thus, the more antitrust enforcement is subject to error under the rule of reason, the more settling companies collude.

Intuitively, these results speak for an amendment toward the rule of per se illegality as only under this rule collusion can entirely be prevented. However, as a main result we show that, contrary to intuition, collusion can be beneficial. We presume that generic companies' incentives to challenge probabilistic patents are restricted due to the high additional costs that challenge processes necessitate. The possibility of collusion therefore provides generic companies additional incentives to challenge probabilistic patents as it increases their expected settlement profits. Additional settlement agreements result where otherwise the holders of probabilistic patents would remain monopolists. As long as collusion under the additional settlements is limited, competition increases, affecting consumer welfare positively.

Under the rule of per se legality settling companies collude maximally, so that consumer welfare does not increase due to the additional patent challenges. Therefore, the rule of per se legality yields the lowest consumer welfare compared to the alternative rules. The rule of reason has the paradoxical advantage over the rule of per se illegality that it induces limited collusion, thereby enhancing generic companies' incentives to challenge probabilistic patents. We show that the rule of reason outperforms the rule of per se illegality when generic companies' incentives to enter are low. In that case, the benefit adhered to collusion, i.e., the benefit of additional settlement agreements, outweighs the cost adhered to collusion, i.e., the cost that under each settlement competition is restrained.

We make a critical assessment of this first result by asking whether there exist alternative incentive mechanism to foster generic entry that are more effective than permitting collusion between originator and generic companies. The US Hatch-Waxman Act of 1984 potentially provides such an alternative incentive mechanism. It stipulates that generic companies first challenging a patent obtain 180 days of marketing exclusivity during which no subsequent generic company may enter. As a second result we find that this prominent incentive device does in fact not have the desired incentive effect and is detrimental to consumer welfare.

In our welfare analysis we only consider effects on static efficiency, resulting from competition among existing products. In Section 1.6, we also consider potential effects on dynamic efficiency, resulting from the creation of new products. On the one hand, the possibility of collusion could have the additional beneficial effect that it increases originator companies' expected settlement profits and thus their incentives to innovate. On the other hand, the possibility of collusion might impair dynamic efficiency as it has the effect that originator companies are able to obtain relatively high profits with weak inventions, which might negatively bias their investment decisions.

In the previous literature the vast majority of researchers also argues for an amendment of the current legal approach. Willig and Bigelow (2004) and Addanki and Daskin (2008) argue in favor of the rule reason because it allows for payments from originator to generic companies and such payments could for various reasons be necessary to facilitate procompetitive settlements.[3] However, they do not take into account that under the rule of reason courts might make errors in their evaluations and approve anticompetitive settlements. As briefly mentioned by Salinger, Ippolito, and Schrag (2007), when there exists the chance that anticompetitive settlements get approved, this might induce companies to conclude not pro- but anticompetitive settlements.[4] Our analysis accounts for the fact that antitrust evaluations under the rule of reason might be subject to error. To the best of our knowledge this is the first formal analysis in which the welfare effects of the different legal rulings toward pay-for-delay settlements are compared.

The structure of the paper is as follows. In Section 1.2, we outline the model. Section 1.3 analyzes to what extent settling companies collude and how high generic companies' incentives to contest probabilistic patents are under the different rules. In Section 1.4, we determine how the legal ruling affects welfare. Section 1.5 assesses the welfare effects of the Hatch-Waxman Act provisions. Results are discussed in Section 1.6 and Section 1.7 concludes.

[3]Other authors arguing that the majority of pay-for-delay settlements should be procompetitive, include Blair and Cotter (2002), Crane (2002), Langenfeld and Li (2003) and Schildkraut (2004).

[4]Bulow (2003), Hemphill (2006, 2009), Hovenkamp, Janis and Lemley (2003), Leffler and Leffler (2004), Lemley and Shapiro (2005), O'Rourke and Brodley (2003), and Shapiro (2003) also argue that pay-for-delay settlements should create a presumption of anticompetitive behavior.

1.2 The Model

We analyze a marketplace for pharmaceuticals, which originator companies (denoted O) have entered with patented products. For each patented product there are two companies sequentially seeking market entry with generic products prior to the patents' expiration.[5] At time $t = 0$ the first generic companies (denoted G_1) decide on entry, and at time $t = \lambda$, with $\lambda \in (0, 1]$, the second generic companies (denoted G_2) decide on entry. Patent exclusivity ends at time $t = 1$.

When entering, the generic companies trigger patent disputes. They incur a fixed cost f_g, which consists of proving bioequivalence and of bringing forward a detailed description of why they believe that the originator's patent is invalid or non-infringed. These fixed costs are not precisely known by other companies. The patent disputes result in either litigations or settlement agreements. As will be shown in Section 1.3, both parties view settlements as superior to litigations because the profits they can obtain by settling are at least as high as expected litigation profits.[6] However, the expected outcome of litigation is the basis for negotiation in the settlement talks. If negotiations break down and no bargain can be reached, the expected outcome of litigation is the value the players receive. When settling, the companies decide on generic entry dates and if permitted on value transfers.[7]

Nature determines the probabilities with which courts would declare the patents valid under litigations. These probabilities, denoted by γ, are common knowledge to originator and generic companies, and uniformly distributed between 0 and 1, i.e., $\gamma \sim U[0, 1]$.[8] The probability of patent validity reflects the strength of the patents. When γ equals zero a patent is invalid, and when γ equals one a patent is ironclad.[9]

---

[5] We consider the case of *two* generic companies *sequentially* seeking market entry because in Section 1.5 we want to investigate the welfare effects of the Hatch-Waxman Act provisions, which stipulate that generic companies first entering award an exclusivity right which delays subsequent generics' entry. Further, the FTC (2002) reports that at most two generic companies challenged probabilistic patents in the past. This number is lower than the number of generic companies seeking market entry *after* patent expiration because a challenge process necessitates additional costs, amounting to \$1 million. Our results also hold for simultaneous entry, in which case $\lambda = 0$. We do not consider this case explicitly for the sake of brevity.

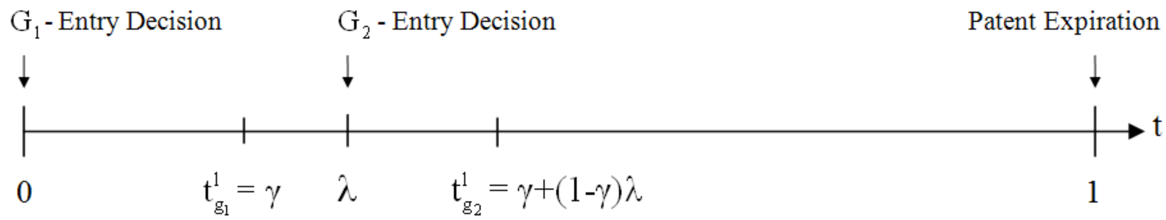
[6] Further reasons why companies prefer settlements can be that settlements costs are lower than litigation costs and that settlements provide legal certainty (see appendices A3 and A4).

[7] Likewise, settling parties could restrict competition by agreeing upon per-unit royalty rates, a fixed price, quantity-restrictions, territory dispartments or mergers. We assume that these other settlement forms are regulated such that none of them yields higher returns than ongoing litigation.

[8] In Appendix A3, we deal with cases in which the companies misperceive γ.

[9] In practice, expected settlement profits also depend on the commercial value of the patents. Further, patents might be strong but easy to 'invent around'. The results we obtain regarding patents of low strength γ also hold true for patents of high commercial value and patents non-infringed with probability γ.

Under litigations generic entry would be uncertain. With probability γ courts would declare the patents valid and generic entry would occur at the end of patent duration, i.e., at $t = 1$, whereas with probability $1-\gamma$ courts would declare the patents invalid and generic entry would occur immediately (G_1 would enter at $t = 0$ and G_2 would enter at $t = \lambda$).[10] By contrast, under settlements the companies can agree on any *certain* future entry date $t_{g_1} \in [0, 1]$ and $t_{g_2} \in [\lambda, 1]$. In order to be able to compare the uncertain entry dates under litigations with the certain entry dates under settlements, we make use of a continuous time model without discounting. We compute probabilistic weightings of the uncertain entry dates under litigations, which we will refer to as expected entry dates under litigations. G_1's expected entry date under litigation is $t^l_{g_1} = \gamma \cdot 1 + (1-\gamma) \cdot 0 = \gamma$ since under litigation G_1 enters with probability γ at time $t = 1$ and with probability $(1 - \gamma)$ at time $t = 0$.[11] For settling companies as well as consumers this entry date is of equal value as a γ percent chance of G_1 entering at the end of patent duration (at $t = 1$) and a $1-\gamma$ percent chance of G_1 entering immediately (at $t = 0$).[12] Accordingly, G_2's expected entry date under litigation is $t^l_{g_2} = \gamma \cdot 1 + (1 - \gamma) \cdot \lambda$ as under litigation G_2 enters with probability γ at time $t = 1$ and with probability $1 - \gamma$ at time $t = \lambda$.



Suppose, for example, that at the point in time G_1 challenges O's patent, the remaining life of the patent is 10 years and the probability of patent validity 20 percent (i.e., $\gamma = 0.2$). Then, G_1's expected entry date under litigation is equal to 2 years from the point in time G_1 challenges the patent ($t^l_{g_1} = \gamma \cdot 10$ years). If G_2 challenges O's patent three years later than G_1 (so that $\lambda = 0.3$), the remaining life of the patent is 7 years. Thus, G_2's expected entry date under litigation is equal to 1.4 years from

[10]We assume that a dispute is triggered and resolved at the same point in time. If we denoted two separate points in time, both parties' reference entry dates would postpone, and with it the entry dates under the rule of per se illegality and under the rule of reason. The welfare implications would not change.

[11]In accordance with our assumption that companies view settlements as superior to litigation, G_1 expects G_2 to settle or to stay off the market when itself litigates.

[12]Note that settlements are not necessarily desirable from a consumer welfare perspective as they only lead to *inter partes* clarification (among settling parties) while litigations, if the outcome is in favor of the generic companies, have *erga omnis* obligation (also apply to third parties). If O and G_1 settle, G_2 will only be able to enter when it has the means to litigate itself. While if O and G_1 litigate and the patent gets invalidated, G_2 will be free to enter immediately. Settlements can, however, not be prohibited, they can only be regulated.

the point in time G_2 challenges the patent and 4.4 years from the point in time G_1 challenges the patent ($t^l_{g_2} = \gamma \cdot 10$ years $+ (1-\gamma) \cdot \lambda \cdot 10$ years).

For simplification, we assume that no other substitutes are available to the originators' patents. Thus, in the absence of generic entry, the market exhibits a monopoly and O makes profits π^m. If G_1 enters, the market will exhibit a duopoly and O and G_1 will make profits π^d_o and π^d_g. If G_2 additionally enters, the market will exhibit a triopoly and companies will make profits π^t_o and π^t_g. We assume that all companies have identical production technologies. However, the profits G_1 and G_2 can make are strictly lower than that of O as their products are perceived inferior by consumers.[13]

Settlements are ruled by courts. The ruling affects the settlement terms, i.e., the generic entry dates that settling companies agree upon, and generic companies' incentives to enter and to challenge probabilistic patents in the first place. Specifically, the more settling companies collude under a rule, meaning the more they delay generic entry compared to litigation, the higher are their expected settlement profits and thus generic companies' incentives to enter.

Collusion increases settling companies' profits because it implies a prolongation of monopoly (or duopoly), through which a surplus s is generated. We assume that settling companies divide this surplus equally among them, that is, the originator company receives the same share in surplus as any of the generic companies. The higher the share in surplus that generic companies receive, the higher are their expected settlement profits, and thus their incentives to enter when collusion is possible.[14]

We apply the consumer welfare standard to assess which rule is most favorable. Absent generic entry, the market exhibits a monopoly between $t = 0$ and $t = 1$. Whereas with generic entry, the market exhibits a monopoly between $t = 0$ and t_{g_1}, a duopoly between t_{g_1} and t_{g_2}, and a triopoly between t_{g_2} and $t = 1$. It follows that collusion affects consumer welfare negatively as it implies a delay in t_{g_1} and t_{g_2} under each concluded settlement. Collusion may, however, also have a positive effect as generic companies' incentives to enter increase, so that more settlements are concluded. Consumer welfare under the additional settlements is higher than under monopoly given $t_{g_1} < 1$ and $t_{g_2} \leq 1$, i.e., given collusion is limited. Thus, there exists the trade-off that the more settling companies collude under a rule, the more competition is restrained under each concluded settlement, but the higher is the number of concluded settlements.

[13]For statistics showing that generic products are perceived as inferior by consumers see, e.g., European Commission (2008, para. 171, 189, and Table 12).

[14]In Appendix A3, we do not assume that settling companies share the surplus equally but compute the Nash Bargaining solution. Under the Nash Bargaining solution generic companies also receive a positive share in surplus. Note, if they received no share in surplus, collusion would not have an incentive effect and then the rule of per se illegality would always outperform the rule of reason.

1.3 Equilibrium Analysis

In this section, we analyze how the choice of the rule affects the degree of collusion and generic companies' incentives to enter. To assess the degree of collusion we ask by how much settling companies delay generic entry compared to litigation. And to assess generic companies' incentives to enter we ask for which values of patent strength they find it profitable to enter.

1.3.1 Equilibria under the Rule of Per Se Illegality

Under the rule of per se illegality (denoted pi) payments from originator to generic companies within settlement agreements are per se illegal. The companies can negotiate over generic entry but originator companies cannot compensate generic companies for a delay in entry compared to litigation. Because compensations are not possible, no party is willing to accept less favorable entry terms than the ones expected under litigation. The companies therefore agree upon $t^{pi} = t^l$. They do not collude. Accordingly, the companies' settlement profits are equal to their expected litigation profits:[15]

$$\pi_o^{pi} \equiv \pi_o^l = \begin{cases} \gamma\pi^m + (1-\gamma)[\lambda\pi_o^d + (1-\lambda)\pi_o^t] & \text{if } \gamma \in [0, \gamma_{g2}^{pi}] \\ \gamma\pi^m + (1-\gamma)\pi_o^d & \text{if } \gamma \in (\gamma_{g2}^{pi}, \gamma_{g1}^{pi}] \\ \pi^m & \text{if } \gamma \in (\gamma_{g1}^{pi}, 1] \end{cases}$$

$$\pi_{g_1}^{pi} \equiv \pi_{g_1}^l = \begin{cases} (1-\gamma)[\lambda\pi_g^d + (1-\lambda)\pi_g^t] & \text{if } \gamma \in [0, \gamma_{g2}^{pi}] \\ (1-\gamma)\pi_g^d & \text{if } \gamma \in (\gamma_{g2}^{pi}, \gamma_{g1}^{pi}] \\ 0 & \text{if } \gamma \in (\gamma_{g1}^{pi}, 1] \end{cases}$$

$$\pi_{g_2}^{pi} \equiv \pi_{g_2}^l = \begin{cases} (1-\gamma)(1-\lambda)\pi_g^t & \text{if } \gamma \in [0, \gamma_{g2}^{pi}] \\ 0 & \text{if } \gamma \in (\gamma_{g2}^{pi}, 1]. \end{cases}$$

Here, $\gamma_{g_1}^{pi}$ and $\gamma_{g_2}^{pi}$ describe the critical values of patent strength for which the generic companies are indifferent between entering or not. They are defined by

$$\pi_{g_1}^l(\gamma) + \frac{s^{pi}}{2} - f_g = 0 \quad \Leftrightarrow \quad \gamma_{g_1}^{pi} = 1 - \frac{f_g - \frac{s^{pi}}{2}}{\pi_g^d},$$

$$\pi_{g_2}^l(\gamma) + \frac{s^{pi}}{3} - f_g = 0 \quad \Leftrightarrow \quad \gamma_{g_2}^{pi} = 1 - \frac{f_g - \frac{s^{pi}}{3}}{(1-\lambda)\pi_g^t}.$$

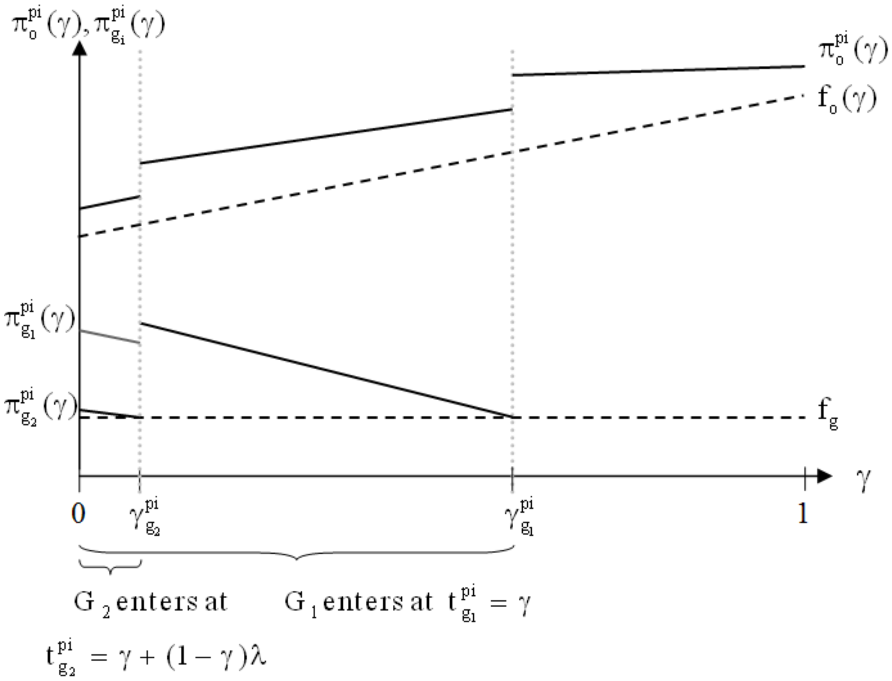
The surplus generated through settlements compared to litigations, s, is zero under this rule.[16] γ_{g_1} and γ_{g_2} show for which values of patent strength the companies find it profitable to enter. Here, G_1 enters for $\gamma \in [0, \gamma_{g_1}^{pi}]$ and G_2 enters for $\gamma \in [0, \gamma_{g_2}^{pi}]$.

[15] For simplification we do not consider litigation and settlement costs here but in Appendix A3.

[16] The surplus is divided by three (O, G_1, and G_2) if G_1 and G_2 enter and by two if only G_1 enters.

Since G_2's entry decision is delayed, its expected settlement profits are lower than G_1's, and so $\gamma_{g2}^{pi} < \gamma_{g1}^{pi}$. This is illustrated in Figure 1.1, which plots generic and originator companies' expected settlement profits (π_o^{pi}, π_{g1}^{pi}, π_{g2}^{pi}) and their fixed entry costs ($f_o(\gamma)$, f_g) against values of patent strength (γ). We look at a marketplace, in which originators hold patents that would be declared valid by courts with probabilities $\gamma \in [0, 1]$.

Figure 1.1: Market outcomes under the rule of per se illegality



The figure shows that the generic companies' expected settlement profits decrease with the probability of patent validity, i.e., with the patents' strength. This is because the companies will settle upon later entry dates if the probability of patent validity is higher. For each value of patent strength G_1's expected settlement profits are higher than that of G_2 as it enters earlier. Consequently, G_1 can cover the fixed costs, f_g, for values of patent strength $\gamma \in [0, \gamma_{g1}^{pi}]$, while G_2 can cover f_g only for $\gamma \in [0, \gamma_{g2}^{pi}]$, with $\gamma_{g2}^{pi} < \gamma_{g1}^{pi}$. It follows, when $\gamma \in [0, \gamma_{g2}^{pi}]$ both generic companies enter. In that case, monopoly lasts from $t = 0$ to $t_{g1}^{pi} = \gamma$, duopoly from $t_{g1}^{pi} = \gamma$ to $t_{g2}^{pi} = \gamma + (1-\gamma)\lambda$, and triopoly from $t_{g2}^{pi} = \gamma + (1-\gamma)\lambda$ to $t = 1$. Instead, when $\gamma \in (\gamma_{g2}^{pi}, \gamma_{g1}^{pi}]$ only G_1 enters and duopoly lasts from $t_{g1}^{pi} = \gamma$ to $t = 1$. Further, when $\gamma \in (\gamma_{g1}^{pi}, 1]$ no generic company enters and monopoly lasts from $t = 0$ to $t = 1$.

Result 1.1 *Under the rule of per se illegality value transfers within settlement agreements are illegal, so that originator companies cannot compensate generic companies* *for a delay in entry compared to litigation. The companies settle on generic entry terms that would in expectation result under litigation. Collusion does not arise.*

1.3.2 Equilibria under the Rule of Per Se Legality

Under the rule of per se legality (denoted pl) payments from originator to generic companies are per se legal. The companies can negotiate over generic entry and originator companies can compensate generic companies for any delay in entry compared to litigation. They therefore settle on entry terms that maximize their joint profits. Since joint profits,

$$\Pi^{pl} = t^{pl}_{g_1}\pi^m + (t^{pl}_{g_2} - t^{pl}_{g_1})(\pi^d_o + \pi^d_g) + (1 - t^{pl}_{g_2})(\pi^t_o + 2\pi^t_g),$$

are an increasing function of $t^{pl}_{g_1}$ and $t^{pl}_{g_2}$, they choose the latest possible entry dates, i.e., $t^{pl}_{g_1} = t^{pl}_{g_2} = 1$.[17] Monopoly is sustained for the whole patent duration. That is, collusion is maximal.

Settling companies create a surplus compared to litigation equal to

$$\begin{aligned} s^{pl}_1 &= (1-\gamma)\left[\pi^m - \lambda(\pi^d_o + \pi^d_g) - (1-\lambda)(\pi^t_o + 2\pi^t_g)\right] && \text{if } \gamma \in [0, \gamma^{pl}_{g2}], \\ s^{pl}_2 &= (1-\gamma)\left[\pi^m - \pi^d_o - \pi^d_g\right] && \text{if } \gamma \in (\gamma^{pl}_{g2}, \gamma^{pl}_{g1}]. \end{aligned}$$

Because the generated surplus is higher than under the rule of per se illegality, the companies' expected settlement profits increase. As a result the critical levels of patent strength, for which the generic companies are indifferent between entering or not, are higher:

$$\gamma^{pl}_{g_1} = 1 - \frac{f_g - \frac{s^{pl}_2}{2}}{\pi^d_g} \quad \text{and} \quad \gamma^{pl}_{g_2} = 1 - \frac{f_g - \frac{s^{pl}_1}{3}}{(1-\lambda)\pi^t_g}.$$

This means, generic entry takes place more often than under the rule of per se illegality. However, competition does not enhance. Under all concluded settlements monopoly lasts from $t = 0$ until $t = 1$.

Result 1.2 *Under the rule of per se legality value transfers within settlement agreements are legal, so that originator companies can compensate generic companies for a delay in entry compared to litigation. The companies maximize their joint profits by delaying generic entry until the end of patent duration. Collusion is maximal.*

[17] Agreements determining generic entry dates later than $t^{pl} = 1$ would be illegal by competition law as at stage $t = 1$ the status of the patent terminates. Accordingly, an agreement that guaranteed to Hoechst Marion Rousselt that its generic competitor, Andrx, would, for the price of \$10 million per quarter, refrain from marketing its generic version of Cardizem CD even after it had obtained FDA approval, has been judged anticompetitive.

1.3.3 Equilibria under the Rule of Reason

The rule of reason (denoted rr) is usually implemented as a three-step process. Initially, the plaintiffs may show whether there are adverse effects on competition. Subsequently, the defendants may "establish procompetitive redeeming virtues of the action." And finally, the plaintiffs may "show that the same procompetitive effects could not be achieved through an alternative means that is less restrictive of competition."[18] In case of patent settlements, plaintiffs can only make use of indicators for adverse effects on competition in the first step.[19] Consequently, the insights that courts obtain will presumably be incomplete. This means, courts will presumably only have vague ideas of whether particular settlement agreements are pro- or anticompetitive, making consumers better or worse off compared to litigation. Therefore, we assume that courts predict the expected generic entry dates under litigation with errors ϵ, such that $\hat{t}^l = t^l + \epsilon$.[20] These errors, ϵ, are uniformly distributed between e and $-e$, that is, $\epsilon \sim U[-e, e]$. An increase in e means courts' predictions of t^l become less precise, or put differently, the chance that courts approve anticompetitive settlements increases.

Accordingly, courts prohibit settlements if the generic entry date that the companies agreed upon lies post to the predicted entry date under litigation ($t^{rr} > \hat{t}^l$). In case of prohibition, the companies have to execute less restrictive agreements where such agreements are available. We assume that the resulting entry terms are neutral in case of prohibition ($t^{rr} = t^l$). We further assume, if courts prohibit a first settlement between O and G_1, they will also prohibit a second settlement between O and G_2 that assigns a corresponding entry date.[21] The probability with which courts prohibit first settlements is

$$\text{Prob}(t_{g_1}^{rr} > \hat{t}_{g_1}^l) = \text{Prob}(\epsilon < t_{g_1}^{rr} - \gamma) = \frac{t_{g_1}^{rr} - \gamma + e}{2e}. \tag{1.1}$$

[18] Citing Clorox Co. vs. Sterling Winthrop, 117 F.3d 50 (2d Cir. 1997).

[19] Indicators for adverse effects include the amounts of value transfers relative to the patents' market value or the agreed upon generic entry dates. Further, plaintiffs can gather evidence through search for prior art, and through examinations of backward citations and patent claims in the patent applications.

[20] It is natural to assume that predictions of $\hat{t}^l > 1$ or $\hat{t}^l < 0$ do not occur as these would be predictions that a patent is valid with more than 100% or less than 0%, respectively.

[21] If the first settlement assigns $t_{g_1}^{rr}$, the corresponding entry date of a second settlement would be $t_{g_2}^{rr} = t_{g_1}^{rr} + (1 - t_{g_1}^{rr})\lambda$ because $t_{g_1}^l = \gamma$ and $t_{g_2}^l = \gamma + (1-\gamma)\lambda$.

Given this probability, companies' expected joint settlement profits are

$$\Pi^{rr} = \begin{cases} \frac{t_{g1}^{rr}-\gamma+e}{2e}\Big[\gamma\pi^m \;+\; (1-\gamma)[\lambda(\pi_o^d+\pi_g^d) \;+(1-\lambda)(\pi_o^t+2\pi_g^t)]\Big] \\ \\ +\frac{e-t_{g1}^{rr}+\gamma}{2e}\Big[t_{g1}^{rr}\pi^m+(1-t_{g1}^{rr})[\lambda(\pi_o^d+\pi_g^d)+(1-\lambda)(\pi_o^t+2\pi_g^t)]\Big] \quad \text{if } \gamma\in[0,\ \gamma_{g2}^{rr}], \\ \frac{t_{g1}^{rr}-\gamma+e}{2e}\Big[\gamma\pi^m+(1-\gamma)(\pi_o^d+\pi_g^d)\Big] \\ \\ +\frac{e-t_{g1}^{rr}+\gamma}{2e}\Big[t_{g1}^{rr}\pi^m+(1-t_{g1}^{rr})(\pi_o^d+\pi_g^d)\Big] \quad \text{if } \gamma\in(\gamma_{g2}^{rr},\ \gamma_{g1}^{rr}]. \end{cases}$$

The companies maximize these expected joint settlement profits by agreeing upon entry dates $t_{g1}^{rr} = \min[\gamma+e/2\ ,\ 1]$ and $t_{g2}^{rr} = \min\big[t_{g1}^{rr}+(1-t_{g1}^{rr})\lambda\ ,\ 1\big]$. This gives us the following result.

Result 1.3 *Under the rule of reason companies are allowed to transfer values within settlements but courts prohibit settlements which they regard anticompetitive. The less precisely courts evaluate settlements, the more companies collude because the higher is the chance that their anticompetitive settlements get approved.*

When deciding on generic entry dates settling companies face the following trade-off. The more they collude, the higher profits they obtain in case of settlement approval but the higher is also the probability that their settlement gets prohibited. If courts evaluate settlement agreements more precisely, so that anticompetitive settlements get more likely prohibited, it will pay less for the companies to choose particularly late generic entry dates.

In this section, we restrict our attention to the case when companies optimally choose $t_{g1}^{rr} = \gamma + e/2 < 1$.[22] Inserting $t_{g1}^{rr} = \gamma + e/2$ into equation (1.1) shows that courts approve settlements with probability $\text{Prob}(t_{g1}^{rr} < \hat{t}_{g1}^{l}) = 1/4$. In that case, the companies generate a surplus compared to litigation equal to

$$\begin{array}{lll} s_1^{rr} & = \frac{e}{2}\left[\pi^m-\lambda(\pi_o^d+\pi_g^d)-(1-\lambda)(\pi_o^t+2\pi_g^t)\right] & \text{if } \gamma\in[0,\gamma_{g2}^{rr}], \\ s_2^{rr} & = \frac{e}{2}\left[\pi^m-\pi_o^d-\pi_g^d\right] & \text{if } \gamma\in(\gamma_{g2}^{rr},\gamma_{g1}^{rr}]. \end{array}$$

The critical levels of patent strength, for which the generic companies are indifferent between entering or not, are given by

$$\gamma_{g1}^{rr} = 1-\frac{f_g-\frac{s_2^{rr}}{8}}{\pi_g^d} \quad \text{and} \quad \gamma_{g2}^{rr} = 1-\frac{f_g-\frac{s_1^{rr}}{12}}{(1-\lambda)\pi_g^t}.$$

[22] We analyze the case when companies optimally choose $t_{g1}^{rr} = 1$ in Appendix A1.

It is easy to see that the generated surplus and thus the critical levels of patent strength increase with e, i.e., with the imprecision of antitrust evaluations. The larger e, the more the companies collude and the higher are generic companies' incentives to enter.

Result 1.4 *By colluding, settling companies generate a surplus which they can divide among each other. That way, expected settlement profits increase, and generic companies obtain additional incentives to enter.*

1.4 Welfare Analysis

A central question of this paper is under which conditions which rule toward pay-for-delay settlements is preferable from a consumer welfare perspective. The previous analysis has shown that under the rule of per se legality settling companies sustain monopoly for the whole patent duration. Thus, initiated challenges do not lead to increased competition. For all values of patent strength, monopoly lasts from $t = 0$ to $t = 1$. This is different under the rule of per se illegality and under the rule of reason. Under these alternative rules initiated challenges of weaker patents lead to increased competition. Therefore, we obtain the following result.

Result 1.5 *Only under the rule of per se legality patent challenges do not lead to increased competition. Thus, the rule of per se legality yields the lowest consumer welfare.*

What remains unanswered is under which conditions the rule of reason or the rule of per se illegality yields higher consumer welfare. We know that under these two rules the companies settle upon:

$$\begin{aligned} t_{g1}^{pi} &= \gamma & \text{and} \quad t_{g2}^{pi} &= \gamma + (1-\gamma)\lambda, \\ t_{g1}^{rr} &= \gamma + \tfrac{e}{2} & \text{and} \quad t_{g2}^{rr} &= \gamma + \tfrac{e}{2} + (1-\gamma-\tfrac{e}{2})\lambda. \end{aligned}$$

Here, e shows how precisely courts evaluate patent settlements under the rule of reason. If courts are able to evaluate patent settlements without error, e equals zero. In that case, generic entry and consumer welfare is the same under both rules. Thus, when a marginal increase in e, at the point where $e = 0$, improves consumers welfare under the rule of reason, patent settlements should be treated under this standard. We obtain the following the result.

Proposition 1.1 *The rule of reason yields higher consumer welfare than the rule of per se illegality if generic companies' incentives to challenge probabilistic patents are* *sufficiently weak, i.e., if* $\gamma_{g1}^{rr} < \gamma_{g1}^{rr'} \equiv \frac{\Delta_1}{\Delta_1 + 2\pi_g^d}$ *and* $\gamma_{g2}^{rr} < \gamma_{g2}^{rr'} \equiv \frac{\Delta_2}{\Delta_2 + 3(1-\lambda)\pi_g^t}$ *with* $\Delta_1 = \pi^m - \pi_o^d - \pi_g^d$ *and* $\Delta_2 = \pi^m - \lambda(\pi_o^d + \pi_g^d) - (1-\lambda)(\pi_o^t + 2\pi_g^t)$.

The proof is relegated to Appendix B. Proposition 1.1 shows that central to the condition ensuring that the rule of reason outperforms the rule of per se illegality is, how strong generic companies' incentives to challenge probabilistic patents are. If generic companies' incentives to enter are low, the rule of reason likely outperforms the rule of per se illegality. Generic companies' incentives to enter depend negatively on their fixed entry costs f_g and positively on their expected settlement profits π_g (see also Figure 1.1). As shown in Appendix B, the condition in Proposition 1.1 can be rewritten as

$$f_{g_1} > \frac{2\pi_g^{d^2}}{2\pi_g^d + \Delta_1} \quad \text{and } f_{g_2} > \frac{3\big((1-\lambda)\pi_g^t\big)^2}{(1-\lambda)\pi_g^t + \Delta_2}.$$

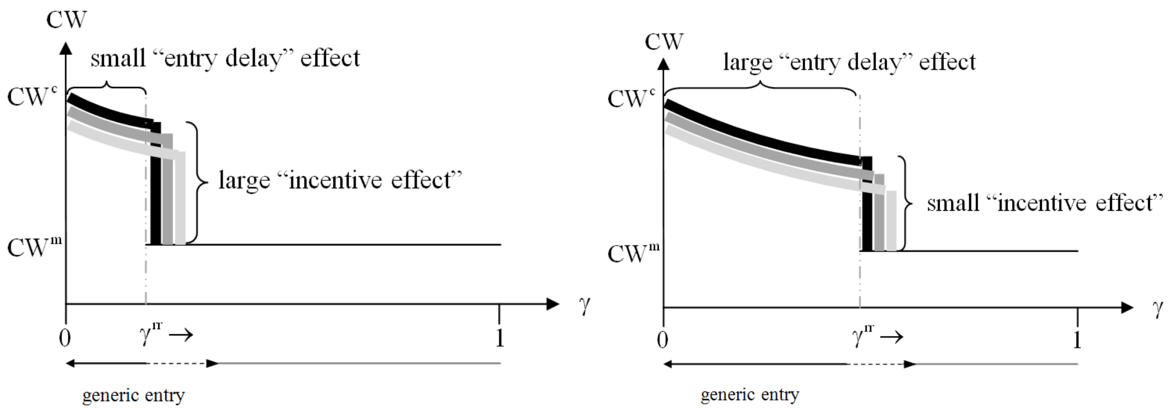
The fixed costs consist of proving bioequivalence and of bringing forward arguments why the originator's patent could potentially be invalid or non-infringed. They vary with the type of challenge. A non-infringement claim, if readily available, is generally easier to conduct than an invalidity claim.[23] Generic companies' expected settlement profits depend on numerous factors, *inter alia* on the extent to which consumers perceive generic products as inferior to original products.

The explanation for why the rule of reason outperforms the rule of per se legality when generic companies' incentives to enter are low is the following. The rule of reason outperforms the rule of per se illegality when a marginal increase in e, at the point where $e = 0$, has positive welfare implications. A marginal increase of e induces settling companies to collude. This affects consumer welfare positively when the negative effect adhered to collusion that under each concluded settlement generic entry is delayed, is outweighed by the positive effect adhered to collusion that additional settlements result as generic companies obtain higher incentives to enter. The negative 'entry delay' effect is small when generic companies' incentives to enter are low because in that case, overall, few settlements are concluded. Since under each settlement entry is equally delayed, the total negative effect remains small. At the same time, the positive 'incentive effect' is big when generic companies' incentives to enter are low because the additional patents that get challenged are of relatively weak strength, which implies that under the additional settlements the companies choose relatively early entry dates and so, competition increases strongly.

[23] In Schering-Plough vs. FTC, for instance, the active ingredient of Schering-Plough's pharmaceutical was an unpatented potassium salt, so that two generic companies could relatively easily come up with alternative, non-infringing means of achieving bioequivalence.

This can be inferred from Figure 1.2. Here, consumer welfare (CW) is plotted against patent strength (γ). CW^c denotes consumer welfare under competition when generic entry occurs immediately and CW^m denotes consumer welfare under monopoly when generic entry occurs at the end of patent duration. Generic companies enter for values of patent strength between 0 and γ^{rr}. So, a low γ^{rr} indicates that generic companies' incentives to enter are low. The figure shows the effects of successive marginal increases in e. An increase in e has the negative effect that under each concluded settlement consumer welfare decreases as the companies delay generic entry. It can be seen, when γ^{rr} is low, the number of concluded settlements is low, so that the aggregate entry delay effect remains low. Further, an increase in e has the positive effect that γ^{rr} increases, meaning generic companies obtain higher incentives to enter and additional settlements are concluded. When γ^{rr} is low, consumer welfare increases strongly because the additional patents that get challenged are of relatively low strength, implying that the companies settle upon early generic entry dates. When entry occurs earlier, consumer welfare increases by more.

Figure 1.2: Effects of a marginal increase of e on consumer welfare



1.5 Welfare Implications of the Hatch-Waxman Act

The welfare analysis showed that the application of the rule of reason toward pay-for-delay settlements can yield higher consumer welfare than the application of the rule of per se illegality because settling companies are induced to collude, which increases generic companies' expected settlement profits and with it their incentives to challenge probabilistic patents. Incentivization is achieved through restraints of competition between originator and generic companies. If there was no lack of incentives, the rule of reason would clearly yield lower consumer welfare than the rule of per se illegality.

The question arises how strong generic companies' incentives to challenge contestable patents are in practice. And moreover, we may bring into question whether it is effective to let settling parties collude in order to achieve incentivization.

Due to lack of data, researchers have not empirically tested how strong generic companies' incentives to challenge probabilistic patents are in practice. We know, however, if it comes to patent litigation, the risk that the patents will be declared invalid or non-infringed by courts is substantial. The FTC (2002) calculated that in 73% of Hatch-Waxman cases (see below), the generic company was found not to have infringed a valid patent.[24] This indicates that generic companies are only willing to involve originator companies in patent disputes when their chances of winning are relatively high. Presumably, generic companies only challenge probabilistic patents of relatively low strength or high commercial value.[25]

To foster generic entry, competition authorities could alternatively strengthen generic companies' market power. For instance, they could encourage the prescription of generic products. However, in the first instance this would presumably lead to an increase in the number of generic companies in the market, and thus not necessarily provide each individual generic company stronger incentives to enter.

A prominent alternative incentive device to foster generic entry is the provision of an exclusivity right to first generic entrants as implemented within the Hatch-Waxman Act (HWA) of 1984 in the US.[26] The HWA awards first generic applicants to file an Abbreviated New Drug Application (ANDA) containing a paragraph IV certification with 180 days of marketing exclusivity, during which the Food and Drug Administration (FDA) may not approve a subsequent generic applicant's ANDA for the same pharmaceutical product (21 U.S.C. §355(j)(5)(B)(iv)).[27]

[24]Further information about generic companies' incentives to challenge probabilistic patents across countries can be inferred from factors like number of settlements in relation to product market values (see, e.g., EC 2008, Figure 97).

[25]Judge Posner stressed the importance of generic companies' incentives to challenge patents in Asahi Glass Co. v. Pentech Pharmaceuticals. The Asahi approach has been repeated and approved in cases *Tamoxifen*, *Schering-Plough vs. FTC*, *In re Cipro*, and further in papers by Balto (2004), and Schildkraut (2004).

[26]The Hatch-Waxman Act is formally known as Drug Price Competition and Patent Restauration Act of 1984, Pub.L.No 98-417, 98 Stat. 1585 (1984).

[27]The 180-day exclusivity period is calculated from either (i) first commercial marketing by the first generic applicant, or (ii) a decision of a court holding the relevant patents to be invalid or not infringed. The marketing exclusivity forfeits and subsequent applicants can enter at the same time as the first applicant when (i) an appeals court has ruled the relevant patents invalid or not infringed and (ii) 75 days after the effective date or 30 month after application filing elapsed. When the originator files within 45 days a patent infringement suit against any generic company that submits an ANDA, FDA's approval of the ANDA stays for at least 30 month during which time no generic can be launched.

The 180 days of marketing exclusivity implies a restraint of competition between generic companies. It increases first generic companies' expected profits and thus provides them higher incentives to enter and to challenge probabilistic patents, given subsequent generic companies would have entered.[28] The question is, whether it also provides them higher incentives to challenge *additional* probabilistic patents. In what follows, we will show that this is not the case. Since subsequent generic companies' incentives to enter decline, the effect of the HWA provisions is anticompetitive.

Proposition 1.2 *The effect of the HWA provisions is anticompetitive because first generic companies' incentives to challenge additional probabilistic patents do not improve ($\gamma_{g_1}^{hwa} \equiv \gamma_{g_1}$) but subsequent generic companies' incentives to challenge probabilistic patents decline ($\gamma_{g_2}^{hwa} < \gamma_{g_2}$).*

The proof is relegated to Appendix B. It has been presumed that the provision of exclusivity rights to first generic companies improves their incentives to enter, so that additional challenges result. However, the HWA provisions only lead to an increase in first generic companies' profits, improving their incentives to enter, when they find it profitable to enter anyway. This is because first generic companies have higher incentives to enter than subsequent generic companies, and the HWA provisions only lead to restraints of competition when subsequent entry would have occurred.[29] Thus, only when first generic companies find it profitable to enter anyway they obtain additional profits due to the HWA provisions. Put another way, when first generic companies do not find it profitable to enter and an increase in their profits would be desirable, leading to additional patent challenges, subsequent generic companies do not find it profitable to enter either, and therefore, an exclusivity right that restricts generic competition does not improve first generic companies' incentives to enter.

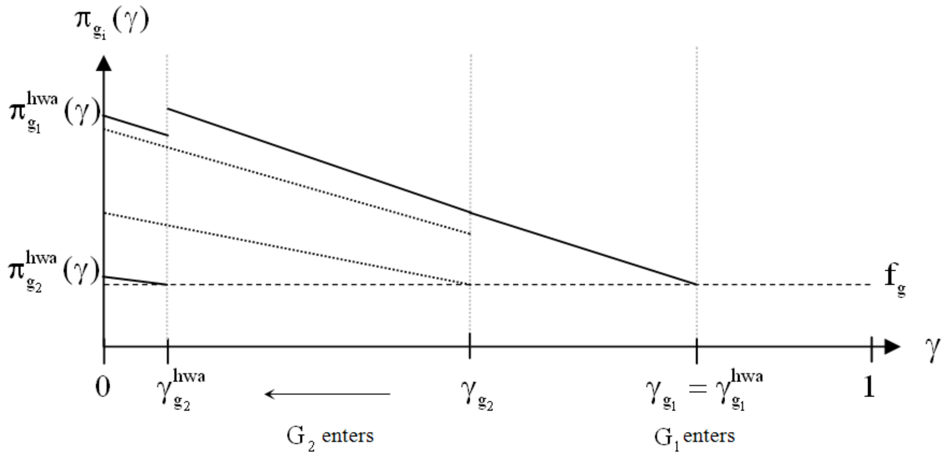
The argument is illustrated in Figure 1.3. It shows that second generic companies' expected settlement profits and thus their incentives to enter decrease under the HWA provisions ($\gamma_{g_2}^{hwa} < \gamma_{g_2}$). First generic companies' expected settlement profits, on the other hand, increase, but only for values of $\gamma \in [0, \gamma_{g_2})$. Since first generic companies' expected settlement profits do *not* increase for values of $\gamma \in (\gamma_{g_2}, 1]$ their incentives to enter remain unchanged ($\gamma_{g_1}^{hwa} \equiv \gamma_{g_1}$).

Arguably, if second generic companies obtained, despite their entry delay, higher settlement profits than first generic companies (e.g., due to lower fixed entry costs or

[28] Apotex, for instance, reported it earned between \$150 million and \$200 million from its marketing exclusivity on the antidepressant Paxil.

[29] Subsequent generic companies have lower incentives to enter than first generic companies as their entry is delayed ($\lambda > 0$).

Figure 1.3: THE EFFECT OF THE HWA PROVISIONS ON GENERIC COMPANIES' INCENTIVES TO ENTER



Note: The drawn through lines show generic companies' expected settlement profits under the HWA, while the dotted lines show their expected settlement profits absent the HWA.

higher productive efficiency), the HWA would be effective in providing first generic companies incentives to enter for additional probabilistic patents. In that case, such incentivization would, however, not improve consumer welfare either as first generic companies would challenge additional probabilistic patents that second generic companies would challenge absent the HWA provisions anyway, since in that case second generic companies have higher incentives to enter than first generic companies (i.e., $\gamma_{g2} > \gamma_{g1}$).

Further, one could argue that the HWA provisions lead to *earlier* entry by first generic companies as they provide generic companies additional incentives to be the first in the market. Similar to the effect that patent races have on entry dates of originator companies, the award of an exclusivity right to first generic companies might have an effect on entry dates of first generic companies. But even if this was the case, earlier entry of first generic companies would be accompanied by later entry of subsequent generic companies.

It follows that due to the HWA provisions no additional patents get challenged, fewer patents get challenged by both generic companies and, further, if patents get challenged by both generic companies, entry by second generic companies is delayed. We can therefore conclude that the HWA provisions do not establish the desired incentive effect, they only restrict competition when this is undesirable. Therefore, we recommend to abolish the provision of exclusivity rights to first generic entrants.[30]

1.6 Discussion

In this section, we discuss our results. We ask whether the problem of anticompetitive pay-for-delay settlements could be alleviated by raising patent quality across-the-board. We explain problems adhered to the rule of reason approach. And further, we ask whether in case of pay-for-delay settlements static efficiency is an acceptable criterion for long-run consumer welfare.

1.6.1 Patent Quality and Effective Opposition

We showed that patent settlements are a means to restrict competition. Arguably, policymakers could get to the root of the problem by strengthening the examination process, thereby raising patent quality across-the-board. Better funding of patent examinations, higher standards for initial review, better incentives that make it easier and more desirable for examiners to reject rather than grant patents and better incentives for applicants to disclose prior art could weed out weak patents in the first place. As several authors have stated, the optimal error rate at a patent office is, however, not zero. Because only very few patents have commercial significance perfect examination would not be cost effective.[31] Since patent offices lack information about which patents matter ahead of time they also cannot focus their examinations on the few important patents. Thus, government relies on litigation and on an effective opposition system to fix all errors.[32] However, as we pointed out in this paper, once an error has been made in the examination process and a patent dispute arises, the problem emerges that generic companies actually do not have incentives to litigate or make use of an opposition system, inducing a second review process, but to settle their disputes out of court. As

[30]Because the HWA provisions have the additional detrimental effect that originator companies basically only need to pay to delay entry by first generic companies in order to also delay entry by subsequent generic companies, Lemley and Hemphill (2011) suggest that first generic companies should only be awarded with an exclusivity right if they successfully defeat the originator companies (e.g., by invalidating their patents). We would like to point out that this kind of regulation would have the additional beneficial effect that generic companies' incentives to *litigate* would increase. When first generic companies are successful in litigations, this benefits subsequent generic companies and leads to increased competition overall (see supra note 12).

[31]Lemley (2001) points out that ninety-five percent of patents will either never be used, or will be used in circumstances that do not crucially rely on the determination of validity.

[32]For information on the opposition systems in the US and in the EU see, e.g., Harhoff, Scherer and Vopel (2003), Farrell and Merges (2004) and USPTO (2009).

long as courts apply the rule of per se legality toward pay-for-delay settlements, errors made in the examination process are not corrected for to the detriment of consumers. Thus, there is an urgent need to reconsider patent law's presumption of validity and to control patent settlements by competition authorities. Only the rule of reason and the rule of per se illegality provide an effective mechanism to correct for errors made in the examination process.

1.6.2 Problems Adhered to the Rule of Reason Approach

A drawback of the rule of reason approach is that ex post competition authorities have an incentive to prohibit every pay-for-delay settlement, knowing that it is optimal for companies to settle on anticompetitive terms. As a consequence, e, which shows to which degree antitrust evaluations of pay-for-delay settlements are subject to error, might become very small, even smaller than the theoretically optimal e.[33] However, low levels of e are desirable because they are likely to cause high positive incentive and low negative entry delay effects. In contrast, high levels of e evidently lead to welfare losses as the negative entry delay effects grow so large that they cannot be outweighed by the positive incentive effects. So the commitment problem faced by antitrust enforcers when evaluating pay-for-delay settlements could in fact help to implement desirable levels of e.

It should be further noted that antitrust enforcers are likely to face practical difficulties implementing the optimal level of e anyway. The optimal level of e depends on many different factors and fine-tuning on e is only possible through gathering more information on the merits of each case.[34] Therefore, it could make sense to implement alternatively to the rule of reason, a per se rule, under which value transfers from originator to generic companies within settlement agreements are permitted up to a specific amount. Such an approach would have similar beneficial effects as the rule of reason approach but would simplify matters. By linking the permitted value transfers to originator companies' returns and keeping the permitted amounts relatively low, competition authorities could appropriately balance entry delay against generic companies' incentives to challenge probabilistic patents and at the same time assure legal certainty.

[33]The optimal 'evaluation error' e is derived in Appendix A2.

[34]Alternatively, competition authorities could prevent companies from choosing highly anticompetitive settlement terms by imposing a fine which the companies have to pay in case of settlement prohibition or by imposing an *ad valorem* tax on payments made by originator companies. These measures would make collusion more costly for the companies, so that in effect the companies would collude less.

Also, the administration costs would be lower than under a full-fledged rule of reason analysis.

1.6.3 Static vs. Dynamic Efficiency

In the welfare analysis we focused on static efficiency, neglecting effects that the choice of rule might have on dynamic efficiency. Static efficiency means economic efficiency at a static level, resulting from competition among existing products, whereas dynamic efficiency means economic efficiency at a dynamic level, resulting from the creation of new products. Antitrust law is mainly concerned with static efficiency, defending market competition, whereas patent law is mainly concerned with dynamic efficiency, conferring market power to innovating companies. The basis for patent law is the Schumpeterian argument, which contends that the prospect of market power incentivizes companies to innovate. Accordingly, while competition promotes static efficiency, it might infer with dynamic efficiency. Different long-run consumer welfare implications might arise when considering dynamic efficiency. There are, however, a number of reasons one should not rely exclusively on the Schumpetarian argument. In some situations competition rather than the promise of market power promotes dynamic efficiency as it pushes companies to adopt the most efficient technologies and to invest in R&D.[35] Competition often spurs faster innovation and induces companies to innovate in different ways, resulting in a variety of different technologies. Thus, static and dynamic efficiency may go hand in hand.

When conducting a welfare analysis, we should nevertheless assess in case of restraints of competition how these restraints affect dynamic efficiency. Under the rule of per se legality and under the rule of reason competition is restrained due to collusion. So, the question arises whether collusion under these rules promotes or impairs dynamic efficiency. According to the Schumpeterian argument, originator companies might be incentivized to innovate due to the additional profits they can obtain by colluding. However, there is also an argument to the contrary. Dynamic efficiency might impair as the possibility of collusion effects that originator companies obtain relatively high profits when holding weak patents. Under the rule of per se legality, originator companies profit from collusion more if their patents are weaker. And also under the rule of reason, originator companies only profit from collusion when holding weak and not ironclad patents. This might cause originator companies to invest more in weak instead of ironclad patents. Since the effect of collusion on dynamic efficiency is not clear from the outset, this is a topic for further research.

[35]This argument has long been associated with Arrow (1962).

1.7 Concluding Remarks

In the EU, regulation of pay-for-delay settlements is still in its infancy. US courts apply the rule of per se legality toward pay-for-delay settlements. This leads to tremendous welfare losses as settling companies are able to use probabilistic patent terms to legitimate restraints of competition.[36] As an alternative to the rule of per se legality, courts could apply the rule of per se illegality, prohibiting settlement agreements that involve value transfers from originator to generic companies, or the rule of reason, evaluating pro- against anticompetitive settlement effects. We showed that in contrast to the rule of per se illegality, the rule of reason induces limited collusion between settling companies when antitrust enforcement under this rule is imperfect. On the one hand, this affects consumer welfare negatively as generic entry under each settlement is delayed compared to litigation. On the other hand, it affects consumer welfare positively as companies' expected settlement profits increase, which provides generic companies additional incentives to challenge probabilistic patents. Additional settlement agreements are concluded where otherwise the holders of probabilistic patents remained monopolists. We showed that if generic companies' incentives to challenge probabilistic patents are low, the negative entry delay effect is outweighed by the positive incentive effect, so that the rule of reason yields higher consumer welfare than the rule of per se illegality.

Under the rule of reason generic entry is promoted through restraints of competition between originator and generic companies. Instead, the Hatch-Waxman Act of 1984 aims to promote generic entry through restraints of competition between generic companies. More specifically, first generic companies are awarded with an exclusivity right, which restricts entry by subsequent generic companies. The analysis revealed that this very prominent incentive device fails to have the desired incentive effect. An increase in first generic companies' profits through restraints of competition would be desirable, leading to additional patent challenges, in cases in which first generic companies do not find it profitable to enter otherwise. However, in these cases subsequent generic companies also do not find it profitable to enter as they usually have lower incentives to enter than first generic companies as their entry is delayed. Thus, the provision of an exclusivity right to first generic companies is effectless in these cases. It still has the effect that competition is restrained when subsequent generic companies find it profitable to enter. Therefore, we recommend to abolish this regulation.

[36]According to the FTC's estimations the costs to consumers in the US increase through pay-for-delay settlements by approximately $3.5 billion each year (see Brief of the United States, available at http://www.justice.gov/atr/cases/f259300/259325.htm).

Although our analysis has shown that the application of antitrust rules toward pay-for-delay settlements would be beneficial from a consumer welfare perspective, it is very difficult to justify the application of antitrust rules toward pay-for-delay settlements in front of courts. This is because pay-for-delay settlements involve valid (but probabilistic) patents. Thus, further legal research needs to be done in this direction.

1.8 Appendix A: Extensions

A1: Settlements under the Rule of Reason when $\gamma > 1 - \frac{e}{2}$

In this appendix, we show that the results computed in Sections 1.3.3 and 1.4 for the rule of reason only hold for the case $\gamma > 1 - e/2$ when the critical level of patent strength, for which G_1 is indifferent between entering or not, does not exceed $\gamma + e/2$.

When $\gamma > 1 - e/2$ settling companies optimally choose $t_{g_1}^{rr'} = t_{g_2}^{rr'} = 1$ under the rule of reason. Inserting $t_{g_1}^{rr'} = 1$ into equation (1.1) shows that courts then prohibit first settlements with probability $\text{Prob}(t^{rr'} > \hat{t}^l) = (1 - \gamma + e)/2e$. The critical levels of patent strength, for which the generic companies are indifferent between entering or not, change to

$$\gamma_{g_1}^{rr'} = \frac{f_g - \frac{s_2^{rr'}+24}{2}e - \pi_g^d}{1 - \pi_g^d}, \text{ and}$$

$$\gamma_{g_2}^{rr'} = \frac{f_g - \frac{s_1^{rr'}+24}{3}e - (1-\lambda)\pi_g^t}{1 - (1-\lambda)\pi_g^t},$$

when they exceed $1 - e/2$. Then, the rule of per se illegality always yields higher consumer welfare than the rule of reason because an increase in e only has a negative entry delay but no positive incentive effect. A negative entry delay effect arises because settling parties are induced to collude more for values of patent strength below $1 - \frac{e}{2}$. A positive incentive effect, however, does not arise because settling parties are not induced to collude more for values of patent strength above $1 - \frac{e}{2}$. For these values settling companies constantly choose $t^{rr'} = 1$, regardless of whether e increases or not. Thus, generic companies' expected settlement profits do not increase with an increase in e for values around $\gamma_{g_1}^{rr'}$ and $\gamma_{g_2}^{rr'}$, and so their incentives to enter remain unchanged.

When the critical levels of patent strength for which the generic companies are indifferent between entering or not do *not* exceed $1 - \frac{e}{2}$, monopoly results for all γ-units above $1 - e/2$. In that case, an increase in e still has the same effects as described in Sections 1.3.3 and 1.4. It induces settling companies to collude more for values around $\gamma_{g_1}^{rr'}$ and $\gamma_{g_2}^{rr'}$ and, thus, enhances generic companies' incentives to enter.

A2: The Socially Optimal Level of *e*

Consumer welfare under the rule of reason approach is given by equation (1.5). Differentiating consumer welfare with respect to e yields

$$\begin{aligned}\frac{\partial CW^{rr}}{\partial e} = \; & \frac{\partial \gamma_{g_1}^{rr}}{\partial e}\left(1-\gamma_{g_1}^{rr}-\frac{e}{8}\right)\left(CW^d-CW^m\right) - \frac{\gamma_{g_1}^{rr}}{8}\left(CW^d-CW^m\right) \\ & + \frac{\partial \gamma_{g_2}^{rr}}{\partial e}\left(1-\gamma_{g_2}^{rr}-\frac{e}{8}\right)(1-\lambda)\left(CW^t-CW^d\right) - \frac{\gamma_{g_2}^{rr}}{8}(1-\lambda)\left(CW^t-CW^d\right).\end{aligned} \tag{1.2}$$

For the derivation of the consumer welfare maximizing level of e we assume that the companies compete in quantities. The originator companies' inverse demand function is given by $p_o = 1 - q$, and the generic companies' inverse demand function is given by $p_g = \alpha - q$, with $0.5 \leq \alpha \leq 1$ and $q = q_o + \sum_{i=1}^{2} q_{g_i}$. The differentiation factor α reflects the degree to which the generic products are perceived inferior to the original product by consumers. The lower α is, the more demand originator companies receive in comparison to generic companies when charging the same price.

We can substitute

$$\gamma_{g_2}^{rr} = \frac{(2\alpha-1)(63-19\lambda-2\alpha(27-7\lambda))e - 216(16f_g+(4(1-\alpha)\alpha-1)(1-\lambda))}{216(1-2\alpha)^2(1-\lambda)},$$

$$\gamma_{g_1}^{rr} = \frac{64((1-2\alpha)^2-9f_g)+(4(8-5\alpha)\alpha-11)e}{64(1-2\alpha)^2},$$

$CW^m = 1/8$, $CW^d = ((1+\alpha)/3)^2$, and $CW^t = 3/8\ \alpha^2 + 3/8\ \alpha + 3/32$ into equation (1.2) and solve it for e. This gives us the consumer welfare maximizing level of e:

$$\begin{aligned} e^* = & \big[1728(486(2\alpha-1)^3(12\alpha(1+\alpha)-1) - 9(2\alpha(3383+6484\alpha+648\alpha^2)-721)f_g \\ & + (-108(2\alpha-1)^3(76\alpha+92\alpha^2-7)+(551+2\alpha(4\alpha(8253+13618\alpha))-3729)f_g)\lambda \\ & + 54(2\alpha-1)^3(44\alpha+76\alpha^2-5)\lambda^2)\big] / \\ & \big[(2\alpha-1)(81(65+4\alpha(2125+\alpha(648\alpha(23\alpha-3)-19201))) \\ & - 9(4\alpha(82869+\alpha(8\alpha(29079\alpha-19531)-204777))-22535)\lambda \\ & + 16(14\alpha-19)(94\alpha-35)(44\alpha+76\alpha^2-5)\lambda^2)\big].\end{aligned}$$

It shows, that e^* decreases with generic companies' fixed entry costs (f_g), i.e., with generic companies' incentives to enter. Further, e^* increases with the delay of G_2's entry decision (λ) and with the degree to which generic products are perceived inferior to original products by consumers (α). This means, the higher f_g and λ and the lower α, the more consumer welfare increases when antitrust evaluations of pay-for-delay settlements under the rule of reason are subject to error.

A3: Litigation Costs and Expectational Asymmetry

This appendix analyzes the level of collusion and generic companies' incentives to challenge probabilistic patents under the consideration of (i) litigation and settlement costs ($l \geq \varsigma$) and (ii) originator and generic companies misconceiving their respective likelihoods of success in the litigations ($\gamma_o \lesseqgtr \gamma,\ \gamma_g \lesseqgtr \gamma$). For simplification we assume in this and the following appendix that only one generic company seeks market entry. The underlying economic intuition of our results remains the same. While we assumed before that companies divide any surplus they generate through settlements compared to litigation by the number of companies involved in the settlement talks, we here compute the Nash Bargaining solutions.

Expected litigation profits of originator and generic companies are

$$\pi_o^l = \gamma_o \pi^m + (1-\gamma_o)\pi_o^d - l,$$
$$\pi_g^l = (1-\gamma_g)\pi_g^d - l.$$

The companies' expected profits pursuant to a settlement under the rule of per se illegality (payment $P = 0$) and under the rule of per se legality (payment $P \geq 0$) are

$$\pi_o = t\pi^m + (1-t)\pi_o^d - P - \varsigma,$$
$$\pi_g = (1-t)\pi_g^d + P - \varsigma.$$

The Nash bargaining solution is determined by

$$\max_t \left(\pi_o - \pi_o^l\right)\left(\pi_g - \pi_g^l\right).$$

As negotiated entry date in the Nash bargaining solution we receive

$$t = \frac{\gamma_o + \gamma_g}{2} + \frac{(l-\varsigma)(\pi^m - \pi_o^d - \pi_g^d)}{2(\pi^m - \pi_o^d)\pi_g^d} + \frac{P(\pi^m - \pi_o^d + \pi_g^d)}{2(\pi^m - \pi_o^d)\pi_g^d}. \tag{1.3}$$

In what follows, we analyze the effects of γ_o, γ_g, l and ς on the degree of collusion and the generic companies' incentives to enter.

Under the rule of *per se legality* originator companies are allowed to make payments to generic companies ($P \geq 0$). Since originator companies can compensate generic companies for any delay in entry compared to litigation, the companies settle on entry terms that maximize their joint profits and negotiate about the division of the generated surplus. They choose $t^{pl} = 1$ as generic entry date, independent of their individual perceptions of patent strength and of litigation and settlement costs. Since the degree of collusion remains the same, the welfare implications also remain the same.

Under the rule of *per se illegality* P equals 0 as payments from originator to generic companies are not allowed. Hence, companies settle upon

$$t^{pi} = \frac{\gamma_o + \gamma_g}{2} + \frac{(l - \varsigma)(\pi^m - \pi_o^d - \pi_g^d)}{2(\pi^m - \pi_o^d)\pi_g^d}.$$

Differentiating t^{pi} with respect to $(\gamma_o + \gamma_g)/2$ and $l - \varsigma$ shows that the level of collusion under the rule of per se illegality increases with originator companies being relatively confident and generic companies being relatively unconfident regarding their chances of winning litigation. Further, collusion increases with the difference between litigation and settlement costs.

Generic companies' expected settlement profits are

$$\pi_g^{pi} = \left(1 - \frac{\gamma_o + \gamma_g}{2} - \frac{(l - \varsigma)(\pi^m - \pi_o^d - \pi_g^d)}{2(\pi^m - \pi_o^d)\pi_g^d}\right)\pi_g^d - \varsigma.$$

Thus, the less confident originator companies and the more confident generic companies, the higher are generic companies' incentives to enter. Further, the lower litigation and settlement costs, the higher are generic companies' incentives to enter.

Under the *rule of reason* originator companies are allowed to make payments to generic companies, so that $P \geq 0$. Settling companies choose entry terms that maximize their expected joint profits and then negotiate about the division of the generated surplus. The companies' expected settlement profits are

$$\pi_o^{rr} = \frac{t^{rr} - \gamma_o + e}{2e}\left[\gamma_o \pi^m + (1 - \gamma_o)\pi_o^d\right] + \frac{e - t^{rr} + \gamma_o}{2e}\left[t^{rr}\pi^m + (1 - t^{rr})\pi_o^d - P\right] - \varsigma \quad \text{and}$$
$$\pi_g^{rr} = \frac{t^{rr} - \gamma_g + e}{2e}(1 - \gamma_g)\pi_g^d + \frac{e - t^{rr} + \gamma_g}{2e}\left[(1 - t^{rr})\pi_g^d + P\right] - \varsigma.$$

Thus, they choose as generic entry dates

$$t^{rr} = \frac{\gamma_o(\pi^m - \pi_o^d) - \gamma_g \pi_g^d}{\pi^m - \pi_o^d - \pi_g^d} + \frac{e}{2}.$$

It follows that the degree of collusion increases with the companies' overconfidence but is independent of litigation and settlement costs.

As under the rule of per se illegality, generic companies' incentives to enter decrease with originator companies' confidence and increase with their own confidence. Further, generic companies' incentives to enter decrease with litigation and settlement costs.

From the analysis we can derive the following policy implications. The higher litigation costs relative to settlement costs, the more settling companies collude under the rule of per se illegality, so that in comparison the rule of reason becomes more desirable. If originator companies are overconfident, collusion under the rule of per se illegality and under the rule of reason increases. Under the rule of reason collusion increases by more but through prohibitions of settlements neutral outcomes can still be achieved.

Further, the higher litigation costs relative to settlement costs, the lower are generic companies' incentives to challenge probabilistic patents under the rule of per se illegality and under the rule of reason. It follows, that incentivization as achieved under the rule of reason becomes more desirable. The same applies if originator companies become more overconfident or generic companies less confident. In that case, generic companies' bargaining position is weakened, which derogates their incentives to enter.

A4: Risk Aversion

Contrary to settlement, litigation poses a risk to companies. Generic companies risk early versus late flow of profits and originator companies risk early versus late loss of monopoly profits. Companies that have more to risk or that are more risk averse tend to accept less favorable settlement terms in order to avoid that settlement fails. Absent any compensation generic companies, which prefer earlier entry to later, might be willing to postpone entry somewhat past the expected entry date under litigation if the postponement is not so protracted that the cost to it in lost profits is more than what is saved in avoided risk. Similarly, originator companies, which prefer later entry to earlier, might be willing to accelerate entry relative to the expected generic entry date under litigation. Because risk aversion has the same effect as an increase in litigation costs we treat the cost of bearing risk as a 'risk premium'.

In the US, generic companies have not made infringing sales that would give rise to claims for damages or incurred production costs when triggering a patent dispute. Hence, their litigation risk may be rather small.[37] They only may risk bankruptcy when litigation takes too long. Originator companies face potentially larger consequences if they lose litigation as their profits would drop dramatically. The bargaining strength of companies that bear a higher risk is weakened.

[37] In the EU, generic companies either enter at risk or await a declaratory judgment (after having indicated their intention to enter and received a notice by the originator company that it intends to sue the generic companies for infringement in case of entry). The prerequisites for declaratory judgments differ among EU states.

Under the rule of per se illegality (P=0) and under the rule of reason (P≥0) the assumption of $l_o \neq l_g$ takes the following effects. Expected litigation profits are

$$\pi_o^l = \gamma\pi^m + (1-\gamma)\pi_o^d - l_o,$$
$$\pi_g^l = (1-\gamma)\pi_g^d - l_g.$$

The companies' expected profits pursuant to a settlement are

$$\pi_o = t\pi^m + (1-t)\pi_o^d - P - \varsigma,$$
$$\pi_g = (1-t)\pi_g^d + P - \varsigma.$$

As Nash bargaining solution we receive

$$t = \gamma - \frac{l_o}{2(\pi^m - \pi_o^d)} + \frac{l_g}{2\pi_g^d} - \frac{\varsigma(\pi^m - \pi_o^d - \pi_g^d)}{2(\pi^m - \pi_o^d)\pi_g^d} + \frac{P(\pi^m - \pi_o^d + \pi_g^d)}{2(\pi^m - \pi_o^d)\pi_g^d}. \tag{1.4}$$

It follows that under the rule of per se illegality settlements become more procompetitive when originator companies' litigation costs increase, and more anticompetitive when generic companies' litigation costs increase. In line with the results of the previous appendix, collusion under the rule of per se illegality decreases compared to under the rule of reason when the originator companies become more risk averse. This means, the rule of per se illegality becomes more favorable compared to the rule of reason. The opposite is true, and the rule of reason becomes more favorable, when generic companies become more risk averse.

As generic companies' bargaining position is strengthened when originator companies become more risk averse, their incentives to enter improve. Incentivization, as achieved under the rule of reason, becomes less important. Thus, when originator companies become more risk averse the rule of per se illegality prevails. Again, the opposite is true and the rule of reason becomes more favorable when generic companies become more risk averse.

1.9 Appendix B: Proofs

Proof of Proposition 1.1

In order to judge whether the rule of reason or the rule of per se illegality is preferable we need to analyze whether consumer welfare is higher under the rule of reason when $e = 0$ or when $e > 0$. We therefore look at whether a marginal increase in e, given $e = 0$ in the first place, has positive welfare implications.

Consumer welfare under the rule of reason is

$$\begin{aligned} CW^{rr} = & \int_0^{\gamma_{g_2}^{rr}(e)} \left[\frac{3}{4}\left[\gamma CW^m + (1-\gamma)\left[\lambda CW^d + (1-\lambda)CW^t\right]\right]\right. \\ & \left. + \frac{1}{4}\left[\left(\gamma + \frac{e}{2}\right)CW^m + \left(1-\gamma-\frac{e}{2}\right)\left[\lambda CW^d + (1-\lambda)CW^t\right]\right]\right] d\gamma \\ & + \int_{\gamma_{g_2}^{rr}(e)}^{\gamma_{g_1}^{rr}(e)} \left[\frac{3}{4}\left[\gamma CW^m + (1-\gamma)CW^d\right] + \frac{1}{4}\left[\left(\gamma+\frac{e}{2}\right)CW^m + \left(1-\gamma-\frac{e}{2}\right)CW^d\right]\right] d\gamma \\ & + \int_{\gamma_{g_1}^{rr}(e)}^{1} CW^m \, d\gamma \\ = & \left[\left(1-\frac{e}{8}\right)\gamma_{g_2}^{rr}(e) - \frac{\gamma_{g_2}^{rr}(e)^2}{2}\right](1-\lambda)(CW^t - CW^d) \\ & + \left[\left(1-\frac{e}{8}\right)\gamma_{g_1}^{rr}(e) - \frac{\gamma_{g_1}^{rr}(e)^2}{2}\right](CW^d - CW^m) + CW^m. \end{aligned} \tag{1.5}$$

Thus, the effect of a marginal increase in e on consumer welfare is

$$\begin{aligned} \frac{\partial CW^{rr}}{\partial e} = & \underbrace{\frac{\partial \gamma_{g_1}^{rr}}{\partial e}\left(1-\gamma_{g_1}^{rr}-\frac{e}{8}\right)\left(CW^d - CW^m\right)}_{\text{Incentive Effect}-G_1} - \underbrace{\frac{\gamma_{g_1}^{rr}}{8}\left(CW^d - CW^m\right)}_{\text{Entry Delay Effect}-G_1} \\ & + \underbrace{\frac{\partial \gamma_{g_2}^{rr}}{\partial e}\left(1-\gamma_{g_2}^{rr}-\frac{e}{8}\right)(1-\lambda)\left(CW^t - CW^d\right)}_{\text{Incentive Effect}-G_2} - \underbrace{\frac{\gamma_{g_2}^{rr}}{8}(1-\lambda)\left(CW^t - CW^d\right)}_{\text{Entry Delay Effect}-G_2}. \end{aligned} \tag{1.6}$$

Equation (1.6) shows, when e (the imprecision of antitrust enforcement under the rule of reason) increases, settling companies agree upon later generic entry dates than under the rule of per se illegality, i.e., they collude more. Under each settlement competition is restrained, which leads to consumer welfare losses. As shown by equation (1.6), for all values of patent strength for which settlements would also result absent an increase in e, that is for all intramarginal $\gamma_{g_1}^{rr}$- and $\gamma_{g_2}^{rr}$-units, consumer welfare decreases by $1/8\left(CW^d - CW^m\right)$ and $1/8\,(1-\lambda)\left(CW^t - CW^d\right)$, respectively.

Through collusion the companies' expected settlement profits increase and with it the critical levels of patent strength for which the generic companies are indifferent between entering or not ($\gamma_{g_1}^{rr}$ and $\gamma_{g_2}^{rr}$). The generic companies find it profitable to also contest patents of higher strength. This affects consumer welfare positively when collusion under the additional settlements is limited. As shown by equation (1.6), a marginal increase in e leads to an increase in $\gamma_{g_1}^{rr}$ by $\partial\gamma_{g_1}^{rr}/\partial e$, meaning first generic companies enter for additional $\partial\gamma_{g_1}^{rr}/\partial e$ units. For each of these units consumer welfare increases by $\left(1-\gamma_{g_1}^{rr}-e/8\right)\left(CW^d-CW^m\right)$. A marginal increase in e also leads to an increase in $\gamma_{g_2}^{rr}$ by $\partial\gamma_{g_2}^{rr}/\partial e$. Thus, second generic companies enter for additional $\partial\gamma_{g_2}^{rr}/\partial e$ units. For each of these units consumer welfare increases by $\left(1-\gamma_{g_2}^{rr}-e/8\right)(1-\lambda)\left(CW^t-CW^d\right)$.

Thus, the imperfection of antitrust enforcement under the rule of reason has two countervailing effects on consumer welfare. It induces settling parties to delay entry, affecting consumer welfare negatively. At the same time, it enhances generic companies' incentives to challenge probabilistic patents, affecting consumer welfare positively.

Next, we analyze under which conditions the negative entry delay effect of a marginal increase in e is outweighed by the positive incentive effect, implying that the rule of reason outperforms the rule of per se illegality. This is the case when the marginal effect of an increase in e on consumer welfare is positive, given $e=0$ in the first place. Setting the first line of equation (1.6) equal to zero, inserting $e=0$, and solving for $\gamma_{g_1}^{rr}$ yields

$$\gamma_{g_1}^{rr'} \equiv \frac{\Delta_1}{\Delta_1+2\pi_g^d},$$

where $\Delta_1=\pi^m-\pi_o^d-\pi_g^d$. Accordingly, setting the second line of equation (1.6) equal to zero, inserting $e=0$, and solving for $\gamma_{g_2}^{rr}$ yields

$$\gamma_{g_2}^{rr'} \equiv \frac{\Delta_2}{\Delta_2+3(1-\lambda)\pi_g^t},$$

where $\Delta_2=\pi^m-\lambda(\pi_o^d+\pi_g^d)-(1-\lambda)(\pi_o^t+2\pi_g^t)$. It follows that consumer welfare is higher under the rule of reason than under the rule of per se illegality when $\gamma_{g_1}^{rr}<\gamma_{g_1}^{rr'}$ and $\gamma_{g_2}^{rr}<\gamma_{g_2}^{rr'}$.

We can further specify these conditions by substituting in $\gamma_{g_1}^{rr}=1-f_g/\pi_g^d$ and $\gamma_{g_2}^{rr}=1-f_g/(1-\lambda)\pi_g^t$. We obtain that the rule of reason yields higher consumer welfare than the rule of per se illegality if

$$f_{g_1} > \frac{2\pi_g^{d^2}}{2\pi_g^d+\Delta_1} \quad \text{and } f_{g_2} > \frac{3\big((1-\lambda)\pi_g^t\big)^2}{(1-\lambda)\pi_g^t+\Delta_2}. \quad \blacksquare$$

Proof of Proposition 1.2

The proof that the HWA provisions have anticompetitive effects proceeds in two steps. We first show that the HWA provisions alleviate second generic companies' incentives to enter ($\gamma_{g2}^{hwa} < \gamma_{g_2}$). We then show that the HWA provisions do not provide first generic companies additional incentives to enter ($\gamma_{g1}^{hwa} \equiv \gamma_{g_1}$).

When the HWA provisions apply, G_2's expected entry date under litigation postpones to $t_{g2}^{l|hwa} = \gamma + (1-\gamma)(\lambda+\nu)$, where ν denotes the 180 days marketing exclusivity period during which the FDA may not approve any subsequent ANDA. When G_2 wants to enter and G_1 settled with O before, G_2 has to file an ANDA (para. IV), whereupon it can be sued by O for infringement. With probability γ G_2 would then loose litigation and enter at time $t = 1$, whereas with probability $(1-\gamma)$ it would win litigation, trigger the 180 days exclusivity period at time $t = \lambda$, and enter at time $t = \lambda + \nu$.

Because G_2's expected entry date under litigation postpones, it has less negotiation power and therefore settles with O on later entry dates. Under the rule of reason O and G_2 agree upon $t_{g2}^{rr|hwa} = \gamma + e/2 + (1 - \gamma - e/2)(\lambda + \nu)$, whereas under the rule of per se illegality they agree upon $t_{g2}^{pi|hwa} = \gamma + (1 - \gamma)(\lambda + \nu)$.

Under the *rule of reason* not only G_2's entry date postpones, the companies also generate a lower surplus when G_2 finds it profitable to enter:

$$s_1^{rr|hwa} = \frac{e}{2}\left[\pi^m - (\lambda + \nu)(\pi_o^d + \pi_g^d) - (1 - \lambda - \nu)(\pi_o^t + 2\pi_g^t)\right].$$

Because G_2's entry date postpones and a lower settlement surplus is generated when it enters, its incentives to enter deplete. That is, the critical level of patent strength, for which G_2 is indifferent between entering or not, decreases to

$$\gamma_{g2}^{rr|hwa} = 1 - \frac{f_g - \frac{s_1^{rr|hwa}}{12}}{(1 - \lambda - \nu)\pi_g^t} < \gamma_{g2}^{rr}.$$

Under the *rule of per se illegality*, settling companies still generate no surplus, i.e., $s^{pi|hwa} \equiv s^{pi} = 0$. But since G_2's entry date postpones, the critical level of patent strength, for which G_2 is indifferent between entering or not, decreases to

$$\gamma_{g2}^{pi|hwa} = 1 - \frac{f_g - \frac{s^{pi|hwa}}{3}}{(1 - \lambda - \nu)\pi_g^t} < \gamma_{g2}^{pi}.$$

Next, we analyze whether the HWA provisions provide G_1 incentives to challenge patents that it would not challenge otherwise, i.e., patents of strength $\gamma \in (\gamma_{g_1}, 1]$. This would be the case if G_1's expected settlement profits increased due to the HWA provisions for patents of strength $\gamma \in (\gamma_{g_1}, 1]$. Indeed, G_1's expected settlement increase due to the HWA provisions for patents of strength $\gamma \in (0, \gamma_{g2}]$, as for these values competition between G_1 and G_2 is restrained as either G_2 enters at a later point in time (if $\gamma \in [0, \gamma_{g2}^{hwa}]$), or G_2 does not enter at all (if $\gamma \in (\gamma_{g2}^{hwa}, \gamma_{g2}]$). However, since $\gamma_{g_2} < \gamma_{g_1}$ G_1's expected settlement do not increase for patents of strength $\gamma \in (\gamma_{g_1}, 1]$ as for these patents G_2 would also, even absent the HWA provisions, not enter. An exclusivity right that restricts competitions between G_1 and G_2 is effectless here, as competition would not arise anyway. Therefore, G_1's incentives to challenge additional patents of strength $\gamma \in (\gamma_{g_1}, 1]$ do not improve, that is, $\gamma_{g_1}^{hwa} \equiv \gamma_{g_1}$.

Note that under the rule of per se legality basically the same mechanisms apply. But an increase in generic companies' incentives to enter would not have a positive effect on consumer welfare anyway as collusion under this rule is maximal. ■

Chapter 2

Can Naked Exclusion Be Procompetitive?[0]

In many recent antitrust cases incumbent upstream firms were alleged of having used exclusive contracts to deter potentially more efficient entrants, thereby harming consumers.[1] In these cases courts need to balance anticompetitive effects caused by entry deterrence or increased wholesale prices against potential efficiency gains created through exclusive contracting within the vertical production chain.

We point out that exclusive contracting can have *procompetitive* effects even if no efficiency gains are generated, provided downstream competition is moderate and downstream firms can breach exclusive contracts. The intuition is the following. Suppose the downstream firms signed the exclusive contract with the incumbent. In that case, the entrant may nevertheless find it profitable to enter since it can induce the downstream firms to breach the contract. Because breaching downstream firms have to pay expectation damages to the incumbent, they only breach when they can obtain sufficiently large downstream profits. Therefore, the entrant needs to sell its input at a relatively low wholesale price. Using a framework developed by Simpson and Wickelgren (2007), we show that for moderate degrees of downstream competition this mechanism leads to lower final consumer prices than without exclusive contracting, and therefore to a rise in welfare.

Our result stands in stark contrast to the previous literature, which asserts that exclusive contracting has anticompetitive or at best neutral effects if no efficiencies are

[0]This chapter is based on joint work with Prof. Dr. Markus Reisinger.

[1]Recent examples are *United States vs. Transitions Optical*, *United States vs. Dentsply*, *Pernod Ricard and Campbell Distillers vs. Bacardi-Martini*, *Langnese-Iglo vs. European Commission*, and *United States vs. Microsoft*.

generated. As is well known, *Chicago School* scholars (e.g., Posner, 1976, and Bork, 1978) argue for a neutral effect. They assume that downstream buyers are independent monopolists (or final consumers). In this situation, where downstream firms do not compete, the incumbent's gain in profit through entry deterrence is lower than the downstream firms' loss in profit. Therefore, the incumbent is unable to compensate the downstream firms for signing exclusive contracts, given no efficiencies are generated. Rasmusen, Ramseyer, and Wiley (1991) and Segal and Whinston (2000) challenge this argument, pointing out that the entrant may not be able to reach minimum efficient scale when selling only to a fraction of buyers, implying that downstream firms exert negative externalities on each other when signing exclusive contracts. The incumbent can induce downstream firms to sign exclusive contracts by exploiting these externalities.

Fumagalli and Motta (2006) analyze the case in which downstream buyers are not independent monopolists but perfect Bertrand competitors and argue for a neutral effect. With perfect downstream competition the entrant needs to sell only to a single downstream firm to reach minimum efficient scale, which removes the negative externalities that signing downstream firms exert on the other. To bring out this effect Fumagalli and Motta (2006) assume that downstream firms face a fixed fee of being active in the downstream market. When all but one firm signed the contract, the firms that signed face a higher wholesale price and therefore stay inactive, which enables the single firm that did not sign to earn monopoly profits. As a consequence, each downstream firm demands the monopoly profit as compensation for signing, so that exclusive contracting becomes too costly for the incumbent. Several follow-up papers show that a different picture emerges once the assumption on the fixed fee of being active is dropped.[2] These papers show that it becomes easier for the incumbent to induce downstream firms to sign if downstream competition increases.[3] The reason is that signed downstream firms stay active, thereby exerting competitive pressure on downstream firms that did not sign. This limits the profits that downstream firms can obtain by not signing. Thus, the compensation that the incumbent needs to offer for signing decreases.[4]

[2] See, for example, Abito and Wright (2008), Wright (2008, 2009), and Kitamura (2010).

[3] A similar argument is put forward in earlier works by Stefanidis (1998), Yong (1999), and Simpson and Wickelgren (2001). In these papers, though, the authors assume that the incumbent can commit to a certain wholesale price when offering exclusive contracts.

[4] These results refer to the case in which the entrant's efficiency advantage is non-drastic. If instead the entrant's cost advantage was drastic, the incumbent would be unable to induce the downstream firms to sign, independent of whether the downstream firms have to pay a fixed fee for being active or not.

An important limitation of these papers is the assumption that once downstream firms have signed exclusive contracts, they cannot breach them later. If all firms have signed, this inevitably leads to entry deterrence. Common law, however, provides each party to a contract the opportunity to breach by paying expectation damages to the injured party. While in some situations breach of contract may indeed be prohibitively costly due to reputational reasons or high litigation costs, it seems unreasonable to assume generally that contract breach is not feasible. Simpson and Wickelgren (2007) (henceforth SW) provide an insightful model in which they incorporate the possibility of contract breach. They analyze the cases in which downstream firms are (i) independent monopolists or (ii) (almost) perfect Bertrand competitors. They find that only in the latter case the incumbent is able to induce downstream firms to sign exclusive contracts. Signing, however, does not lead to entry deterrence due to the possibility of contract breach. Nevertheless the effects of exclusive contracts are anticompetitive when downstream firms are perfect Bertrand competitors because the entrant induces just a single downstream firm to breach, and this firm monopolizes the downstream market, which results in higher final consumer prices. In accordance with the previous literature, SW therefore argue that absent efficiency gains the effect of exclusive contracting is anticompetitive or at best neutral.

We extend the analysis by SW to account for general degrees of downstream competition. Particularly moderate degrees of competition are relevant and therefore important to consider as products are often physically differentiated and consumers often have different preferences over products. As a central result we find that for such moderate degrees of downstream competition exclusive contracting can have *procompetitive* effects, even if no efficiencies are generated. Intuitively, when downstream competition is moderate, the entrant induces not just a single downstream firm but *both* downstream firms to breach because it can receive sizable profits from both these firms. When breaching, the firms have to pay expectation damages to the incumbent. Thus, in order to render breaching profitable the entrant must set its wholesale price sufficiently low. In particular, the wholesale price it needs to set lies below the one that the upstream firms would set absent exclusive contracting. As a consequence, final consumer prices fall and total welfare rises.[5]

Downstream firms sign an exclusive contract although they may later breach it because *not* signing implies that their downstream profits will be small due to competi-

[5] As shown by Mathewson and Winter (1987) and Bernheim and Whinston (1998) if two *incumbent* upstream firms compete for exclusive contracts, the effect can be procompetitive as well. However, the mechanisms leading to this effect are very different from the one identified in our analysis, in which exclusive dealing is used for entry deterrence reasons.

tion. When both sign, the incumbent obtains monopoly profits, even if the downstream firms breach, as it is subject to expectation damages. These monopoly profits it can partly use to compensate the downstream firms for signing. Thus, exclusive deals enable contracting parties to extract some of the entrants' rents via expectation damages.

The paper proceeds as follows. Section 2.2 sets out the model. In Section 2.3 we present our result with a general demand function. Section 2.4 provides an application with a linear demand function, and Section 2.5 concludes.

2.2 The Model

In this section, we outline the model which follows SW. Everything described below is common knowledge to all agents. We analyze an industry with an upstream and a downstream market. In the upstream market an incumbent firm I and a potential entrant E produce a homogeneous input good. In the downstream market two differentiated firms i and j process the input good at a one-to-one technology and compete in prices for final consumers.

For tractability reasons we assume that downstream firms i and j are symmetric. Downstream firm i's demand function when setting a price p_i and when the rival sets a price p_j is given by $D(p_i, p_{j;}\gamma)$, with $\partial D(p_i, p_{j;}\gamma)/\partial p_i < 0$, $\partial D(p_i, p_{j;}\gamma)/\partial p_j \geq 0$ and $|\partial D(p_i, p_{j;}\gamma)/\partial p_i| \geq |\partial D(p_i, p_{j;}\gamma)/\partial p_j|$. A downstream firm's demand is falling in its own price, it is rising in its rival's price, and the absolute effect of its own price is larger than the effect of its rival's price. In this demand function, $\gamma \in [0, 1)$ is a parameter representing the degree of downstream competition or product differentiation, that is, $\partial D(p_i, p_{j;}\gamma)/\partial p_i < 0$ is weakly decreasing and $\partial D(p_i, p_{j;}\gamma)/\partial p_j \geq 0$ is strictly increasing in γ. For $\gamma = 0$, the two products are independent, implying that each downstream firm is a monopolist, that is, $\partial D(p_i, p_{j;}0)/\partial p_j = 0$ and $|\partial D(p_i, p_{j;}\gamma)/\partial p_i|$ is minimal. As $\gamma \to 1$, the two products become perfect substitutes, implying perfect Bertrand competition, that is, $\lim_{\gamma \to 1} \partial D(p_i, p_{j;}\gamma)/\partial p_j = \infty$ and $\lim_{\gamma \to 1} \partial D(p_i, p_{j;}\gamma)/\partial p_i = -\infty$ as long as both demands are strictly positive. We impose two technical assumptions, $\partial^2 D(p_i, p_{j;}\gamma)/\partial p_i^2 \leq 0$ (or not too positive) and $\partial^2 D(p_i, p_{j;}\gamma)/(\partial p_i \partial p_j) \geq 0$, which guarantee that each downstream firm's demand function is concave and that equilibrium prices are strategic complements, i.e., $\partial p_i/\partial p_j > 0$. They also ensure that firm i's profit is increasing in the cost of firm j.

The timing of the game is as follows (see also Table 2.1). In the first stage, I makes simultaneous nondiscriminatory exclusive contract offers to the downstream firms.[6] An

[6]Our results would not change if we assumed that I makes sequential or discriminatory offers.

exclusive contract is a compensation x from I to the downstream firms in exchange for the downstream firms' commitment to purchase exclusively from I. After observing these offers, the downstream firms simultaneously decide whether to accept or reject them. In the second stage, E decides on entry. In stage 3.1, active upstream firms set wholesale prices to each available downstream firm. I is able to discriminate between those downstream firms that have signed the exclusive contract (captive downstream firms) and those who have not (free downstream firms). It offers a wholesale price w_c to captive downstream firms and a wholesale price w_f to free downstream firms. E offers a wholesale price w_e to free downstream firms.[7] Captive downstream firms can become free by breaching and paying expectation damages to I in stage 3.2. In accordance with common law I's expectation damages are based on its lost profits. It needs to be restored to the position it would have been in had the contract been performed.[8] We assume, if both downstream firms breach, each one pays half of the expectation damages. In stage 3.3, I and E produce the input good. Free downstream firms purchase the input good from E if $w_e \leq w_f$ and from I if $w_e > w_f$. Captive downstream firms purchase from I at w_c.[9] Downstream firms compete for consumers by setting prices p_i and p_j.

Table 2.1: Timeline

Stage 1	Stage 2	Stage 3.1	Stage 3.2	Stage 3.3
I offers excl. contract	E enters or not	I sets prices w_f, w_c	i,j can breach	i,j buy input
i,j accept or reject		E sets price w_e		i,j compete

To simplify the notation, we denote the equilibrium downstream price vector $\mathbf{p}(w_i, w_j) = [p_i(w_i, w_j), p_j(w_j, w_i)]^{\mathrm{T}}$ when needed as an argument in firm i's demand and $\mathbf{p}(w_j, w_i) = [p_j(w_j, w_i), p_i(w_i, w_j)]^{\mathrm{T}}$ when needed as an argument in firm j's demand.

We assume that E is more efficient than I, enjoying lower marginal costs of production ($c_E < c_I$). To enter E has to pay a sunk cost f. Throughout the analysis we assume that E is sufficiently more efficient than I that it can cover this sunk cost when selling to both downstream firms at a wholesale price of w'_E, where w'_E is the wholesale price that E must charge to induce one downstream firm to breach provided the rival downstream

[7] In accordance with SW, we restrict our attention to the case of linear wholesale prices. For a brief discussion on two-part tariffs see the conclusion.

[8] As SW we consider the situation in which breaching downstream firms are subject to expectation damages. The main mechanism driving our result would also be at work in case of liquidated damages, which are considered by Aghion and Bolton (1987), Innes and Sexton (1994), and Spier and Whinston (1995). For a general discussion on the difference between expectation and liquidated damages see Brodley and Ma (1993).

[9] Our results are invariant to renegotiations, that is, even if I could change w_c after the downstream firms decided whether to breach, all of our results would hold.

firm does not breach.[10] Specifically, we impose that $2(w'_E - c_E)D(\mathbf{p}(w'_E, w'_E);\gamma) > f$, where $D(\mathbf{p}(w'_E, w'_E);\gamma)$ denotes a downstream firm's demand given both downstream firms face an input price of w'_E and set their downstream prices accordingly.[11] The underlying economics driving our main result are not affected by this assumption.

To avoid the epsilon notation on prices and compensations, we assume that the downstream firms sign the exclusive contract when they are indifferent between signing or not, they breach the exclusive contract when they are indifferent between breaching or not, and they buy from E when they are indifferent between buying from E or I.

Our equilibrium concept is subgame perfection with the additional refinement that if multiple equilibria arise, the downstream firms play the equilibrium that is Pareto dominant from their perspective. This assumption is necessary because in stage 3.2 of the game multiple equilibria can arise in which either both downstream firms or none of them breaches the exclusive contract.

Finally, we assume that $D(\mathbf{p}(c_I, c_I);\gamma)(p(c_I, c_I) - c_I) \geq D(\mathbf{p}(w_c, c_I);\gamma)(p(w_c, c_I) - c_I)$, where w_c solves the maximization problem $\max_w D(\mathbf{p}(w, c_I);\gamma)(w - c_I)$. This assumption implies that a downstream firm is better off when it competes in the downstream market on the basis of its true costs c_I than on costs $w_c > c_I$, given that the rival faces costs c_I.[12] This assumption simplifies the proofs of the arguments but is not crucial for our main effect to hold. It is satisfied for many commonly used demand functions such as the linear one considered in Section 2.4, CES, logit or Hotelling.

2.3 The Effect in General Form

We first look at the equilibrium in the downstream market when downstream firm i faces a wholesale price w_i while downstream firm j faces a wholesale price w_j. Firm i's profit function is

$$\pi_i = D(p_i, p_j;\gamma)(p_i - w_i).$$

[10] The explicit definition of w'_E is given in equation (2.4) in the appendix.

[11] A similar assumption is imposed by SW who assume $2(c_I - c_E)D(\mathbf{p}(c_I, c_I);\gamma) > f$. Our assumption adjusts the one by SW to the case of differentiated products.

[12] In general, it is well-known from the literature on vertical restraints (e.g., Bonanno and Vickers, 1985) that competing on the basis of higher costs than the true costs can be beneficial for a firm as it induces the rival firm to react less aggressively. This argument, however, relies on the fact that, at the true input costs of firm D_i, a change in these costs has only a second-order effect on the optimal choice of firm D_i but a first-order effect on the choice of the rival firm. By contrast, in our case w_c is set according to a different maximization problem, implying that it is biased upwards to a large extent.

The first-order conditions are given by

$$\frac{\partial D(p_i, p_j;\gamma)}{\partial p_i}(p_i - w_i) + D(p_i, p_j;\gamma) = 0, \quad i \neq j, \quad i,j = 1,2. \tag{2.1}$$

These first-order conditions characterize the equilibrium of the downstream game. Since we assume that $|\partial D(p_i, p_j;\gamma)/\partial p_i| > |\partial D(p_i, p_j;\gamma)/\partial p_j|$, the equilibrium is unique. Due to the other assumptions, we have the natural properties that in equilibrium $dp_i/dw_i > 0$ and $dp_i/dw_j > 0$.

Since $\gamma \to 1$ implies $\partial D(p_i, p_j)/\partial p_i \to -\infty$, we obtain that profits become zero when products are undifferentiated. By contrast, when $\gamma = 0$, implying that the downstream firms are independent monopolists, profits are largest.

In the following, we assess for which levels of downstream competition the incumbent can profitably make use of exclusive contracts.

Lemma 2.1 *If downstream competition is sufficiently strong, i.e., $\gamma \geq \widehat{\gamma}$, in any equilibrium both downstream firms sign the exclusive contract with the incumbent.*

Lemma 2.1 shows that the incumbent can profitably offer positive payments to the downstream firms for signing exclusive contracts if downstream competition is sufficiently intense. The intuition for this result is the following. The downstream firms sign the exclusive contract when the incumbent offers them as compensation at least the profit they can obtain when rejecting it. Exclusive contracts enable the incumbent to earn the monopoly profit since it obtains the right to expectation damages in case of contract breach. When downstream competition is intense, double marginalization is only a minor problem, implying that the incumbent's monopoly profit is relatively high. At the same time, the profit that the downstream firms can obtain by rejecting the contract, i.e., the compensation that the incumbent has to offer them for signing, is relatively low. Therefore, the incumbent is able to profitably make use of exclusive contracts, inducing the downstream firms to sign, when downstream competition is sufficiently strong.

The result of Lemma 2.1 is close to that by SW who show that the incumbent can induce both downstream firms to sign exclusive contracts if downstream competition is almost perfect. We now move one step further and determine the optimal pricing decision of the entrant given both downstream firms have signed the contract.

Lemma 2.2 *If $\underline{\gamma} \leq \gamma \leq \overline{\gamma}$, with $\underline{\gamma} \geq \widehat{\gamma}$, both signed downstream firms breach the exclusive contract and buy from the entrant at a wholesale price $w_e < c_I$.*

This Lemma shows that both signed downstream firms are induced to breach at a wholesale price below c_I if the degree of downstream competition is in an intermediate range. The intuition for this result is the following. Breaching downstream firms have to pay expectation damages to the incumbent. Thus, they only breach when they can obtain sufficiently high profits in the downstream market. The profits they can obtain by breaching decrease when the other downstream firm also breaches. The negative externality that the downstream firms exert on each other when breaching increases with the degree of downstream competition. Consequently, the entrant must reduce its wholesale price with increasing degree of downstream competition when it wants to induce breach of both downstream firms. In particular, when the degree of downstream competition lies above a certain threshold, i.e., if $\gamma \geq \underline{\gamma}$, the entrant needs to price below the Bertrand duopoly price c_I to induce breach of both downstream firms. If the degree of downstream competition increases further and $\gamma > \overline{\gamma}$, the entrant needed to set such a low wholesale price to ensure breach of both downstream firms that it finds it more profitable to set a higher wholesale price and induce breach of just one downstream firm. When inducing just one downstream firm to breach it can keep its wholesale price at a higher level but receives less demand.

From Lemma 2.2 we know, if $\underline{\gamma} \leq \gamma \leq \overline{\gamma}$ both signed downstream firms breach the exclusive contract and face a wholesale price $w_e < c_I$. If exclusive contracts were prohibited, they would face a wholesale price equal to c_I. This gives us the main result.

Proposition 2.1 *If $\underline{\gamma} \leq \gamma \leq \overline{\gamma}$, exclusive contracts have procompetitive effects.*

The proposition shows that naked exclusion can indeed be procompetitive. For $\underline{\gamma} \leq \gamma \leq \overline{\gamma}$ the entrant induces both signed downstream firms to breach the exclusive contracts by setting its wholesale price below c_I. Because downstream firms obtain the input good at a lower cost, they set lower prices to final consumers, which then leads to an increase in consumer surplus and total welfare.

If products are sufficiently differentiated, i.e., if $\gamma < \underline{\gamma}$, the possibility of exclusive contracting is competitively innocuous because either the incumbent cannot pay the downstream firms to accept exclusive contracts or the entrant can induce both downstream firms to breach at a price equal to c_I. In both cases welfare is unaffected by the possibility of exclusive contracting.

The effect of exclusive contracting on welfare is unclear for $\gamma > \overline{\gamma}$. Compared to the situation in which exclusive dealing is not possible, the free downstream firm faces a lower wholesale price but the captive downstream firm a higher wholesale price. Thus, the downstream firms might set higher prices to final consumers than they would set absent exclusive contracting, in which case exclusive contracting would have anticompetitive effects. As shown by SW, the anticompetitive effect prevails if $\gamma \to 1$.

Downstream firms sign exclusive contracts although they may later breach them because by doing so they can increase their joint profits with the incumbent. When both downstream firms sign, the incumbent earns monopoly profits even if the downstream firms later breach as it becomes subject to expectation damages. These monopoly profits the incumbent can partly use to compensate the downstream firms for signing. This applies whenever downstream competition is relatively strong. In that case the incumbent's monopoly profits when both downstream firms sign are particularly high as double marginalization is reduced and the downstream firms' profits when *not* signing, i.e., the compensations the incumbent needs to offer for signing, are particularly low. Thus, even if the downstream firms later breach, exclusive contracts enable the incumbent to extract some of the entrants' rents via expectation damages.

Our analysis shows that for a general class of demand functions naked exclusion can have procompetitive effects. Because the analysis so far does not allow us to draw conclusions on how large the specific regions for γ are, we provide a linear demand example in the next section.

2.4 An Application with Linear Demand

In this section, we show that with a commonly used linear demand function exclusive contracting has procompetitive effects in a sizable range, in which the degree of product differentiation between the downstream firms is moderate.

We assume that demand is defined by the standard representative consumer model (see e.g., Vives, 1999), where a consumer's utility is given by

$$U(q_i, q_j) = (q_i + q_j) - \frac{(q_i^2 + q_j^2) + 2\gamma q_i q_j}{2} + v.$$

Here, q_i is the amount of consumption from downstream firm i and v is the consumption of an outside good whose marginal utility is normalized to one. The parameter $\gamma \in [0, 1)$ again reflects the degree of product differentiation between the downstream firms. For $\gamma = 0$ the two goods are independent, while for $\gamma \to 1$ they become perfect substitutes. If consumers maximize this utility subject to an income constraint, the inverse demand of downstream firm i becomes $p_i = 1 - q_i - \gamma q_j$. It is straightforward to derive the Nash equilibrium in the downstream market by maximizing the downstream firms' profit for given wholesale prices w_i and w_j. Downstream firm i's price in this equilibrium is given by

$$p_i = \frac{(1-\gamma)(2+\gamma) + 2w_i + \gamma w_j}{(2-\gamma)(2+\gamma)}.$$

Both downstream firms receive positive demand only if their prices are sufficiently close to each other. If their prices strongly diverge, the higher priced downstream firm receives no demand, while the lower priced downstream firm captures the entire market. Specifically, downstream firm i's demand function is given by

$$q_i = \begin{cases} 1 - p_i & \text{if} \quad 0 < p_i \leq \frac{-1+\gamma+p_j}{\gamma}, \\ \frac{1-\gamma-p_i+\gamma p_j}{(1-\gamma^2)} & \text{if} \quad \frac{-1+\gamma+p_j}{\gamma} < p_i < 1-\gamma+\gamma p_j, \\ 0 & \text{if} \quad 1-\gamma+\gamma p_j \leq p_i. \end{cases}$$

We measure the entrant's efficiency advantage by θ, where $c_I = \theta w^m(c_E) + (1-\theta)c_E$.[13] Here, $w^m(c_E)$ denotes the monopoly wholesale price when a firm's marginal cost is c_E, i.e., $w^m(c_E) = (1+c_E)/2$. Hence, $\theta = 0$ implies that the entrant has no efficiency advantage, while $\theta = 1$ implies that the entrant's efficiency advantage is just drastic. To simplify the exposition we assume that $\theta \geq 0.121$ and $f = 0$.[14] The first assumption is the equivalent to $c_E < w_E'$ in the general demand case. Together with the second assumption it rules out the case in which entry is not profitable for E.

We first assess for which degrees of downstream competition the incumbent can profitably make use of exclusive contracts.

Lemma 2.1′ *Both downstream firms sign the exclusive contracts if and only if* $\gamma \in [0.5, 1)$.

Next, we determine the wholesale prices that the upstream firms set in equilibrium.

Lemma 2.2′

- *The entrant sells to both downstream firms if (i)* $\gamma \in [0, 0.714)$, *or (ii)* $\gamma \in [0.714, 0.899) \wedge \theta \geq \hat{\theta}(\gamma)$.

- *The entrant sells to one downstream firm if (i)* $\gamma \in [0.714, 0.899) \wedge \theta < \hat{\theta}(\gamma)$, *or (ii)* $\gamma \in [0.899, 1)$.

- *The entrant sets* $w_e = c_I$ *if* $\gamma \in [0, 0.618)$, *and* $w_e < c_I$ *if* $\gamma \in [0.618, 1)$.

Here, $\hat{\theta}(\gamma) : [0.714, 0.899] \to [0.114, 1]$ *is a strictly increasing function.*[15]

[13] This notation of the efficiency advantage follows Abito and Wright (2008).

[14] Here and in the following numbers are rounded up to three decimals.

[15] The explicit definition of $\hat{\theta}(\gamma)$ is given by equation (2.13) in the appendix.

Congruent with Lemma 2.2, this lemma shows that dependent on the degree of product differentiation the entrant induces either both downstream firms or one downstream firm to breach. However, in this example not only the degree of product differentiation but also the size of the entrant's efficiency advantage is important. More precisely, when $\gamma \in [0.714, 0.899)$ the entrant only induces breach of both downstream firms if its efficiency advantage is sufficiently high. Intuitively, when the entrant induces both downstream firms to breach it receives a higher demand but in order to induce both downstream firms to breach it has to set a lower wholesale price. So it faces the tradeoff that if it induces both downstream firms to breach it gains from a higher demand but looses from having to set a lower price. The key insight is that when the entrant's efficiency advantage is high it can gain more by receiving a higher demand and thus rather chooses to induce both downstream firms to breach.

As Lemma 2.2, Lemma 2.2′ shows that increased downstream competition makes it more difficult for breaching firms to raise the damage payment, so that it becomes necessary for the entrant to set its wholesale price below c_I when $\gamma \geq 0.618$.

In the following, we assess the effect of exclusive contracting on welfare and consumer surplus. Here, we obtain the following result.

Proposition 2.1′ *The effect of exclusive contracting on welfare and consumer surplus is neutral when* $\gamma \in [0, 0.618)$*, positive when (i)* $\gamma \in [0.618, 0.714)$ *or (ii)* $\gamma \in [0.714, 0.899)$ $\wedge\ \theta \geq \hat{\theta}(\gamma)$*, and negative when (i)* $\gamma \in [0.714, 0.899) \wedge \theta < \hat{\theta}(\gamma)$ *or (ii)* $\gamma \in [0.899, 1)$*.*

Proposition 2.1′ shows that with a linear demand function the effect of exclusive contracting has procompetitive effects in a sizable range in which downstream competition is moderate. The range increases with the entrant's efficiency advantage.

2.5 Concluding Remarks

We have shown that naked exclusion has procompetitive effects if downstream firms can breach exclusive contracts and competition between them is moderate. In this situation, both downstream firms sign the contract with the incumbent but also *both* downstream firms breach it and buy from the entrant later on. Because downstream firms have to pay expectation damages to the incumbent when breaching, the entrant must set its wholesale price sufficiently low. In particular, it must set its wholesale price lower than absent exclusive contracting. As a consequence, downstream firms set lower prices to final consumers, which leads to a rise in consumer surplus and welfare.

Our analysis challenges the view that naked exclusion is anticompetitive or at best neutral—a conclusion emanating from the previous literature. We find that exclusive dealing, although being intended by an incumbent firm as an entry deterring and therefore anticompetitive device, can have procompetitive effects. This speaks against a per se approach toward exclusive dealing.

A limitation of our model is that we assumed that entry costs are sufficiently small. This way we wanted to rule out cases in which the rival does not find it profitable to enter. If we dropped this assumption, exclusive contracting would more readily lead to entry deterrence, and thus it would more likely be anticompetitive.

Following SW we confined our attention to linear upstream prices. The case of two-part tariffs is much more complicated and therefore beyond the scope of this work. If upstream firms could offer two-part tariffs the analysis would change in the following way. The incumbent firm would be able to avoid double marginalization implying that, when downstream firms sign exclusive deals, it would obtain the monopoly profit of the industry regardless of the degree of downstream competition. Downstream firms had to pay higher damages to the incumbent firm in case of contract breach. Thus, the entrant needed to offer an even lower wholesale price to render breaching profitable, given that negative fixed fees are not possible (e.g., due to moral hazard issues). So if the entrant was sufficiently more efficient, the effect identified in this work would carry over to the case of two-part tariffs. In fact, exclusive dealing would lead to even lower wholesale prices than in case of linear upstream pricing.

2.6 Appendix: Proofs

Proof of Lemma 2.1

The proof proceeds in three steps. We first calculate the compensation x_2, which the incumbent has to offer so that both downstream firms sign the exclusive contract. We then calculate the compensation x_1, which I has to offer so that exactly one downstream firm signs the exclusive contract. Finally, we compare the net profits that I makes when it induces both, one or neither downstream firm to sign.

In the following, we denote the number of signed downstream firms by $S \in \{0,1,2\}$. The compensation x_2 must equal the additional profit that a downstream firm can make when rejecting the contract given the other downstream firm accepts it:

$$x_2 = \pi^f_{i|S=1} - \pi^c_{i|S=2}.$$

Here, $\pi^f_{i|S=1}$ denotes a downstream firm's profit when rejecting the contract while the rival downstream firm accepts it. $\pi^c_{i|S=2}$ denotes a downstream firm's profit when both accept it. For any compensation above x_2 accepting is strictly preferred by the downstream firms but I makes lower profits.

If both downstream firms accept the exclusive contract, I's maximization problem is[16]

$$\max_{w_i, w_j} \quad D\left(\mathbf{p}(w_i, w_j);\gamma\right)(w_i - c_I) + D(\mathbf{p}(w_j, w_i);\gamma)(w_j - c_I).$$

Since both downstream firms are symmetric, the optimal input prices, w_i and w_j, are identical. Let us denote the solution to this problem $w_i^* = w_j^* = w_I$. When both downstream firms are captive, I charges the monopoly wholesale price to them as it receives the same profits from them whether they breach or not. I's profit is then $\Pi_{I|S=2} = 2D(\mathbf{p}(w_I, w_I);\gamma)(w_I - c_I)$ and a downstream firm's profit, excluding the compensation payment, is $\pi^c_{i|S=2} = D(\mathbf{p}(w_I, w_I);\gamma)(p(w_I, w_I) - w_I)$.

Now suppose that one downstream firm rejects the contract. In the subsequent price game I and E compete for free downstream firms. Note that the captive downstream firm can also become free by breaching the contract. The standard Bertrand argument implies that I offers a wholesale price $w_f = c_I$ and E offers a wholesale price $w_e \leq c_I$ to free downstream firms. It could be optimal for E to set $w_e < c_I$ to induce the captive

[16] In the following we use $D(p_i, p_j;\gamma)$ as a short-cut for $\max\{0, D(p_i, p_j;\gamma)\}$, that is, we do not explicitly write out if a demand function becomes zero. We do so to reduce the notational burden.

downstream firm to breach. In order to verify this, we determine whether the captive downstream firm has an incentive to breach if E sets $w_e = c_I$.

If the captive downstream firm does not breach the contract, its input price is w_c. Since I gets the same profit from the captive downstream firm whether it breaches or not, w_c is $\arg\max_w D(\mathbf{p}(w, c_I); \gamma)(w - c_I)$. This yields $w_c > c_I$. The captive downstream firm's profit when not breaching is $D(\mathbf{p}(w_c, c_I); \gamma)(p(w_c, c_I) - w_c)$. If the captive downstream firm instead breaches, its profit is $D(\mathbf{p}(c_I, c_I); \gamma)(p(c_I, c_I) - c_I)$ net the damage payment $D(\mathbf{p}(w_c, c_I); \gamma)(w_c - c_I)$ that it has to pay to I. Thus, breaching is profitable for the captive downstream firm if

$$D(\mathbf{p}(c_I, c_I); \gamma)(p(c_I, c_I) - c_I) - D(\mathbf{p}(w_c, c_I); \gamma)(w_c - c_I) \geq D(\mathbf{p}(w_c, c_I); \gamma)(p(w_c, c_I) - w_c)$$

or

$$D(\mathbf{p}(c_I, c_I); \gamma)(p(c_I, c_I) - c_I) \geq D(\mathbf{p}(w_c, c_I); \gamma)(p(w_c, c_I) - c_I),$$

which is satisfied by our assumption of Section 2.2. Hence, the captive downstream firm breaches the contract when E sets $w_e = c_I$, so that it is optimal for E to set $w_e = c_I$ and no lower wholesale price. E finds it optimal to enter since by assumption $2(c_I - c_E)D(\mathbf{p}(c_I, c_I); \gamma) > f$. As the captive downstream firm breaches, the downstream firm that did not sign the contract makes profits equal to $\pi^f_{i|S=1} = D(\mathbf{p}(c_I, c_I); \gamma)(p(c_I, c_I) - c_I)$. We can deduce that I has to offer

$$x_2 = D(\mathbf{p}(c_I, c_I); \gamma)(p(c_I, c_I) - c_I) - D(\mathbf{p}(w_I, w_I); \gamma)(p(w_I, w_I) - w_I)$$

as compensation to each downstream firm for accepting the exclusive contract.

We now derive the compensation x_1 that I has to offer to induce a single downstream firm to sign the exclusive contract. This compensation must equal the additional profit that a downstream firm can make when rejecting the exclusive contract provided the other downstream firm rejects it, i.e.,

$$x_1 = \pi^f_{i|S=0} - \pi^c_{i|S=1}.$$

Here, $\pi^f_{i|S=0}$ denotes a downstream firm's profit when both firms reject the contract, while $\pi^c_{i|S=1}$ denotes a downstream firm's profit when it signs the contract while the rival firm rejects it. If both downstream firms reject the contract, E enters and the subsequent price game between the upstream firms results in simple Bertrand duopoly wholesale prices, i.e., both upstream firms set wholesale prices equal to c_I. Thus, when both downstream firms reject the contract, they make profits equal to $\pi^f_{i|S=0} = D(\mathbf{p}(c_I, c_I); \gamma)(p(c_I, c_I) - c_I)$.

From the analysis above we know that a downstream firm's profit when it signs the contract, while the rival firm rejects it, is $\pi^c_{i|S=1} = D(\mathbf{p}(c_I, c_I);\gamma)(p(c_I, c_I) - c_I) - D(\mathbf{p}(w_c, c_I);\gamma)(w_c - c_I)$. We can deduce that I must offer

$$x_1 = D(\mathbf{p}(w_c, c_I);\gamma)(w_c - c_I)$$

as compensation in order to induce one downstream firm to sign the exclusive contract.

We now compare the net profits that I makes when inducing both downstream firms, one or neither downstream firm to sign the exclusive contract. When it induces both downstream firms to sign the exclusive contract, its net profit is

$$2\big[D(\mathbf{p}(w_I, w_I);\gamma)(w_I - c_I) + D(\mathbf{p}(w_I, w_I);\gamma)(p(w_I, w_I) - w_I) - D(\mathbf{p}(c_I, c_I);\gamma)(p(c_I, c_I) - c_I)\big].$$

It offers x_2 to each downstream firm as compensation for signing and receives the monopoly profit whether the downstream firms breach or not. When it induces one downstream firm to sign its net profit is zero. It pays x_1 as compensation for signing to one downstream firm, makes zero profit and receives a damage payment equal to x_1 because the signed downstream firm breaches. Since I and E are perfect Bertrand competitors but E is more efficient, I also makes zero net profit when inducing neither downstream firm to sign. Hence, I makes use of exclusive contracting only if it is able to profitably induce *both* downstream firms to accept the exclusive contract, i.e., if

$$D(\mathbf{p}(w_I, w_I);\gamma)(p(w_I, w_I) - c_I) \geq D(\mathbf{p}(c_I, c_I);\gamma)(p(c_I, c_I) - c_I). \tag{2.2}$$

If the products are independent of each other, i.e., if $\gamma = 0$, the right-hand side is larger than the left-hand side since no double marginalization takes place. To the converse, if the products are (almost) perfect substitutes, i.e., if $\gamma \to 1$, the right-hand side is zero since $p(c_I, c_I) \to c_I$, while the left-hand side is still positive since $p(w_I, w_I) \to w_I > c_I$. Therefore, there must exist at least one intermediate value of γ, denote it $\widehat{\gamma}$, at which (2.2) holds with equality. If there are multiple γ satisfying (2.2), let $\widehat{\gamma}$ be the largest. It follows that both downstream firms sign the exclusive contract if $\gamma \geq \widehat{\gamma}$. ■

Proof of Lemma 2.2

If both downstream firms signed the exclusive contract, I charges the monopoly wholesale price $w_c = w_I$ to captive downstream firms and the Bertrand duopoly price $w_f = c_I$ to free downstream firms. Hence, E is constraint in its pricing decision to free downstream firms by $w_e \leq c_I$. It may choose to induce both downstream firms or one downstream firm to breach. To induce both downstream firms to breach it needs to set a wholesale price $w_e \leq w_E$, where w_E is defined by

$$D(\mathbf{p}(w_E, w_E);\gamma)(p(w_E, w_E) - w_E) - D(\mathbf{p}(w_I, w_I);\gamma)(w_I - c_I)$$
$$-D(\mathbf{p}(w_I, w_E);\gamma)(p(w_I, w_E) - w_I) = 0. \tag{2.3}$$

E optimally sets its wholesale price so that the downstream firms are indifferent between breaching or not. The first term denotes the profit that a downstream firm obtains when breaching provided the other downstream firm also breaches, the second term denotes the damage payment to I in case of contract breach, which is half the profit that I makes when none of the downstream firms breaches, and the third term denotes the profit that a downstream firm makes when not breaching provided the other downstream firm breaches.

If $\gamma = 0$, it is easy to see that there exists an equilibrium in which both downstream firms breach when E sets $w_E = c_I$ since $D(p(c_I))(p(c_I) - c_I) > D(p(w_I))(p(w_I) - c_I)$. If, however, γ becomes sufficiently large, E needs to set $w_E < c_I$ for such an equilibrium to exist. To see this note that the first term of (2.3) goes to zero when downstream competition becomes very intense as $p(w_E, w_E) \to w_E$, while the two last terms of (2.3) are negative. Thus, when γ is sufficiently large and E sets $w_E = c_I$, the condition for both downstream firms to breach would be violated. E then needs to set $w_E < c_I$, which increases the first term, does not change the second, and raises the third term, so that (2.3) is satisfied. It follows that there must exist a value of γ, which we denote $\check{\gamma}$, such that E needs to set $w_E = c_I$ if $\gamma \leq \check{\gamma}$ and $w_E < c_I$ if $\gamma > \check{\gamma}$ for an equilibrium in which both downstream firms breach to exist.

We now turn to the case in which E wants to induce just one downstream firm to breach. To do so, E must set $w_e \leq w'_E$, where w'_E is defined by

$$D(\mathbf{p}(w'_E, w_I);\gamma)(p(w'_E, w_I) - w'_E) - 2D(\mathbf{p}(w_I, w_I);\gamma)(w_I - c_I)$$
$$+D(\mathbf{p}(w_I, w'_E);\gamma)(w_I - c_I) - D(\mathbf{p}(w_I, w_I);\gamma)(p(w_I, w_I) - w_I) = 0. \tag{2.4}$$

The first term denotes the profit that a downstream makes when breaching provided the other downstream firm does not breach, the second and the third term denote the damage payment to I in case of contract breach, which is the profit that I makes if the downstream firm does not breach minus the profit that I makes when it breaches, and the fourth term denotes the profit that a downstream firm makes when not breaching provided the other downstream firm does not breach.

Two cases can now occur, either $w'_E \geq w_E$ or $w'_E < w_E$. If $w'_E \geq w_E$, then in any subgame perfect equilibrium one downstream firm breaches if E sets $w_e \in (w_E, w'_E]$ and both downstream firms breach if E sets $w_e \leq w_E$. If, however, $w'_E < w_E$, there are two subgame perfect equilibria if E sets $w_e \in (w'_E, w_E]$, with either both downstream firms or no downstream firm breaching. By assumption the downstream firms are able to coordinate themselves to play the equilibrium that is Pareto dominant from their perspective. Here, the Pareto dominant equilibrium is the equilibrium in which no downstream firm breaches as each firm exerts a negative externality on the other firm when breaching. It follows that in this case E can only induce both downstream firms to breach when setting $w_e = w'_E$.

We will now show that there always exists a region in which E must set $w_e < c_I$ to induce both downstream firms to breach. We do so by showing that at $\gamma = \check{\gamma}$, at which $w_E = c_I$, w'_E lies below c_I. If w'_E lies below c_I at $\gamma = \check{\gamma}$, E must set $w_e < c_I$ to induce both downstream firms to breach. By our assumption on f, E would nevertheless find it profitable to enter.

We know that at $\check{\gamma}$ equation (2.3) can be written as

$$\begin{aligned}D(\mathbf{p}(c_I,c_I);\check{\gamma})(p(c_I,c_I)-c_I) - D(\mathbf{p}(w_I,w_I);\check{\gamma})(w_I-c_I)\\ -D(\mathbf{p}(w_I,c_I);\check{\gamma})(p(w_I,c_I)-w_I) = 0.\end{aligned} \tag{2.5}$$

We need to show that w'_E is lower than c_I when (2.5) is fulfilled, which is equivalent to the left-hand side of (2.4) being negative when $w'_E = c_I$, i.e.,

$$\begin{aligned}D(\mathbf{p}(c_I,w_I);\check{\gamma})(p(c_I,w_I)-c_I) - 2D(\mathbf{p}(w_I,w_I);\check{\gamma})(w_I-c_I)\\ +D(\mathbf{p}(w_I,c_I);\check{\gamma})(w_I-c_I) - D(\mathbf{p}(w_I,w_I);\check{\gamma})(p(w_I,w_I)-w_I) < 0.\end{aligned} \tag{2.6}$$

Subtracting the left-hand side of (2.6) from the left-hand side of (2.5) and rearranging the terms, we obtain

$$\begin{aligned}D(\mathbf{p}(c_I,c_I);\check{\gamma})(p(c_I,c_I)-c_I) + D(\mathbf{p}(w_I,w_I);\check{\gamma})(p(w_I,w_I)-c_I)\\ -D(\mathbf{p}(c_I,w_I);\check{\gamma})(p(c_I,w_I)-c_I) - D(\mathbf{p}(w_I,c_I);\check{\gamma})(p(w_I,c_I)-c_I),\end{aligned}$$

which needs to be positive for our result to hold. We can rewrite the last expression as

$$\begin{aligned}D(\mathbf{p}(c_I,c_I);\check{\gamma})(p(c_I,c_I)-c_I) - D(\mathbf{p}(w_I,c_I);\check{\gamma})(p(w_I,c_I)-c_I)\\ +D(\mathbf{p}(w_I,w_I);\check{\gamma})(w_I-c_I) + D(\mathbf{p}(w_I,w_I);\check{\gamma})(w_I-c_I)\\ -D(\mathbf{p}(c_I,w_I);\check{\gamma})(w_I-c_I) - D(\mathbf{p}(w_I,w_I);\check{\gamma})(w_I-c_I)\\ +D(\mathbf{p}(w_I,w_I);\check{\gamma})(p(w_I,w_I)-w_I) - D(\mathbf{p}(c_I,w_I);\check{\gamma})((p(c_I,w_I)-w_I).\end{aligned} \tag{2.7}$$

We start with the first line of (2.7). We know that (2.3) is just satisfied at $\check{\gamma}$, i.e.,

$$D(\mathbf{p}(c_I,c_I);\check{\gamma})(p(c_I,c_I)-c_I) = D(\mathbf{p}(w_I,w_I);\check{\gamma})(w_I-c_I) + D(\mathbf{p}(w_I,c_I);\check{\gamma})(p(w_I,c_I)-w_I). \tag{2.8}$$

Inserting the right-hand side of (2.8) into the first line of (2.7) gives

$$\begin{aligned} &D(\mathbf{p}(w_I, w_I);\check{\gamma})(w_I - c_I) + D(\mathbf{p}(w_I, c_I);\check{\gamma})(p(w_I, c_I) - w_I) \\ &-D(\mathbf{p}(w_I, c_I);\check{\gamma})(w_I - c_I) - D(\mathbf{p}(w_I, c_I);\check{\gamma})(p(w_I, c_I) - w_I) \\ &= [D(\mathbf{p}(w_I, w_I);\check{\gamma}) - D(\mathbf{p}(w_I, c_I);\check{\gamma})](w_I - c_I) > 0. \end{aligned}$$

The first line of (2.7) is therefore positive since a downstream firm's demand increases in the rival firm's price, i.e. $D(\mathbf{p}(w_I, w_I);\check{\gamma}) > D(\mathbf{p}(w_I, c_I);\check{\gamma})$.

Now we turn to the second and third line. In case both downstream firms have signed the contract, we know that I maximizes

$$\max_{w_i, w_j} \quad D(\mathbf{p}(w_i, w_j);\check{\gamma})(w_i - c_I) + D(\mathbf{p}(w_j, w_i);\check{\gamma})(w_j - c_I)$$

and that w_I is the solution to this maximization problem. Therefore, line 2 of (2.7) equals I's monopoly profit when it charges $w_i = w_j = w_I$, while line 3 of (2.7) displays I's profit when making a suboptimal pricing decision, namely $w_i = c_I$ and $w_j = w_I$. It follows that line 2 and 3 of (2.7) are positive by the definition of w_I.

Closer inspection of line 4 reveals that it is positive if a downstream firm makes higher profits when setting its price on the basis of its true input cost—call it c'—instead of a lower input cost—call it $c < c'$. When basing its pricing decision on different costs, a downstream firm does not only change its own price but also its rival's price. Generally optimality of the cost-based decision requires

$$\max_{c} \quad D(\mathbf{p}(c, y);\check{\gamma})(p(c, y) - c'),$$

which gives a first-order condition of

$$\begin{aligned} &D(\mathbf{p}(c, y);\check{\gamma})\frac{\partial p(c, y)}{\partial c} + \frac{\partial D(\mathbf{p}(c, y);\check{\gamma})}{\partial p(c, y)}\frac{\partial p(c, y)}{\partial c}(p(c, y) - c') \\ &+\frac{\partial D(\mathbf{p}(c, y);\check{\gamma})}{\partial p(y, c)}\frac{\partial p(y, c)}{\partial c}(p(c, y) - c') = 0. \end{aligned} \tag{2.9}$$

Further, the optimality condition for the downstream price $p(c, y)$, resulting from the maximization problem

$$\arg\max_{p(c,y)} \quad D(\mathbf{p}(c, y);\check{\gamma})(p(c, y) - c),$$

is given by

$$\frac{\partial D(\mathbf{p}(c, y);\check{\gamma})}{\partial p(c, y)}(p(c, y) - c) + D(\mathbf{p}(c, y);\check{\gamma}) = 0$$

and can be rewritten as

$$D(\mathbf{p}(c,y);\check{\gamma}) + \frac{\partial D(\mathbf{p}(c,y);\check{\gamma})}{\partial p(c,y)} p(c,y) = \frac{\partial D(\mathbf{p}(c,y);\check{\gamma})}{\partial p(c,y)} c. \tag{2.10}$$

Inserting (2.10) into (2.9) gives

$$\frac{\partial D(\mathbf{p}(c,y);\check{\gamma})}{\partial p(c,y)} \frac{\partial p(c,y)}{\partial c}(c-c') + \frac{\partial D(\mathbf{p}(c,y);\check{\gamma})}{\partial p(y,c)} \frac{\partial p(y,c)}{\partial c}(p(c,y)-c') = 0.$$

The second term is positive while the first term depends on the sign of $c - c'$. Since $\partial D(\mathbf{p}(c,y);\check{\gamma})/\partial p(c,y)$ is negative and $\partial p(c,y)/\partial c$ is positive, optimality requires $c > c'$. Hence, it can never be better for a downstream firm to set a price on the basis of a lower input cost than its true input cost, which implies that line 4 of (2.7) must be positive.

We can conclude that the expression in (2.7) is positive, implying $w'_E < c_I$ at $\check{\gamma}$. By continuity there exists a region around $\check{\gamma}$, such that E must set $w_e < c_I$ to induce the downstream firms to breach. Let us denote the lower bound of this region $\widetilde{\gamma}$ and the upper bound $\overline{\gamma}$, with $\widetilde{\gamma} < \check{\gamma} < \overline{\gamma}$.

Finally, we need to show that $\check{\gamma}$ lies indeed above $\widehat{\gamma}$. First note that at $\gamma = \widehat{\gamma}$ condition (2.2) can be written as

$$D(\mathbf{p}(c_I,c_I);\widehat{\gamma})(p(c_I,c_I)-c_I) = D(\mathbf{p}(w_I,w_I);\widehat{\gamma})(p(w_I,w_I)-w_I)+D(\mathbf{p}(w_I,w_I);\widehat{\gamma})(w_I-c_I). \tag{2.11}$$

However, $\check{\gamma}$ is defined by

$$D(\mathbf{p}(c_I,c_I);\check{\gamma})(p(c_I,c_I)-c_I) = D(\mathbf{p}(w_I,c_I);\check{\gamma})(p(w_I,c_I)-w_I)+D(\mathbf{p}(w_I,w_I);\check{\gamma})(w_I-c_I). \tag{2.12}$$

Equations (2.11) and (2.12) differ in the first terms on the right-hand sides. We know that $D(\mathbf{p}(w_I,c_I);\gamma)(p(w_I,c_I) - w_I) < D(\mathbf{p}(w_I,w_I);\gamma)(p(w_I,w_I) - w_I)$. In addition, we know that for $\gamma = 0$, the left-hand sides of (2.11) and (2.12) are bigger than the respective right-hand sides while for $\gamma \to 1$, the reverse holds true. Since $\widehat{\gamma}$ is defined as the largest γ for which (2.11) holds, it follows that for $\check{\gamma}$ to fulfill (2.12) we must have $\check{\gamma} > \widehat{\gamma}$. To complete the proof, we define $\underline{\gamma} \equiv \max[\widehat{\gamma},\widetilde{\gamma}]$. ■

Proof of Lemma 2.1′

From the analysis above we know that I makes use of exclusive contracting if the monopoly profit that it earns when both downstream firms sign is higher than twice the compensation ($x_2 = \pi^c_{i|S=2} - \pi^f_{i|S=1}$) that it has to offer to each downstream firm for signing. I's monopoly wholesale price is $w_I = (1 + c_I)/2$. Thus, when both downstream firms sign the contract, I obtains a monopoly profit equal to $\Pi_I(w_I, w_I) = (1 - c_I)^2/(2(1+\gamma)(2-\gamma))$ and each downstream firm makes profits equal to $\pi^c_{i|S=2} = (1-c_I)^2(1-\gamma)/(4(1+\gamma)(2-\gamma)^2)$. We know from the previous analysis that a single captive downstream firm is induced to breach when E sets $w_e = c_I$. Therefore, a downstream firm's profit when rejecting the exclusive contract, given the rival firm accepts it and breaches it later, is $\pi^f_{i|S=1} = (1-c_I)^2(1-\gamma)/((2-\gamma)^2(1+\gamma))$. We can deduce that I makes effective use of exclusive contracting if

$$\Pi_I(w_I, w_I) - 2\left[\pi^c_{i|S=2} - \pi^f_{i|S=1}\right] = \frac{(1-c_I)^2}{2(1+\gamma)(2-\gamma)} - \frac{3(1-c_I)^2(1-\gamma)}{2(2-\gamma)^2(1+\gamma)} \geq 0.$$

It is easy to verify that the inequality holds if and only if $\gamma \in [0.5, 1)$. ■

Proof of Lemma 2.2′

From the previous lemma we know that I does not offer exclusive contracts to the downstream firms if $\gamma \in [0, 0.5)$. All downstream firms are free in that case. The subsequent price game between the upstream firms results in the simple Bertrand duopoly prices. Both downstream firms buy from E at $w_e = c_I > c_E$.

If $\gamma \in [0.5, 1)$, we know that I offers exclusive contracts that both downstream firms sign. I charges the monopoly wholesale price, $w_c = w_I$, to captive downstream firms and the Bertrand duopoly price $w_f = c_I$ to free downstream firms, i.e., those firms that breach the contract later on. It follows that E is constraint in its pricing decision to free downstream firms by $w_e \leq c_I$. We first determine w_E, i.e., the price at which a downstream firm is indifferent between breaching or not, given the rival firm breaches. From the proof of Lemma 2.2, we know that w_E is given by (2.3). If firm i adheres to the contract while firm j breaches, firm i only receives positive demand if p_i and p_j are sufficiently close, i.e., if $(-1+\gamma+p_i)/\gamma < p_j$. If p_i and p_j are sufficiently close, condition (2.3) is satisfied for

$$\begin{aligned} w_{E_1} =& \frac{1}{4(2-\gamma^2)(2-2\gamma-\gamma^2)}\Big[4\gamma^4 + 2(3+c_E)\gamma^3 + \theta(1-c_E)(\gamma^3 - 2\gamma) - 4(3+c_E)\gamma \\ & - (1-c_E)(2-\theta)(2+\gamma)\sqrt{(1-\gamma)(2-\gamma^2)(6-8\gamma-\gamma^2+2\gamma^3)} + 16(1-\gamma^2)\Big], \end{aligned}$$

where we replaced c_I by $\theta(1+c_E)/2 + (1-\theta)c_E$.

If downstream competition is relatively strong, firm i receives no demand when adhering to the contract, given firm j breaches. In this case, we are in the region $0 < p_j \leq (-1+\gamma+p_i)/\gamma$ and condition (2.3) is satisfied for

$$w_{E_2} = 1 - \frac{(1-c_E)(2-\theta)\sqrt{(2-\gamma)(1-\gamma)}}{4(1-\gamma)}.$$

We now determine for which regions of γ the wholesale prices w_{E_1} and w_{E_2} are relevant. Determining p_i and p_j for the case in which firm j breaches and buys at w_{E_1} while firm i adheres to the contract and buys at $w_c = w_I$, and inserting these prices into $(-1+\gamma+p_i)/\gamma < p_j$, gives that w_{E_1} is relevant for $\gamma \in [0.5, 0.710)$. Similarly, by determining p_i and p_j for the case in which firm j buys at w_{E_2} and firm i buys at $w_c = w_I$, and inserting these prices into $p_j \leq (-1+\gamma+p_i)/\gamma$, gives that w_{E_2} is relevant for $\gamma \in [0.710, 1)$.[17]

We now turn to w'_E, i.e., E's wholesale price at which a downstream firm is indifferent between breaching or not provided the rival firm does not breach. From the proof of Lemma 2.2, we know that it is given by (2.4). If the captive firm j still receives positive demand when firm i breaches the contract, condition (2.4) is fulfilled for

$$\begin{aligned} w'_{E_1} =& \frac{1}{4(2-\gamma^2)^2}\Big[16 - 8\gamma - 16\gamma^2 + 3\gamma^3 + 4\gamma^4 + (4\gamma - \frac{3}{2}\gamma^3)(c_E(2-\theta)+\theta) \\ & - \frac{1}{2}(2-\theta)(1-c_E)(2+\gamma)\sqrt{48 - 96\gamma + 12\gamma^2 + 76\gamma^3 - 31\gamma^4 - 16\gamma^5 + 8\gamma^6}\Big]. \end{aligned}$$

If the captive firm j receives no demand when firm i breaches the contract, condition (2.4) is satisfied for

$$w'_{E_2} = \frac{2\,(1-\gamma)^2\,(2-\gamma^2) + c_E\,(2+(1-\gamma)\,\gamma^2)\,(2-\theta) + (2+\gamma^2-\gamma^3)\,\theta}{2\,(2-\gamma)^2\,\gamma\,(1+\gamma)}.$$

In the same way as above we can determine for which region of γ the two wholesale prices are relevant. Here, we obtain that w'_{E_1} is relevant for $\gamma \in [0.5, 0.706)$ and w'_{E_2} is relevant for $\gamma \in [0.706, 1]$.

It is straightforward to verify that w'_{E_1} lies below w_{E_1} for $\gamma \in [0.5, 0.706)$. Thus, when E charges w'_{E_1}, in the unique equilibrium *both* downstream firms breach the exclusive contract. When E charges $w_e \in (w'_{E_1}, w_{E_1}]$, there are two equilibria with either both downstream firms or no downstream firm breaching the exclusive contract.

[17] The reason why the threshold values for the two regions coincide at $\gamma = 0.710$ is that firm i's profit function has a kink but no jump at $(-1+\gamma+p_i)/\gamma = p_j$. Thus, the wholesale prices w_{E_1} and w_{E_2} are identical at the value where one switches from one region to the other.

By assumption the downstream firms play the Pareto dominant equilibrium which is the one in which no downstream firm breaches. It follows that it is optimal for E to charge $w_e = \min[w'_{E_1}, c_I]$, inducing both downstream firms to breach. We find that w'_{E_1} lies above c_I for $\gamma \in [0.5, 0.618)$. Since $c_I > c_E$, E sets $w_e = c_I$ and induces both downstream firms to breach if $\gamma \in [0.5, 0.618)$. To analyze whether it is profitable for E to set w'_{E_1} if $\gamma \in [0.618, 0.706)$ we need to compare w'_{E_1} with c_E. Since w'_{E_1} is strictly decreasing in γ, a sufficient condition for w'_{E_1} to be larger than c_E provided $\gamma \in [0.618, 0.706)$ is that $w'_{E_1} > c_E$ at $\gamma = 0.706$. We find that $w'_{E_1} > c_E$ at $\gamma = 0.706$ if $\theta \geq 0.121$, which is fulfilled by assumption. Therefore, in equilibrium E sets $w_e = w'_{E_1}$ and induces both downstream firms to breach if $\gamma \in [0.618, 0.706)$.

We now turn to the case in which $\gamma \in [0.706, 0.710)$. Here the relevant wholesale prices are w'_{E_2} and w_{E_1}. By comparing these wholesale prices we find that $w'_{E_2} < w_{E_1}$, which again implies that for $w_e \in (w'_{E_2}, w_{E_1}]$ multiple equilibria exist in which either both or no downstream firm breaches the contract. By the same argument as above, the downstream firms coordinate on the equilibrium in which none of them breaches since this is Pareto dominant. It is easy to verify that w'_{E_2} is smaller than c_I and that it exceeds c_E for $\gamma \in [0.706, 0.710)$ if $\theta \geq 0.121$. Therefore, it is optimal for E to set $w_e = w'_{E_2}$, inducing both downstream firms to breach if $\gamma \in [0.706, 0.710)$.

Finally, we turn to the case in which $\gamma \in [0.710, 1)$. E can choose between w_{E_2} and w'_{E_2}. For $\gamma \in [0.710, 0.714)$ we find that $w'_{E_2} < w_{E_2}$. In the same way as above, we obtain that it is optimal for E to set $w_e = w'_{E_2}$ and induce both downstream firms to breach if $\theta \geq 0.118$, which is fulfilled by assumption. To determine whether it is more profitable for E to set w_{E_2} or w'_{E_2} when $\gamma \in [0.714, 1)$, we compare the profits that E makes in each case. The profit that E makes when setting w_{E_2}, inducing both downstream firms to breach, is

$$\Pi_E(w_{E_2}, w_{E_2}) = \frac{(1-c_E)^2(2-\theta)\sqrt{(2-\gamma)(1-\gamma)}\left[4(1-\gamma)-(2-\theta)\sqrt{(2-\gamma)(1-\gamma)}\right]}{8(2-\gamma)(1-\gamma)^2(1+\gamma)},$$

and the profit that E makes when setting w'_{E_2}, inducing one downstream firm to breach, is

$$\Pi_E(w'_{E_2}, w_I) = \frac{(1-c_E)^2(2-\theta)(2(2-\gamma^2)(1-\gamma)^2-\theta(2+\gamma^2-\gamma^3))}{8(2-\gamma)^2\gamma^2(1+\gamma)}.$$

By solving $\Pi_E(w_{E_2}, w_{E_2}) = \Pi_E(w'_{E_2}, w_I)$ for θ we obtain

$$\hat{\theta}(\gamma) = \frac{2(2-(2-\gamma)\gamma(3+\gamma(1-2\sqrt{(2-\gamma)(1-\gamma)}-(2-\gamma)\gamma)))}{2-(2-\gamma)\gamma(1+2\gamma)}. \tag{2.13}$$

For all $\theta > \hat{\theta}(\gamma)$ we have that $\Pi_E(w_{E_2}, w_{E_2}) > \Pi_E(w'_{E_2}, w_I)$ and vice versa. It is straightforward to verify that $\hat{\theta}(\gamma)$ is strictly increasing in γ. Inserting $\gamma = 0.714$ into $\hat{\theta}(\gamma)$ yields $\hat{\theta}(0.714) = 0.114$, while $\hat{\theta}(\gamma)$ equals 1 when $\gamma = 0.899$.

We now have to show that it is profitable for E to enter and set either w_{E_2} or w'_{E_2} if $\gamma \in [0.714, 1)$. As w'_{E_2} is relevant when E's efficiency advantage is low, we only need to compare w'_{E_2} with c_E. Since w'_{E_2} is strictly increasing in γ, a sufficient condition for w'_{E_2} to be larger than c_E provided $\gamma \in [0.714, 1)$ is that $w'_{E_2} > c_E$ at $\gamma = 0.714$. We find that $w'_{E_2} > c_E$ at $\gamma = 0.714$ for $\theta \geq 0.114$, which is again fulfilled by our assumption that $\theta \geq 0.121$. Therefore, when $\gamma \in [0.714, 0.899)$ and $\theta < \hat{\theta}(\gamma)$, it is optimal for E to set $w_e = w'_{E_2}$, inducing only one downstream firm to breach, whereas, when $\gamma \in [0.714, 0.899)$ and $\theta \geq \hat{\theta}(\gamma)$, it is optimal for E to set $w_e = w_{E_2}$, inducing both downstream firms to breach. When $\gamma \in [0.899, 1)$, it is always optimal for E to set $w_e = w'_{E_2}$, inducing one downstream firm to breach. ■

Proof of Proposition 2.1′

When exclusive dealing is not possible, E enters and the subsequent price game between the upstream firms results in the simple Bertrand duopoly prices, i.e., both upstream firms set wholesale prices equal to c_I. When exclusive contracting is possible and $\gamma \in [0, 0.5)$, both downstream firms decline I's offer for an exclusive contract. When $\gamma \in [0.5, 0.618)$, both downstream firms sign the exclusive contract, but are induced to breach if E sets $w_e = c_I$. In both these cases the outcome is unaffected by the possibility of exclusive contracting.

When (i) $\gamma \in [0.618, 0.714)$ or (ii) $\gamma \in [0.714, 0.899) \wedge \theta \geq \hat{\theta}(\gamma)$, both downstream firms sign the exclusive contract and are induced to breach at a price $w_e < c_I$. Because the downstream firms acquire the input good at a lower price, they set lower prices to final consumers. As a consequence, consumer surplus and welfare rise.

Finally, when (i) $\gamma \in [0.714, 0.899) \wedge \theta < \hat{\theta}(\gamma)$ or (ii) $\gamma \in [0.899, 1)$, both downstream firms sign the exclusive contract and E induces one to breach at a wholesale price $w_e < c_I$. The breaching downstream firm sets a price to final consumers of $p_{ed} = (c_I + 2\gamma - 1)/2\gamma$, which leads to monopolization of the downstream market. Because p_{ed} is higher than the price that the downstream firms would set absent exclusive dealing, $p_{ned} = (c_I - \gamma + 1)/(2 - \gamma)$, consumer surplus and welfare fall. ■

Chapter 3

Price Discrimination and Fairness Concerns[0]

3.1 Introduction

Standard theory suggests that firms can substantially improve profitability through third degree price discrimination. But standard theory does not take into account that consumers might perceive it 'unfair' to charge different prices to different consumer groups. Due to consumers' fairness concerns, the profitability of third degree price discrimination might be adversely affected. Amazon.com, for instance, antagonized its customers by charging different prices for the same DVD titles. Customers were so outraged that Amazon.com abolished its price discriminating strategy within only three days. It claimed that the price differences were the result of a random 'price test' and refunded all customers who paid the higher prices.

In the present paper, we seek to examine carefully how the profitability of third degree price discrimination is affected by consumers' fairness concerns. Besides the optimal price discriminating tariff, we are interested in whether the provided contextual information matters. In particular, we want to find out whether firms obtain higher profits when consumers know that those consumers who are poorer are charged lower prices compared to when consumers do not know the wealth of the other consumers. This question is motivated by the observation that consumers seem to object price discrimination based on income differences, e.g., student discounts, less than other forms of price discrimination, e.g., price discrimination on the internet based on consumers' purchasing history.

[0]This chapter is based on joint work with Prof. Dr. Florian Englmaier and Prof. Dr. Markus Reisinger.

In the first part of this paper, we analyze the profitability of third degree price discrimination within a laboratory experiment. Subjects are split into groups of three, consisting of one firm, one h-consumer and one l-consumer. H-consumers' demand is less price elastic, so they are charged higher prices than l-consumers under price discrimination. They are also assigned a higher income. First, firms can choose between two different price menus. Thereafter, consumers can spend their income on the firms' product. For each purchasing decision consumers can make, they are informed about their own material payoff and that of the firms.

We analyze firms' profits when charging the discriminating tariff that is optimal under standard theory (denoted *spd* for strong price discrimination), when charging a weaker discriminating tariff *wpd*, and when charging a non-discriminating tariff *npd*. Further, we inquire whether firms' profits are affected when consumers know other consumers' price and income (treatment *i2*) compared to when they only know other consumers' price (treatment *i1*).

If consumers had no fairness preferences, firms would maximize their profits by charging the discriminating tariff *spd*. As a main result we find that the weaker discriminating tariff *wpd* yields on average 5% higher profits than *spd* in treatment *i1* and 7% higher profits in treatment *i2*. Even the non-discriminating tariff *npd* yields on average 2% higher profits than *spd* in treatment *i1*. Thus, firms can increase profitability by choosing a *weaker* discriminating tariff than the one predicted to be optimal under standard theory.

We gain more insight by estimating the effect of consumers' reciprocal reactions on firms' profits. We find that firms' profits from h-consumers are adversely affected by h-consumers' reciprocal reactions under price discrimination, implying that the disadvantaged h-consumers punish firms by reducing their demand. In contrast, firms' profits from l-consumers are positively affected, implying that the advantaged l-consumers reward firms by enhancing their demand. In line with empirical evidence of other studies, the *negative reciprocity effect* on firms' profits from h-consumers is strong and *significant*, whereas the *positive reciprocity effect* on firms' profits from l-consumers is weak and *not significant*. Consequently, firms' overall profits are negatively affected when choosing the discriminating tariffs, *wpd* or *spd*. Since the negative reciprocity effect intensifies compared to the positive reciprocity effect when firms choose the stronger price differential *spd*, firms obtain on average higher profits when choosing *wpd* than when choosing *spd*.

Furthermore, we find that negative and positive reciprocity effects decrease when consumers know other consumers' income. Because the decrease of the negative reciprocity effect is stronger, firms obtain on average 1% higher profits when choosing *spd* and 3% higher profits when choosing *wpd*.

In the second part of this paper, we adopt the model by Falk and Fischbacher (2006) on reciprocity to explain the results of our experiment. We argue that consumers judge firms' pricing decisions by comparing the material payoff they can achieve by purchasing from the firms with the material payoff they believe *other consumers* can achieve. When they know other consumers' income they adjust their price fairness judgment and correct for the income difference. Since h-consumers are charged higher prices than l-consumers under price discrimination, h-consumers believe that the material payoff they can obtain is lower than that of l-consumers. Consequently, they regard firms' pricing decisions as unfair. Their perception of price unfairness intensifies when the price difference gets larger, and it diminishes when they learn that they have a higher income. Accordingly, l-consumers, who are charged lower prices, regard firms' pricing decisions as fair. Their perception of price fairness intensifies when the price difference gets larger, and it diminishes when they learn that they have a lower income. The model ascertains that in reaction to perceived price unfairness h-consumers punish firms by reducing their demand and in reaction to perceived price fairness l-consumers reward firms by enhancing their demand. To the extent that h-consumers' negative reactions are stronger than l-consumers' positive reactions, the model predicts that the profitability of third degree price discrimination will be adversely affected when the price differential increases, and the adverse effect will diminish when consumers know other consumers' income

The remainder of the paper is organized as follows. The paper continues in Section 3.2 with a summary of the previous literature. Section 3.3 provides a detailed description of the experimental design. Section 3.4 highlights the main results of the experiment. In Section 3.5, we lay out a model on reciprocity in the context of third degree price discrimination based on Falk and Fischbacher (2006). Finally, Section 3.6 concludes and points out directions for future research.

3.2 Previous Literature

Price discrimination is an important strategic instrument for firms in many product markets. Not surprisingly, a great deal of theoretical work has been devoted to analyze the *optimal price discriminating tariff*, both under monopoly and oligopoly. For surveys on this literature see, e.g., Armstrong (2006) or Stole (2007). Empirical evidence on the issue is relatively scarce. While some studies have analyzed the profitability of second degree price discrimination, e.g., Nevo and Wolfram (2002) or Busse and Rysman (2005), there is only very little evidence on the profitability of third degree price discrimination. Leslie (2004) analyzes the effectiveness of discount mail coupons targeted to consumers with lower willingness to pay, using data from a Broadway play. Verboven (1996) and Goldberg and Verboven (2001) provide evidence of international price discrimination in the European car market. Borenstein (1991) shows that the differences in retail margins for leaded and unleaded gasoline are correlated with income and availability of leaded gasoline in a particular area.

A large literature studies the profitability of *price increases* under the consideration of consumers' fairness preferences. Kahneman, Knetsch, and Thaler (1986a,b) find that consumers are concerned with firms' intention behind price increases. They propose the *dual entitlement principle*, according to which consumers feel entitled to the terms of their reference transaction but acknowledge that firms are entitled to the terms of their reference transaction as well. Following this, consumers regard price increases as unfair if these price increases are not justified by increased costs and lead to an increase in firms' reference profit. Their arguments are illustrated within a formal model by Rotemberg (2005, 2011), verified by Franciosi et al. (1995), and complemented by Martins and Monroe (1994), Campell (1999), and Bolton, Warlop, and Alba (2003). In this literature, it is presumed that consumers compare their payoff of the current period with that of previous periods (self/self comparisons) and with that of firms (internal self/other comparisons).

Surprisingly little research has been devoted to the question how consumers' fairness preferences impact the effectiveness of *third degree price discrimination*. Rotemberg (2011) offers a start in the analysis. He argues that consumers object third degree price discrimination in case it demonstrates *insufficient firm altruism*. He shows within a theoretical model that altruistic firms would price discriminate based on the income of different consumer groups, charging higher prices to consumers with a higher income. Therefore, selfish firms could profit by mimicking altruistic firms, also adopting price discrimination based on income differences while avoiding price discrimination based on demand elasticities. By contrast, in our model consumers do not judge firms' pricing decisions by their altruistic intentions but by the outcome, which is the material payoff the consumers can obtain by purchasing from the firms compared to the material payoff they believe *other consumers* can obtain (external self/other comparisons). Hence, consumers only object price discrimination when they are charged higher prices than other consumers. They even approve price discrimination when they are charged lower prices. Different profit implications arise. However, similar to Rotemberg (2011) we find that firms obtain higher profits when consumers know that those consumers with a higher income are charged higher prices compared to when the wealth of the different consumer groups is unknown.

Our theory that consumers form their price fairness judgments by comparing their material payoff with that of other consumers and not with that of firms seems natural as consumers often do not know firms' payoff. Also, other consumers are more comparable due to higher similarity. Our theory is supported by the results of a field study by Anderson and Simester (2008). The authors analyze consumers' reaction to premium prices for larger sizes of women's apparel. From a firm's perspective such premium prices for larger sizes of apparel are justified by higher material costs. Thus, when consumers compared their material payoff with that of firms, they should accept such premium prices. The authors find, however, that consumers who demand larger sizes react unfavorably to paying a higher price.

The main difference between the field experiment by Anderson and Simester (2008) and our laboratory experiment lies in firms' motivation for charging varying prices. While in the field experiment firms charge varying prices because of different production costs, in our laboratory experiment firms charge varying prices because consumers differ with respect to their demand elasticities.

In another related study, Shor and Oliver (2006) investigate the effect of couponing on consumers' purchasing probabilities. Couponing can be seen as a device to price discriminate. The authors find that consumers, who do not posses a coupon but are prompted for a coupon on a web site, are less likely to purchase. They partly explain this adverse effect with consumers' belief that they will also be able to obtain a coupon when searching on the internet. Such an effect does not arise in our setting, in which price discrimination is based on consumers' characteristics. In our setting, it is predetermined which consumers will have to pay higher prices.

Price fairness assessments are usually a comparative phenomenon. Specifically, consumers usually use *reference prices* as a basis for their price fairness judgments. A closely related research stream therefore asks how consumers actually form reference prices. Lichtenstein, Bloch, and Black (1988) and Janiszewski and Lichtenstein (1999) propose that consumers use internal memory-based references. Other authors stress the importance of external references, in particular of prices charged by competitors (see, e.g., Büyükkurt, 1986, Urbany, Bearden, and Weilbaker, 1988, Lichtenstein and Bearden, 1989, Alba et. al, 1994, and Dholakia and Simonson, 2005). We stipulate that under third degree price discrimination especially prices charged to other consumers contribute to the formation of consumers' reference prices.

The main contribution of this work is to provide empirical evidence and a theoretical explanation for how the profitability of third degree price discrimination is affected by consumers' fairness concerns. In our setting, different consumer groups are charged varying prices based on their characteristics and the motivation for charging varying prices is that consumers differ in their demand elasticities. As a new explanation for consumers' behavioral reactions we propose that they form their price fairness judgments by comparing the material payoff they can obtain by purchasing from a firm with the material payoff that other consumers can obtain. Thus, we look at a three-player setting, stressing the importance of external self/other comparisons in the context of third degree price discrimination. With this framing we can explain the empirical findings that firms obtain higher profits by charging a weaker price differential than the one predicted to be optimal under standard theory, and further that firms' profitability increases when consumers know that those consumers who are charged lower prices have a lower income compared to when they do not know other consumers' income.

3.3 Experimental Procedures and Design

The experiment was computer-based and conducted at the experimental laboratory MELESSA of the University of Munich in August 2010, using the experimental software z-Tree (Fischbacher, 2007) and the organizational software Orsee (Greiner, 2004). In total, 192 participants were randomly recruited for 8 experimental sessions (graduate students were excluded). In any of the 8 experimental sessions 24 subjects participated. No subject could attend more than one session. On average, subjects earned 13.00 euro (including 4 euro show-up fee, with a minimum of 6.00 euro and a maximum of 21.40 euro) for a duration of approximately 50 minutes.

Upon arrival, subjects were seated at computer terminals in a large room that contains 25 terminals. The computer terminals are partitioned from each other by blinds, so that no subject could see the terminal screen of another participant. Subjects received three-pages instructions that were read aloud by the experimenter.[1] The instructions were framed in terms of a transaction in order to make the experiment less abstract and easier to understand. Before the experiment started subjects were asked to answer test questions that showed whether they understood the scenario, the tasks, and, in particular, the material payoff determination. The experiment started on the computer screen only after everybody had answered the test questions correctly and there were no further questions.

[1] A translation of the experimental instructions is provided in Appendix B.

At the beginning of the experiment subjects were randomly assigned a type: *firm*, *h-consumer* or *l-consumer*. By experimental design, h-consumers' demand was less price elastic, so that they were charged higher prices than l-consumers under price discrimination. Further, h-consumers were assigned income $I_h = 400$ EP and l-consumers income $I_l = 200$ EP, where EP (Experimental Points) was the experimental currency, with an exchange rate such that 1 EP corresponded to 1 euro-cent. Everybody knew that each subject had the same chance to be assigned either type. Subjects kept their type in all parts of the experiment.

The main part of the experiment consisted of three rounds. At the beginning of each round, subjects were randomly put together into groups of three, with one firm, one h-consumer and one l-consumer. In every following round they were randomly reassigned to a new group.[2] Subjects were completely anonymous and not identifiable, i.e., it was impossible for them to build reputations over the three rounds.

In each round a one-shot game was played. The sequence of actions was the following. First, the firm was asked to choose one out of two price menus, where the choice of price menus varied across treatments. The firm was informed about her expected profits under standard theory, i.e., about her expected profits under the assumption that consumers solely maximize their material payoff. The firm's choice $(p_h,\ p_l)$ was made public to consumers, whereupon consumers could spend their income on the firm's product, making a purchasing decision. Consumers were informed about their material payoff when choosing quantities q_h and q_l, which was equal to

$$\pi_h = 400 + 32q_h - 0.8q_h^2 - p_h q_h,$$
$$\pi_l = 200 + 16q_l - 0.2q_l^2 - p_l q_l.$$

Next to the information on their own material payoff, consumers received information about the firm's material payoff, which was equal to

$$\pi_f = \begin{cases} 0 & \text{if } q_h = 0,\ q_l = 0, \\ p_h q_h + 500 & \text{if } q_h > 0,\ q_l = 0, \\ p_l q_l + 500 & \text{if } q_h = 0,\ q_l > 0, \\ p_h q_h + p_l q_l + 1000 & \text{if } q_h > 0,\ q_l > 0. \end{cases}$$

[2]Since in each session 24 subjects participated (8 firms, 8 h-consumers and 8 l-consumers), the number of group combinations was 120. Hence, the chance that subjects were assigned to the same group two times in three rounds was 0.0069%. The chance that consumers were assigned with one particular firm in one group two times in three rounds was 1.37%.

To account for the additional losses that firms usually incur when consumers switch to other firms, we determined that firms only obtained 500 EP extra when consumers did not 'switch', that is, when they did not purchase nothing ($q_i = 0$).

Consumers received the material payoff information as exemplified in Table 3.1. The tables provided consumers clear insights which quantity they had to choose in order to maximize their own material payoff and, if they deviated from that choice, how this would affect their own material payoff and the firm's material payoff. All subjects knew that they would be paid according to the outcome generated by one of their choices, to be selected at random from the three rounds.

Table 3.1: Information provided to consumers

Quantity	Benefit from purchasing	Expenditure	Your Payoff	Sellers' payoff from selling to you
⋮	⋮	⋮	⋮	⋮
8	205	128	477	628
9	223	144	479	644
10	240	160	480	660
11	255	176	479	676
12	268	192	476	692
13	279	208	471	708
⋮	⋮	⋮	⋮	⋮

The objective of the experiment was to learn which price menu firms should choose in order to maximize their profits. Specifically, we wanted to find out whether consumers indeed exhibit no fairness preferences in the context of third degree price discrimination, so that firms can maximize their profits by choosing the price menu predicted to be optimal under standard theory. Therefore, we analyzed firms' profits when choosing the price menu predicted to be optimal under standard theory, in our example $p_h = 16$ and $p_l = 8$ (denoted *spd* for strong price discrimination), when choosing a weaker discriminating price menu *wpd* with $p_h = 14$ and $p_l = 10$, and when choosing a non-discriminating price menu *npd* with $p_h = 12$ and $p_l = 12$. In each round firms could choose between one of these three price menus and an alternative price menu *apd* with $p_h = 40$ and $p_l = 20$. That is, each of the price menus *npd*, *wpd* and *spd* came up in one round of the experiment, always together with *apd*. The sequence in which the price menus came up varied across subjects.

The price menus *npd*, *wpd* and *spd* yielded positive profits under standard theory, whereas the price menu *apd* yielded zero profits under standard theory. We chose *apd* such that firms would in expectation obtain relative low profits because we were only interested in comparisons between *npd*, *wpd* and *spd* and thus we wanted to assure that firms would mainly choose *npd*, *wpd* and *spd* instead of *apd*.

Hence, firms' choice of price pairs was rather limited. As will become clearer in Section 3.5, when we outline the model, if consumers anticipated that firms' choice of price pairs was rather limited, this would have caused them to act less reciprocally. In practice, firms' choice of price pairs is of course unlimited. Thus, in practice consumers' behavioral reactions might be stronger than in our experimental setting.

We were also interested in whether firms can obtain higher profits by providing consumers information about other consumers' income. Consumers therefore received either information about other consumers' price (treatment *i1*), or information about other consumers' price *and* income (treatment *i2*). We kept the information levels constant over the three rounds. Table 3.2 provides an overview of the different treatments.

Table 3.2: Treatments

		i1 (p_j)	i2 (p_j, I_j)
no price discrimination	($p_h = 12,\ p_l = 12$)	*npd_i1*	*npd_i2*
weak price discrimination	($p_h = 14,\ p_l = 10$)	*wpd_i1*	*wpd_i2*
strong price discrimination	($p_h = 16,\ p_l = 8$)	*spd_i1*	*spd_i2*

Having finished the main part of the experiment, subjects were asked to play the standard trust game.[3] The purpose was to verify whether consumers' reciprocal behavior in the experiment is correlated to their reciprocal behavior in the trust game. At last, subjects were asked to answer a short questionnaire about their socio-economic characteristics. Before subjects left, their earnings were paid to them in private by a person that was not the experimenter.

3.4 Experimental Results

In this section, we describe our experimental results. We first report which profits firms obtain on average in the different treatments. We then estimate how these profits are affected by consumers' reciprocal reactions. Further, we briefly examine the impact of self/self price comparisons over time. Finally, we show, as a robustness check, to what extent consumers' behavioral reactions to price discrimination are correlated to their behavioral reactions in the trust game.

[3] For further information on the trust game see Section 3.4.4.

3.4.1 Firms' Average Profits

Table 3.3 provides an overview of the average profits that firms obtain in the different treatments (in EP). We compute the percentage differential between firms' average profits in the experiment and firms' profits under standard theory in parentheses. Firms' profits under standard theory are the profits firms would obtain when consumers had no fairness preferences, i.e., when consumers always chose the quantities that maximize their own material payoff. Hence, a larger percentage differential reported in parentheses shows a larger impact of consumers' behavioral reactions on firms' profits. We report firms' profits from h-consumers, firms' profits from l-consumers, and firms' overall profits.

Table 3.3: Firms' average profits

Firms' average profits		npd	wpd	spd
from l-consumers	(i1)	512 (-17%)	605 (-7%)	620 (-6%)
	(i2)	517 (-17%)	598 (-8%)	606 (-8%)
from h-consumers	(i1)	648 (-2%)	591 (-10%)	516 (-20%)
	(i2)	584 (-12%)	629 (-4%)	540 (-16%)
overall	(i1)	1160 (-8%)	1195 (-8%)	1136 (-14%)
	(i2)	1101 (-13%)	1227 (-6%)	1146 (-13%)

Notes: Percentage differentials between firms' average profits and firms' profits under standard theory are reported in parentheses.

The results computed in Table 3.3 show that the price differential *spd*, which standard theory predicts to be optimal, yields on average 5% lower profits than the weaker price differential *wpd* in treatment *i1* and 7% lower profits in treatment *i2*.[4] In fact, it even yields on average 2% lower profits than the non-discriminating tariff *npd* in treatment *i1*. This suggests that the effectiveness of third degree price discrimination is deterred by negative consumer reactions, especially when the price differential is large. Firms can obtain higher profits by choosing a weaker price differential than the one predicted to be optimal under standard theory.

[4]Using a Mann Whitney U test or a Wilcoxon signed-rank test, we find no statistically significant evidence that firms' average profits differ across treatments. That is, we find no statistically significant evidence that *spd* yields higher profits than *wpd* or *npd*.

As standard theory suggests, firms' average profits from l-consumers are higher when choosing *spd* than when choosing *npd*. Contrary thereto, firms' average profits from h-consumers are *lower* when choosing *spd* than when choosing *npd*. This indicates that the gains from third degree price discrimination are deterred by *h*-consumers' negative reciprocal reactions.

Another finding is that price discriminating firms obtain on average higher profits in *i2*-treatments than in *i1*-treatments. That is, firms' profitability is higher when consumers know other consumers' income (conditional on consumers with a less price elastic demand, who are charged a *higher* price, also having a *higher* income). This indicates that in practice firms can obtain higher profits when they price discriminate based on income differences, charging consumers who are generally regarded as poorer, like students or the elderly, lower prices, compared to when they price discriminate on characteristics that do not reveal consumers' wealth.[5]

3.4.2 Reciprocity Effect on Firms' Profits

In this section, we examine the impact of consumers' reciprocal reactions on firms' profits under third degree price discrimination. We find, in line with the existing literature, that the degree of reciprocity is highly heterogeneous across consumers. 55% of consumers have no reciprocal preferences and choose in all three rounds the quantities that maximize their material payoff. The proportion of material payoff maximizing consumers is higher among l-consumers than among h-consumers (59% vs. 50%) and higher among women than among men (62% vs. 44%). In 10% of all purchasing decisions consumers choose *not* to purchase from the firm, i.e., to punish the firm maximally.

We conduct a multivariate OLS analysis to estimate how firms' profits are affected by consumers' reciprocal reactions. As dependent variable we use the percentage differential between firms' actual profits and firms' profits under standard theory.

In assessing the impact of consumers' reciprocal behavior on firms' profits, we face the following difficulty. If consumers reduce their demand, this can either be seen as inequality-reducing behavior (following consumer/firm comparisons) or as reciprocal punishment (following consumer/consumer comparisons), given the firm price discriminated and charged these consumers higher prices than other consumers. Similarly, when consumers increase their demand, this can either be seen as social-surplus-increasing behavior or as reciprocal reward, given the firm price discriminated and charged these

[5]The large profit differential in treatment *npd_i2* (-13%) compared to treatment *wpd_i2* (-6%) suggests that firms might be able to gain profits by charging consumers who are generally regarded as poorer a lower price, even if the demand of these poorer consumers is not more price elastic.

consumers lower prices than other consumers.[6,7] While inequality-reducing behavior and social-surplus-increasing behavior may arise under both, price discrimination and non-price discrimination treatments, reciprocal behavior may only arise under price discrimination treatments. We want to focus on the reciprocity hypothesis and its ability to explain deviations from standard theory. Thus, in order to isolate the reciprocity effects we include dummy variables for all treatments as dependent variables into the regression with the exception of the dummy variable for treatment *npd_i1* (see Table 3.2). The regression then shows the additional behavioral effects that arise in the price discrimination treatments compared to the non-price discrimination treatment *npd_i1*. We will interpret these additional effects as reciprocity effects.[8]

As control variable we include a gender dummy variable which equals one if a consumer is female. Further, we include a variable denoted *trustgeneral* that is obtained from consumers' answers of the questionnaire at the end of the experiment. There, subjects were asked to what extent they would confirm that one can generally trust others, with possible answers ranging from zero for 'trusting' to three for 'not trusting'. We also include a dummy variable *apd* as control, which indicates whether a consumer was charged the alternative price menu *apd* ($p_h = 40,\ p_l = 20$) in a previous round. Under *apd* subjects are exposed to comparatively high prices and may therefore perceive price offers in subsequent rounds, be it *npd* ($p_h = 12,\ p_l = 12$), *wpd* ($p_h = 14,\ p_l = 10$) or *spd* ($p_h = 16,\ p_l = 8$), as less unfair/more fair. So, we expect the coefficient for *apd* to be positive, implying that consumers punish less or reward more after they have been exposed to relatively high prices under *apd* before. Furthermore, we adjust standard errors for 64 clusters in consumers' identity.

The regression results are presented in Table 3.4. Again, we distinguish between firms' profits from h-consumers, firms' profits from l-consumers and firms' overall profits. The coefficients of interest are the coefficients of the price discrimination treatments. A negative treatment coefficient, for instance, indicates that firms' profits are negatively affected by consumers' reciprocal reactions.

The estimation shows that the difference between firms' actual profits from h-consumers and firms' profits from h-consumers under standard theory, denoted $\Delta\pi_f$

[6]In our setting social surplus-increasing behavior is unlikely as the benefit of firms when consumers increase their demand is only slightly higher than the loss consumers incur by doing so.

[7]Charness and Rabin (2002) provide an interesting analysis on the difference between the mentioned social preferences.

[8]Inequality reducing behavior would not be singled out entirely if consumer/firm comparisons varied across treatments. That would, for instance, be the case if h-consumers believed that firms obtain much higher profits when charging *spd* than when charging *npd*, and in consequence reduced their quantity choice by more when being charged *spd*. The effects we report as reciprocity effects would then be overestimated as they would consist of reciprocity *and* inequity aversion effects.

Table 3.4: Reciprocity impact on firms' profits

VARIABLES	$\Delta\pi_f$ (h-consumers)	$\Delta\pi_f$ (l-consumers)	$\Delta\pi_f$ (overall)
npd_i2	-11.34*	-0.796	-6.338
	(5.705)	(9.417)	(5.481)
wpd_i1	-12.74**	9.832	-2.013
	(6.040)	(7.258)	(4.688)
wpd_i2	-7.429*	7.524	-0.248
	(4.440)	(8.310)	(4.117)
spd_i1	-24.17***	10.76	-7.197
	(8.122)	(9.010)	(5.603)
spd_i2	-19.62***	6.969	-6.202
	(6.734)	(8.271)	(4.832)
gender_h	16.55***		9.851**
	(5.245)		(3.799)
trustgeneral_h	-16.37***		-7.833**
	(4.667)		(3.301)
apd_h	14.22**		12.98***
	(5.934)		(3.245)
gender_l		9.494*	3.543
		(5.060)	(3.246)
trustgeneral_l		-1.007	-0.416
		(2.784)	(2.204)
apd_l		4.609	2.186
		(5.803)	(5.174)
Constant	13.27**	-21.43**	-5.014
	(6.511)	(8.479)	(6.324)
Observations	181	181	181
R-squared	0.223	0.049	0.138

Notes: The table reports coefficients of OLS regressions. The dependent variables are the percentage differentials between firms' actual profits (from h-, l-, h- and l-consumers) and firms' profits (from h-, l-, h- and l-consumers) under standard theory. The treatment *npd_i1* is the baseline. Robust standard errors are reported in parentheses. The regressions are clustered by consumers and therefore control for individual fixed effects. *** denotes $p<0.01$, ** denotes $p<0.05$, and * denotes $p<0.1$.

(h-consumers), is 12.74 percentage points lower in treatment *wpd_i1* than in treatment *npd_i1* (see column 1 of Table 3.4). This suggest that under price discrimination firms' profits from h-consumers are negatively affected due to negative reciprocity.

By contrast, the difference between firms' actual profits from l-consumers and firms' profits from l-consumers under standard theory, denoted $\Delta\pi_f$ (l-consumers), is 9.83 percentage points *higher* in treatment *wpd_i1* than in treatment *npd_i1* (see column 2 of Table 3.4). This implies that under price discrimination firms' profits from l-consumers are positively affected due to positive reciprocity. The *positive reciprocity effect* on firms' profits from l-consumers is, however, *smaller than the negative reciprocity effect* on firms' profits from h-consumers, and it is not significant.

That positive reciprocity is a comparatively weak factor has also been found in other recent experimental studies. Consumers seem to react more when price differences are unfavorable to them, implying that they count negative deviations from the reference outcome more than positive deviations. Xia, Monroe, and Cox (2004) explain this finding with the different emotions that consumers have in the two states. In the context of third degree price discrimination the disadvantaged h-consumers presumably have *strong* negative feelings such as anger or disappointment, while the advantaged l-consumers may have *weak* positive feelings such as egoism-based pleasure or satisfaction.

Because the negative reciprocity effect on firms' profits from h-consumers is higher than the positive reciprocity effect on firms' profits from l-consumers, firms' overall profits are negatively affected by consumers' reciprocal reactions. Specifically, the overall profit differential, denoted $\Delta\pi_f$ (overall), is 2.01 percentage points lower in treatment *wpd_i1* than in treatment *npd_i1* (see column 3 of Table 3.4). This negative reciprocity effect on firms' overall profits is also not significant.

The results shown in Table 3.4 further reveal that $\Delta\pi_f$ (h-consumers) is even lower in treatment *spd_i1* than in treatment *wpd_i1*, implying that the negative reciprocity effect intensifies with the size of the price differential. Compared to treatment *npd_i1*, the profit differential is 12.74 percentage points lower in treatment *wpd_i1* and 24.17 percentage points lower in treatment *spd_i1*.[9]

The positive reciprocity effect also seems to increase with the size of the price differential. Compared to treatment *npd_i1*, the profit differential is 9.83 percentage points higher in treatment *wpd_i1* and 10.76 percentage points higher in treatment *spd_i1*. Clearly, the increase in the positive reciprocity effect is weaker than the increase in the negative reciprocity effect. Both are not significant.

Because *the negative reciprocity effect* on firms' profits from disadvantaged consumers *intensifies with the size of the price differential compared to the positive reciprocity effect* on firms' profits from advantaged consumers, the negative reciprocity effect on firms' overall profits is higher in treatment *spd_i1* than in treatment *wpd_i1*. In particular, the negative reciprocity effect on firms' overall profits is 2.01 percentage points in treatment *wpd_i1* and 7.20 percentage points in treatment *spd_i1*. The increase in the negative reciprocity effect on firms' overall profits is, however, not significant.[10]

The estimation further shows that the negative reciprocity effect on firms' profits from h-consumers is lower in *i2*-treatments than in *i1*-treatments. Specifically, the

[9] The ANOVA results show no significant difference between treatments *wpd_i1* and *spd_i1* (Prob $>$ F = 0.20) but between treatments *wpd_i2* and *spd_i2* (Prob $>$ F = 0.08).

[10] The ANOVA results show that the difference between treatment *wpd_i1* and treatment *spd_i1* with regard to firms' overall profits is not significant (Prob $>$ F = 0.358).

negative reciprocity effect decreases from 12.74 percentage points in treatment *wpd_i1* to 7.43 percentage points in treatment *wpd_i2* and from 24.17 percentage points in treatment *spd_i1* to 19.62 percentage points in treatment *spd_i2*.[11] This suggests that h-consumers punish firms less for charging them a higher price, when they know that they have a higher income.

The positive reciprocity effect on firms' profits from l-consumers is also lower in *i2*-treatments than in *i1*-treatments, suggesting that l-consumers reward firms less for charging them a lower price when they know that they have a lower income. In particular, the positive reciprocity effect decreases from 9.83 percentage points in treatment *wpd_i1* to 7.52 percentage points in treatment *wpd_i2*, and from 10.76 percentage points in treatment *spd_i1* to 6.97 percentage points in treatment *spd_i2*.[12]

So, the negative as well as the positive reciprocity effect are lower in *i2*-treatments than in *i1*-treatments. Since the negative reciprocity decreases by more, *firms' overall profits are less negatively affected in i2-treatments* than in *i1*-treatments. In particular, the negative reciprocity effect on firms' overall profits decreases from 2.01 percentage points in treatment *wpd_i1* to 0.25 percentage points in treatment *wpd_i2*, and from 7.20 percentage points in treatment *spd_i1* to 6.20 percentage points in treatment *spd_i2*.[13] Thus, firms seem to obtain higher profits when consumers know that disadvantaged consumers have a higher income. However, the differences in reciprocity effects between *i2*- and *i1*-treatments are not significant.

Interestingly, the estimations reported in Table 3.4 also show that firms obtain significantly higher profits from *female* consumers, suggesting that female h-consumers punish significantly less and that female l-consumers reward significantly more. Similarly, firms obtain significantly higher profits from h-consumers who rather confirm in the questionnaire that they generally trust others.

[11] The ANOVA results show that the differences between *i1*- and *i2*-treatments with regard to firms' profits from h-consumers are not significant. For *wpd* we obtain Prob > F = 0.384, and for *spd* we obtain Prob > F = 0.729.

[12] The ANOVA results show that the differences between *i1*- and *i2*-treatments with regard to firms' profits from l-consumers are not significant. For *wpd* we obtain Prob > F = 0.864, and for *spd* we obtain Prob > F = 0.761.

[13] The ANOVA results show that the differences in means between the *i1*- and *i2*-treatments with regard to firms' overall profits are not significant. For *wpd* we obtain Prob > F = 0.640 and for *spd* we obtain Prob > F = 0.901.

3.4.3 Self/Self Price Comparisons Over Time

The experiment was designed such that not only self/other price comparisons arise but also self/self price comparisons over time. Clearly, self/self price comparisons over time might influence consumers' price fairness perceptions. When consumers were previously charged a lower price, they might feel entitled to this lower price and perceive subsequent higher prices as less fair/more unfair. Likewise, when they were previously charged a higher price, they might perceive subsequent lower prices as more fair/less unfair. In this section, we account for self/self comparisons by including dummy variables $p_{-1|low}$ and $p_{-1|high}$ in the regression of the reciprocity impact on firms' profits. These dummy variables indicate whether a consumer was charged a lower price in the previous round ($p_{-1|low} = 1$), whether a consumer was charged a higher price in the previous round ($p_{-1|high} = 1$), or none of these cases for observations in the first round ($p_{-1|low} = p_{-1|high} = 0$). Results are shown in Table 3.6 in Appendix A.

We find no statistically significant influence of self/self price comparisons over time. This supports an argument made by Xia, Monroe, and Cox (2004) that self/other comparisons are likely to have a greater effect on consumers' price fairness judgments than self/self comparisons. However, sign and size of the coefficients $p_{-1|low}$ and $p_{-1|high}$ suggest that firms obtain lower profits when consumers were previously charged a lower price, implying that a lower price in the previous round causes consumers to punish more or reward less in the actual round. By contrast, firms seem to obtain higher profits when consumers were previously charged a higher price, implying that a higher price in the previous round causes consumers to punish less or reward more in the actual round. These effects are, however, not significant.

3.4.4 Correlation Between Consumers' Behavior in the Experiment and in the Trust Game

Following the main part of the experiment, subjects were asked to play the standard trust game. In the trust game responders positively reciprocate by rewarding a sender based on both the gains from exchange to the responder as well as the responder's belief about the intention motivating the action of the sender. We assigned consumers the role of responders in order to be able to test whether their (reciprocal) behavior in the experiment is correlated to their reciprocal behavior in the trust game.

Firms, in the role of senders, received 20 EP and could send any amount between 0 and 10 EP to *two* consumers.[14] The amount they sent was tripled by the experimenter.

[14]The conversion rate changed to 1 EP = 10 euro-cent.

Thereupon, consumers had a contingent choice (strategy method of elicitation) to send any amount, they potentially received, back to the firms. They were informed that their decision only affected the outcome of the firms if the firms opted to give them that choice.

We measure the strength of the linear relationship between consumers' reciprocal behavior in the experiment and in the trust game using the Pearson Correlation Coefficient. As an indicator for consumers' reciprocal behavior in the experiment, we use the percentage of material payoff consumers are willing to give up to punish or reward firms ($\Delta\pi_h$ and $\Delta\pi_l$). And as an indicator for consumers' reciprocal behavior in the trust game we use the standard deviation of amounts that consumers choose to send back to firms dependent on the firms' choices (*std_tg*).[15] Table 3.5 shows the results.

Table 3.5: Pearson Correlation Coefficients

	$\Delta\pi_h(npd)$	$\Delta\pi_h(wpd)$	$\Delta\pi_h(spd)$	$\Delta\pi_l(npd)$	$\Delta\pi_l(wpd)$	$\Delta\pi_l(spd)$
std_tg	0.004 (0.975)	-0.256** (0.048)	-0.320** (0.012)	-0.072 (0.586)	-0.158 (0.229)	-0.146 (0.262)

Notes: The table reports Pearson Correlation Coefficients between subjects' reciprocal behavior in the experiment (captured by $\Delta\pi$) and in the trust game (captured by std_tg). Significance levels are reported in parentheses; *** denotes p<0.01, ** denotes p<0.05, and * denotes p<0.1.

Since consumers presumably do not behave reciprocally in the non-price discrimination treatments *npd* we find no statistically significant linear correlations between $\Delta\pi_h$ (*npd*) and $\Delta\pi_l$ (*npd*) on the one hand and *std_tg* on the other hand. In contrast, we do find statistically significant linear correlations between $\Delta\pi_h$ (*wpd*) and $\Delta\pi_h$ (*spd*) on the one hand and *std_tg* on the other hand, with a stronger relationship between $\Delta\pi_h$ (*spd*) and *std_tg* than between $\Delta\pi_h$ (*wpd*) and *std_tg*. This supports the result that h-consumers behave reciprocally under third degree price discrimination, and that they do so the more, the stronger the degree of price discrimination is. Unsurprisingly we find weaker and not significant linear correlations between $\Delta\pi_l$ (*wpd*) and $\Delta\pi_l$ (*spd*) on the one hand and *std_tg* on the other hand, since we already found a weaker and not significant positive reciprocity effect on firms' profits from l-consumers.

[15]The results would be similar when we used the average amount that consumers send back as indicator for consumers' reciprocal behavior in the trust game.

3.5 A Simple Model of Reciprocity in the Context of Third Degree Price Discrimination

Our experimental results suggest that consumers exhibit social preferences when being price discriminated. Depending on the price they are charged compared to other consumers they either regard price discrimination as fair or unfair. In reaction to perceived price fairness or unfairness consumers behave reciprocally, raising or lowering firms' profits by raising or lowering their demand. In this section, we formalize *reciprocity* in the context of third degree price discrimination, adapting the framework developed by Falk and Fischbacher (2006).[16] Our goal is to explain formally (i) how consumers react to price discrimination when they are charged higher and when they are charged lower prices than other consumers, (ii) how their reactions alter when the price differential increases, (iii) how their reactions alter when they get to know not only the price but also the income of other consumers, and most importantly (iv) to what extent the profitability of third degree price discrimination is affected by consumers' fairness concerns.

An extensive literature studies *reciprocity* of economic agents in decision making (for surveys see, e.g., Sobel, 2005 or Fehr and Schmidt, 2006). In Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) fairness evaluations are based upon interpersonal payoff comparisons. Players reciprocate in order to reduce inequity in payoffs. On the other hand, in Rabin (1993) and Dufwenberg and Kirchsteiger (2004) reciprocity is driven by intentions and not necessarily as a desire to reduce inequity.[17] Players positively reciprocate fair intentions and negatively reciprocate unfair ones. In Falk and Fischbacher (2006) both concepts are combined.

Our framework differs from the previous ones in one dimension. While the previous models consider the case in which two players evaluate the intention and/or the outcome of the other player's action and react reciprocally toward the other, our model involves three players, a firm and two consumers. A consumer judges the *intention* behind a firm's pricing decision by its *outcome*, which is not the difference between her material payoff and that of the firm but rather the difference between the material payoff she can obtain by purchasing from the firm and the material payoff she believes the other

[16] For simplification we neglect in our framework that consumers might also be motivated to reduce the difference between their material payoff and that of the firm by reducing their demand (see, e.g., Fehr and Schmidt, 1999), and further that consumers might be willing to increase social surplus, i.e., the joint payoff with the firm, by enhancing their demand (see, e.g., Charness and Rabin, 2003).

[17] Fehr and Schmidt (1999) emphasize that the outcome of an action to some extent also reveals its underlying intention.

consumer can obtain. While in Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) players reciprocate toward a second player to *reduce* inequity between them, in our model players reciprocate toward a second player *because of* inequity between them and a third player, which is caused by the second player. Thus, in our model reciprocity is driven by intentions and intentions are evaluated by caused inequity.

We could adapt any of the intention-based reciprocity models to explain our experimental results on reciprocity in the context of third degree price discrimination. However, for our purposes the Falk and Fischbacher (2006) model is particularly suitable as it considers both the role of intentions *and* inequity aversion as sources of reciprocal behavior.

As in the experiment we consider a one stage game between three agents: a firm, an h-consumer and an l-consumer.[18] Firm F moves first, charging prices p_h and p_l to consumers h and l, whereupon the consumers purchase quantities q_h and q_l from F. Let h-consumers have a higher income than l-consumers, that is, $I_h > I_l$, and let their demand be less price elastic. According to standard theory F then optimally sets $p_h > p_l$, provided price discrimination is feasible.

We account for reciprocity by allowing the utility of consumer i ($i \neq j,\ i,j = h,l$) when choosing quantity q_i to depend not only on her *material payoff* $\pi_i(q_i)$, as standard theory would suggest, but also on her *reciprocity utility.* The reciprocity utility consists of a reciprocity parameter $\rho_i \geq 0$ which measures consumer i's *individual reciprocal preferences*, a kindness term $\varphi_i(\cdot)$ which measures the *kindness* of the firm's pricing decision as perceived by consumer i, and a reciprocation term $\sigma_i(q_i)$ which measures consumer i's *reciprocal response* to the perceived kindness.

Consumer i's utility when choosing quantity q_i is defined as:

$$U_i(q_i) := \pi_i(q_i) + \rho_i\ \varphi_i(\cdot)\ \sigma_i(q_i).$$

The higher the *individual reciprocity parameter* ρ_i, the more weight puts consumer i on her reciprocity utility as compared to her material payoff. In the following, we derive the kindness term $\varphi_i(\cdot)$ and the reciprocation term $\sigma_i(q_i)$.

Falk and Fischbacher (2006) define the *kindness term* as $\varphi_i := \upsilon_i \Delta_i$, where the *intention factor* υ_i captures whether individuals have choice alternatives or not. For instance, if F chooses a price pair that is disadvantageous for consumer i, consumer i will perceive F's pricing decision as less unkind if F could only choose between price

[18]The model by Falk and Fischbacher (2006) actually considers n-stages. To provide an understanding of the nature of reciprocity in the context of third degree price discrimination we concentrate on the one stage game.

pairs that are disadvantageous for consumer i, compared to when F could also have chosen price pairs that are advantageous for consumer i. A restriction in F's choice alternatives is captured by $\upsilon_i < 1$. Since in the context of third degree price discrimination consumers are likely to perceive firms' pricing decisions as fully intentional as firms usually have unrestricted choice alternatives when deciding on prices, we assume that $\upsilon_i = 1$.[19]

Consequently, the kindness of F's pricing decision as perceived by consumer i is measured by the *outcome of F's pricing decision* (Δ_i) only. The outcome of F's pricing decision is the *material payoff that consumer i can obtain* by purchasing from F compared to a *reference standard.* This reference standard could be the payoff that F obtains by selling to consumer i (internal self/other comparison). However, consumers usually do not know which payoff firms obtain. Other consumers are also more comparable than firms as other consumers have more similarities. Therefore, we assume that consumer i compares the material payoff that she can obtain being priced p_i with the material payoff that she believes consumer j can obtain being priced p_j (external self/other comparison).

The outcome term Δ_i is defined as:

$$\Delta_i := \pi_i(p_i, \overline{q}_i) - \pi_j(p_j, \overline{q}_j) + \gamma_i(I_i - I_j),$$

where $\pi_i(p_i, \overline{q}_i)$ denotes consumer i's maximally achievable material payoff, $\pi_j(p_j, \overline{q}_j)$ denotes consumer i's belief about consumer j's maximally achievable material payoff, I_i denotes consumer i's income, I_j denotes consumer i's belief about consumer j's income, and $\gamma_i \in [0, 1)$ captures by how much consumer i accounts for the believed income difference in her price fairness evaluation.

For simplicity we assume, when consumer i receives no information to the contrary, she believes that consumer j has identical characteristics, i.e., she believes that $\pi_j(p_j, \overline{q}_j) = \pi_i(p_j, \overline{q}_i)$ and $I_j = I_i$. This implies that consumer i forms her price fairness judgment by comparing the material payoff that she can maximally achieve being priced p_i with the material payoff she could maximally achieve when she was priced p_j. The outcome term Δ_i will then be negative and she will regard F's pricing decision as unkind, when her price is higher than that of consumer j ($p_i > p_j$). If, on the other hand,

[19] In the experiment, however, firms could only choose between *two* price pairs, and one of these price pairs, *apd*, was clearly inferior to the other. If consumers anticipated that firms had a restricted choice, this would have caused them to perceive firms' pricing decisions as less unkind/kind (since $\upsilon_i < 1$ then) and thus to reciprocate less. The reciprocity effects we observed in the experiment would then be lower than the reciprocity effects to be observed in real market environments.

she is charged a lower price than consumer j, the outcome term Δ_i will be positive and she will regard F's pricing decision as kind.

The higher the price consumer i is charged compared to consumer j, the lower she will believe is her maximally achievable material payoff compared to that of consumer j, and thus she will regard F's pricing decision as more unkind (Δ_i decreases $p_i - p_j > 0$). Vice versa, the lower the price consumer i is charged compared to consumer j, the higher she will believe is her maximally achievable material payoff compared to that of consumer j, and thus she will regard F's pricing decision as more kind (Δ_i increases with $p_i - p_j < 0$). This means, the disadvantaged h-consumers will perceive F's pricing decision as more unfair when the price differential increases, while the advantaged l-consumers will perceive F's pricing decision as more fair when the price differential increases.

We allow for implicit price comparisons by incorporating $\gamma_i(I_i - I_j)$ into the outcome term. The term $(I_i - I_j)$ denotes the difference between consumer i's income and consumer i's belief about consumer j's income. The parameter γ_i captures the extent to which consumer i accounts for the believed income difference in her price fairness judgment. If consumer i does not know consumer j's income, she believes that consumer j has the same income, so that the term $\gamma_i(I_i - I_j)$ cancels out. The correction of consumer i will then be zero. If, however, consumer i knows consumer j's income, *she will accept a higher price more when she has a higher income* than consumer j (i.e., Δ_i will be higher if $I_i - I_j > 0$), and *she will feel more entitled to a lower price when she has a lower income* than consumer j (i.e., Δ_i will be lower if $I_i - I_j < 0$). The model therefore predicts that with the income information h-consumers will accept a higher price more, whereas l-consumers will accept a lower price less. The larger the income difference, the more it will influence consumers' price fairness judgments.

We now derive the *reciprocation term* σ_i, which captures how consumer i alters her quantity choice (q_i) and thus F's payoff in response to the experienced kindness.

The reciprocation term σ_i *is defined as:*

$$\sigma_i(q_i) := \begin{cases} \pi_f(q_i) - \pi_f(\overline{q}_i) & \text{if } \Delta_i \leq 0 \\ \alpha_i \left[\pi_f(q_i) - \pi_f(\overline{q}_i)\right] & \text{if } \Delta_i > 0 \end{cases},$$

where $\pi_f(q_i)$ denotes F's material payoff when consumer i purchases q_i. Further, $\overline{q}_i$ denotes the quantity choice that would maximize consumer i's material payoff, and $\alpha_i \in [0, 1)$ denotes an individual discount factor for positive reciprocity.

When $\Delta_i < 0$ holds and consumer i perceives F's pricing decision to be unfair, she can increase her utility by negatively reciprocating, i.e., by choosing a lower quantity than $\overline{q}_i$, thereby lowering F's profit. By contrast, when $\Delta_i > 0$ holds and consumer i perceives F's pricing decision to be fair, she can increase her utility by positively reciprocating, i.e., by choosing a higher quantity than $\overline{q}_i$, thereby enhancing F's profit. While in the former case the reciprocation term σ_i will be negative, in the latter case it will be positive. The more unkind consumer i perceives F's pricing decision, i.e., the lower Δ_i, the more will consumer i be able to increase her utility by negatively reciprocating, choosing a lower quantity than $\overline{q}_i$. Vice versa, the more kind consumer i perceives F's pricing decision, i.e. the more positive Δ_i, the more will consumer i be able to increase her utility by positively reciprocating, choosing a higher quantity than $\overline{q}_i$.

The empirical evidence from our experiment suggests that consumers count negative deviations from the reference outcome more than positive deviations. Thus, when consumer i's individual reciprocity parameter ρ_i is positive, her utility loss from a disadvantageous price differential is presumably larger than her utility gain from an equally sized advantageous price differential. We account for that by incorporating α_i into the reciprocation term, which discounts the utility that consumer i can obtain by positively reciprocating. This implies that negative consumer reactions will be stronger than positive consumer reactions.

The assumption that $\alpha_i < 1$ has important implications for the profitability of third degree price discrimination. It implies that the profitability of third degree price discrimination will be *negatively* affected by consumers' reciprocal reactions, especially when Δ_h is high. That is, the more unfair the disadvantaged h-consumers perceive F's pricing decision the more negatively profits will be affected due to reciprocity. To which extent h-consumers perceive F's pricing decision as unfair depends first of all on the price differential. *The higher the price h-consumers are charged compared to l-consumers, the stronger will be h-consumers' negative reactions compared to l-consumers' positive reactions, averting the profitability of third degree price discrimination.* Thus, the negative effect on the profitability of third degree price discrimination will increase with the price differential $|p_i - p_j|$.

Further, the extent to which h-consumers perceive F's pricing decision as unfair depends on the disclosure of income information. If h-consumers know that they have a higher income than l-consumers, they perceive a higher price as less unfair and in reaction punish less. L-consumers, on the other hand, perceive a lower price as less fair and in reaction reward less. Due to the assumption that $\alpha_i < 1$, the positive acceptance effect on the side of h-consumers will have stronger profit implications than the negative entitlement effect on the side of l-consumers. Put differently, when consumers are informed about other consumers' income, the gains that F obtains from h-consumers due to increased acceptance of a higher price will be larger compared to the losses that F incurs from l-consumers due to decreased appreciation of a lower price. Thus, the negative effect on the profitability of third degree price discrimination will be lower when consumers are informed about other consumers' income (and the consumer group with the less price elastic demand has a higher income).

As argued by Rotemberg (2011), consumers might perceive F's *intention behind price discrimination* as more benevolent when they obtain the income information. H-consumers will then perceive F's pricing decision as less unfair (in line with the positive acceptance effect put forward in our model), while l-consumers will then perceive F's pricing decision as more fair (in contrast with the negative entitlement effect put forward in our model). The results of our experiment suggest that even if l-consumers perceived F's intention behind price discrimination to be more benevolent when they obtain the income information, the negative entitlement effect (as described in our model) prevails. This is because the positive reciprocity effect on firms' profits from l-consumers is *lower* in *i2*-treatments than in *i1*-treatments.

3.6 Concluding Remarks

We conducted an experimental study which showed that the profitability of third degree price discrimination is negatively affected by consumers' fairness concerns. The higher the price differential that firms charge, the stronger are negative reactions by disadvantaged consumers compared to positive reactions by advantaged consumers. As a consequence, firms obtain higher profits by charging a weaker price differential than the one predicted to be optimal under standard theory. Furthermore, we found that price discriminating firms obtain higher profits when consumers are informed about other consumers' income. This is because the disadvantaged consumers, who have a higher income in our setting, react less negative and the advantaged consumers react less positive. Overall, the negative reactions attenuate compared to the positive reactions.

We explained these results within a theoretical framework, that is based on Falk and Fischbacher (2006). The model stipulates the following. When consumers have no reciprocal preferences, then, regardless of whether they perceive firms' pricing decisions as fair or not, they will optimally choose the quantity that maximizes their material payoff. If, on the other hand, they have reciprocal preferences, then, depending on whether they perceive firms' pricing decisions as fair or unfair, they will optimally choose a higher or lower quantity than the quantity that maximizes their material payoff, thereby either rewarding or punishing the firms. Whether consumers regard firms' pricing decisions as fair or unfair depends on the price they are charged compared to other consumers. If they are charged a higher price they will generally regard the pricing decision as unfair. However, when they know that they have a higher income than the other consumers they will regard it as less unfair. Vice versa, consumers who are charged lower prices than other consumers will regard the pricing decision as fair and less so when they know that they have a lower income. Negative consumer reactions will be stronger than positive consumer reactions, in particular when the price differential is large. Thus, the negative reciprocity effect on the profitability of third degree price discrimination will increase with the price differential and it will be lower when consumers are informed about other consumers' income.

Future research should explore the impact of consumers' fairness concerns on the profitability of third degree price discrimination more broadly and consider a number of other individual and contextual factors, such as short-term versus long-term customer/seller relationships, or consumers' switching options. Consumers in long-term customer/seller relationships might feel entitled to lower prices and therefore punish firms more when being negatively price discriminated. Further, the adverse effects on the profitability of third degree price discrimination might aggravate when consumers do not suffer high losses when they switch to other firms.

A limitation of our study is that it focuses on short-run profit implications. It would be very interesting to also explore the long-run profit implications. It could well be possible that disadvantaged consumers only initially perceive higher prices as unfair and accept them *over time*, so that higher price differentials become more profitable over time.

3.7 Appendix A: Regression Results

Table 3.6: RECIPROCITY IMPACT ON FIRMS' PROFITS CONSIDERING SELF/SELF COMPARISONS

VARIABLES	$\Delta\pi_f$ (h-consumers)	$\Delta\pi_f$ (l-consumers)
spd_i1	-20.97**	7.936
	(8.187)	(10.06)
spd_i2	-15.48*	2.864
	(9.009)	(9.543)
gender_h	16.20**	
	(6.912)	
trustgeneral_h	-15.82**	
	(6.333)	
apd_h	15.92*	
	(8.818)	
$p^h_{-1\|low}$	-6.857	
	(10.86)	
$p^h_{-1\|high}$	2.048	
	(5.825)	
gender_l		11.77*
		(6.100)
trustgeneral_l		-3.798
		(3.547)
apd_l		0.941
		(9.560)
$p^l_{-1\|low}$		-7.022
		(12.77)
$p^l_{-1\|high}$		-1.416
		(8.249)
Constant	11.62	-14.38
	(8.956)	(11.67)
Observations	90	90
R-squared	0.228	0.070

Notes: The table reports coefficients of OLS regressions. The dependent variables are the percentage differentials between firms' actual profits (from h- or l-consumers) and firms' profits (from h- or l-consumers) under standard theory. The treatment *npd_i1* is the baseline. Robust standard errors are reported in parentheses. The regressions are clustered by consumers and therefore control for individual fixed effects. *** denotes p<0.01, ** denotes p<0.05, and * denotes p<0.1.
Here, we only include the treatment variables *spd_i1* and *spd_i2* and compare them to the treatment *npd_i1*. This is because the experiment was designed in such a way that in the first rounds only *npd* and *spd* treatments came up, not *wpd* treatments. Thus, we do not have observations for *wpd_i1* or *wpd_i2* in the first rounds, where $p_{-1|low} = p_{-1|high} = 0$. So, if we included the treatment variables *wpd_i1* and *wpd_i1* in the regressions, we would not have a base category against which the dummy variables $p_{-1|low}$ and $p_{-1|high}$ would be assessed, which would lead to perfect multicollinearity.

3.8 Appendix B: Translation of the Instructions

General Information about the Course of the Experiment

This experiment analyses economic behavior in markets. During the experiment you and the other participants can earn money by making decisions. The amount of money you earn depends on your own decisions as well as on the decisions of the other participants and is determined by the rules that will be explained in the following in detail.

The entire experiment takes about 60 minutes. At the beginning you will receive detailed instructions. If you have questions after these instructions, please raise your hand. The experimenter will then come to you and answer your questions privately. Each participant is given a number by which he may be identified during the course of the experiment. Due to linguistic simplicity, we only use male terms in these instructions. These are supposed to be understood gender neutral.

Anonymity

The main part of the experiment consists of 3 rounds. At the beginning of each round you will be randomly assigned to a group consisting of 3 participants. You will not learn the identity of the participants that you are in a group with, neither during nor after the experiment. Also other participants will not learn about your role, your decisions and how much you earned. We will analyze the data from the experiment only **anonymously**. At the end of the experiment, you must sign an acknowledgment of the receipt of payoff. But this is only for the accounting.

Groups

At the beginning of the experiment you are randomly assigned the role of a seller or the role of a buyer. This role you will keep over all three rounds. At the beginning of each round you will be randomly assigned to a group, consisting of one seller and two buyers. In each subsequent round, your group will be randomly re-assembled.

Payoffs

Your payoffs will be paid to you at the end of the experiment. We randomly choose the result of one of the three rounds of the main part of the experiment. Following the main part of the experiment, you will be asked to make further decisions and to provide additional information. For this, you will receive additional payment.

During the experiment the currency is not euros but experimental points (EP). Your earnings in the course of the experiment will be calculated in EP. At the end of the experiment, all EP that you earned will be converted into euros. The conversion rate is: 1 experimental point = 1 euro-cent.

The payoff of a seller arises from the sale of a good, and the payoff of a buyer arises from the purchase of that good. A **seller** must choose in each round a pair of prices at which he wants to sell the good to the two buyers. He can offer the two buyers different prices. **A seller's payoff is equal to the quantities that the buyers purchase, multiplied by the prices he has set.** So depending on how much the buyers buy to the prices he set his payoff rises or falls. A seller receives an extra payment from the experimenter of 500 EP if he sells a positive quantity to a buyer. A positive quantity means a quantity greater than zero.

Consider the following example. A seller sets buyer 1 a price of 30 EP per unit, and buyer 2 a price of 20 EP per unit. Buyer 1 buys 10 units, and buyer 2 buys 15 units. The total payoff of the seller in this example is:

$$\underbrace{30\text{ EP}}_{\text{price 1}} \cdot \underbrace{10}_{\text{quantity 1}} + \underbrace{500\text{ EP}}_{\text{extra payment 1}} + \underbrace{20\text{ EP}}_{\text{price 2}} \cdot \underbrace{15}_{\text{quantity 2}} + \underbrace{500\text{ EP}}_{\text{extra payment 2}} = \underbrace{1600\text{ EP}}_{\text{payoff}}$$

Buyers receive at the beginning of each round a budget from the experimenter which they can use to buy the good. Each buyer can only buy exactly as many units of the good as he can afford with his budget. A table, as in the example below, shows a buyer what his benefits and his expenditures are when purchasing a certain quantity. **The payoff of a buyer is the difference between benefits and expenditures of the quantity he chooses, plus his budget.**

Consider the following example, which refers to the table below. A seller sets a price of 10 EP. Buyer 1 receives a budget of 100 EP, and may thus buy a maximum of 10 units. He chooses the quantity 8, so that his benefit from purchasing the goods is 96 EP and his expenditure 80 EP. Hence, his payoff in selecting quantity 8 is:

$$\underbrace{96\text{ EP}}_{\text{benefit}} - \underbrace{80\text{ EP}}_{\text{expenditure}} + \underbrace{100\text{ EP}}_{\text{budget}} = \underbrace{116\text{ EP}}_{\text{payoff}}$$

Buyer 1's expenditures correspond to the set price multiplied by the quantity he purchases (see column 3). If the seller set a higher price than 10 EP, buyer 1's expenditures per quantity would increase and thus his payoff would decrease.

The right column of the table indicates which payoff the seller obtains from selling to buyer 1. Note that the seller's total payoff is composed of the payoff from the sale to buyer 1 and of the payoff from the sale to buyer 2.

Quantity	**Benefit** from purchasing the good	**Expenditure** (price · quantity)	**Payoff** of buyer 1 (budget + benefit - expenditure)	Sellers' payoff from selling to buyer 1 (price · quantity + extra payment)
0	0	100	0	0
1	19	10	109	510
2	36	20	116	520
3	51	30	121	530
4	64	40	124	540
5	75	50	125	550
6	84	60	124	560
7	91	70	121	570
8	**96**	**80**	**116**	**580**
9	99	90	109	590
10	100	100	100	600

Questions

We would like to ask you to answer the following two questions. Suppose you are assigned the role of a buyer and get the information shown in the table above.
o What would be your payoff if you chose as quantity 6?
o What payoff would the seller obtain from selling to you if you chose as quantity 4?

After the main part of the experiment

The conversion rate is now: 1 experimental point = 10 euro-cent. You are assigned as sender or receiver in a group consisting of one sender and two receivers. The sender receives 20 EP from the experimenter. Of these 20 EP the sender can send between 0 and 10 EP to each receiver. The amount must be the same for both receivers. The experimenter will triple the amount sent, which we denote y. The receivers can then return any amount between 0 and $3 \cdot y$ EP to the sender.

Consider the following example. The sender sends the receivers the amount 5 EP. So, y equals 5 in this example. The experimenter triples the sent amount y. The receivers can then return any amount between 0 and $3 \cdot y$ EP, that is, any amount between 0 and 15 EP, to the sender.

Bibliography

[1] Abito, Jose M., and Julian Wright (2008), "Exclusive Dealing with Imperfect Downstream Competition," *International Journal of Industrial Organization*, 26, 227-246.

[2] Addanki, Sumanth, and Alan J. Daskin (2008), "Patent Settlement Agreements," book chapter 85, published in: *Issues in Competition Law and Policy*, American Bar Association.

[3] Aghion, Philippe, and Patrick Bolton (1987), "Contracts as a Barrier to Entry," *American Economic Review*, 77, 388-401.

[4] Alba, Joseph W., Susan M. Broniarczyk, Terence A. Shimp, and Joel E. Urbany (1994), "The Influence of Prior Beliefs, Frequency Cues, and Magnitude Cues on Consumers' Perceptions of Comparative Price Data," *Journal of Consumer Research*, 21, 219-235.

[5] Anderson, Eric T., and Duncan I. Simester (2008), "Does Demand Fall When Customers Perceive That Prices Are Unfair? The Case of Premium Pricing for Large Sizes," *Marketing Science*, 27(3), 492-500.

[6] Armstrong, Mark (2006), "Recent Developments in the Economics of Price Discrimination," in *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress of the Econometric Society*, eds. R. Blundell, W. Newey, and T. Persson, Cambridge University Press, 97-141.

[7] Arrow, Kenneth J. (1961), "Economic Welfare and the Allocation of Resources for Invention," in: Nelson, R. (ed.): *The Rate and Direction of Inventive Activity: Economic and Social Factors*, Princeton University Press, 609-626.

[8] Balto, David (2004), "Bringing Clarity to the Patent Settlement Debate: Judge Posner's Asahi Decision," *Biotechnology Law Report*, 23(2), 168-171.

[9] Bernheim, B. Douglas, and Michael D. Whinston (1998), "Exclusive Dealing," *Journal of Political Economy*, 106, 64-103.

[10] Blair, Roger D., and Thomas F. Cotter (2002), "Are Settlements of Patent Disputes Illegal Per Se?," *The Antitrust Bulletin*, 47, 491-539.

[11] Bolton, Gary E., and Axel Ockenfels (2000), "A Theory of Equity, Reciprocity and Competition," *American Economic Review*, 90, 166-193.

[12] Bolton, Lisa E., Luk Warlop, and Joseph W. Alba (2003), "Consumer Perceptions of Price (Un)Fairness," *Journal of Consumer Research*, 29, 474-91.

[13] Bonanno, Giacomo, and John Vickers (1988), "Vertical Separation," *Journal of Industrial Economics*, 36, 257-265.

[14] Borenstein, Severin (1991), "Selling Costs and Switching Costs: Explaining Retail Gasoline Margins," *RAND Journal of Economics*, 22(3), 354-369.

[15] Bork, Robert H. (1978), *The Antitrust Paradox*, New York: Basic Books.

[16] Brodley, Joseph F., and Ching-to Albert Ma (1993), "Contract Penalties, Monopolizing Strategies, and Antitrust Policy," *Stanford Law Review*, 45, 1161-1213.

[17] Bulow, Jeremy (2003), "The Gaming of Pharmaceutical Patents," *Stanford Research Paper* No. 1804.

[18] Busse, Meghan, and Marc Rysman (2005), "Competition and Price Discrimination in Yellow Pages Advertising," *RAND Journal of Economics*, 36(2), 378-390.

[19] Büyükkurt, B. Kemal (1986), "Integration of Serially Sampled Price Information: Modeling and Some Findings," *Journal of Consumer Research*, 13, 357-373.

[20] Campbell, Margaret C. (1999), "Perceptions of Price Unfairness: Antecedents and Consequences," *Journal of Marketing Research*, 36, 187-99.

[21] Charness, Gary, and Matthew Rabin (2002), "Understanding Social Preferences with Simple Tests," *The Quarterly Journal of Economics*, 117, 817-869.

[22] Crane, Daniel A. (2002), "Exit Payments in Settlement of Patent Infringement Lawsuits: Antitrust Rules and Economic Implications," *Florida Law Review*, 54, 747-798.

[23] Dholakia, Utpal M., and Itamar Simonson (2005), "The Effect of Explicit Reference Points on Consumer Choice and Online Bidding Behavior," *Marketing Science*, 24(2), 194-205.

[24] Dufwenberg, Martin, and Georg Kirchsteiger (2004), "A Theory of Sequential Reciprocity," *Games and Economic Behavior*, 47, 268-298.

[25] European Commission, Competition DG, (2008), "Pharmaceutical Sector Inquiry—Preliminary Report," *DG Competition Staff Working Paper*, available online at: http://ec.europa.eu/competition/sectors/pharmaceuticals/inquiry/ preliminary_report.pdf.

[26] European Commission, Competition DG, (2009), "Pharmaceutical Sector Inquiry—Final Report," available online at: http://ec.europa.eu/competition/ sectors/pharmaceuticals/inquiry/staff_working_paper_part1.pdf.

[27] European Commission, Competition DG, (2010), "1^{st} Report on the Monitoring of Patent Settlements," available online at: http://ec.europa.eu/competition/ sectors/pharmaceuticals/inquiry/patent_settlements_report1.pdf.

[28] Falk, Armin, and Urs Fischbacher (2006), "A Theory of Reciprocity," *Games and Economic Behavior*, 54, 293-315.

[29] Farrell, Joseph, and Robert P. Merges (2004), "Incentives to Challenge and Defend Patents: Why Litigation Won't Reliably Fix Patent Office Errors and Why Administrative Patent Review Might Help," *Boalt Working Papers in Public Law.*

[30] Federal Trade Commission (2002), "Generic Drug Entry Prior to Patent Expiration," available at http://www.ftc.gov/os/2002/07/genericdrugstudy.pdf.

[31] Fershtman, Chaim, and Kenneth L. Judd (1987), "Equilibrium Incentives in Oligopoly," *American Economic Review*, 77, 927-940.

[32] Fehr, Ernst, and Klaus M. Schmidt (1999), "A Theory of Fairness, Competition and Cooperation," *The Quarterly Journal of Economics*, 114(3), 817-868.

[33] Fehr, Ernst, and Klaus M. Schmidt (2006), "The Economics of Fairness, Reciprocity and Altruism - Experimental Evidence and New Theories," in: Handbook on the Economics of Giving, Reciprocity and Altruism, Vol. 1, Serge-Christophe Kolm, and Jean Mercier Ythier (ed.), Amsterdam: Elsevier, 615-691.

[34] Fischbacher, Urs (2007), "Z-Tree: Zurich Toolbox for Ready-Made Economics Experiments," *Experimental Economics*, 10, 171-178.

[35] Franciosi, Robert, Praveen Kujal, Roland Michelitsch, Vernon Smith, and Gang Deng (1995), "Fairness: Effect on Temporary and Equilibrium Prices in Posted-Offer Markets," *The Economic Journal*, 105, 938-950.

[36] Fumagalli, Chiara, and Massimo Motta (2006), "Exclusive Dealing and Entry, when Buyers Compete," *American Economic Review*, 96, 785-795.

[37] Goldberg, Pinelopi K., and Frank Verboven (2001), "The Evolution of Price Dispersion in the European Car Market," *The Review of Economic Studies*, 68(4), 811-848.

[38] Greiner, Ben (2004), "An Online Recruitment System for Economic Experiments," in K. Kremer and V. Macho, eds., Forschung und wissenschaftliches Rechnen 2003, GWDG Bericht 63, Ges. für Wiss. Datenverarbeitung, Göttingen, Germany, 79-93.

[39] Harhoff, Dietmar, Frederic M. Scherer, and Katrin Vopel (2003), "Citations, Family Size, Opposition and the Value of Patent Rights," *Research Policy*, 32, 1343-1363.

[40] Hemphill, C. Scott (2006), "Paying for Delay: Pharmaceutical Patent Settlement As a Regulatory Design Problem," *New York University Law Review* 1553.

[41] Hemphill, C. Scott (2009), "An Aggregate Approach to Antitrust: Using New Data and Rulemaking to Preserve Drug Competition," *Columbia Law Review*.

[42] Hemphill, C. Scott, and Mark A. Lemley (2011), "Earning Exclusivity: Generic Drug Incentives and the Hatch-Waxman Act," *Antitrust Law Journal*, forthcoming.

[43] Hovenkamp, Herbert, Mark Janis, and Mark A. Lemley (2003), "Anticompetitive Settlements of Intellectual Property Disputes," *Minnesota Law Review*, 87, 1719-1766.

[44] Huang, Jen-Hung, Ching-Te Chang, and Cathy Yi-Hsuan Chen (2005), "Perceived Fairness of Pricing on the Internet," *Journal of Economic Psychology*, 26, 343-361.

[45] Innes, Robert, and Richard J. Sexton (1994), "Strategic Buyers and Exclusionary Contracts," *American Economic Review*, 83, 566-584.

[46] Janiszewski, Chris, and Donald R. Lichtenstein (1999) "A Range Theory Account of Price Perception," *Journal of Consumer Research*, 25, 353-368.

[47] Kahneman, Daniel, Jack L. Knetsch, and Richard Thaler (1986a), "Fairness as a Constraint on Profit Seeking: Entitlements in the Market," *American Economic Review*, 74(4), 728-741.

[48] Kahneman, Daniel, Jack L. Knetsch, and Richard Thaler (1986b), "Fairness and the Assumptions of Economics," *Journal of Business*, 59(4), 285-300.

[49] Kitamura, Hiroshi (2010), "Exclusionary Vertical Contracts with Multiple Entrants," *International Journal of Industrial Organization*, 28, 213-219.

[50] Langenfeld, James, and Wenqing Li (2003), "Intellectual Property and Agreements to Settle Patent Disputes: The Case of Settlement Agreements with Payments from Branded to Generic Drug Manufacturers," *Antitrust Law Journal*, 70(3), 777-818.

[51] Leffler, Keith, and Cristofer Leffler (2004), "Efficiency Trade-Offs in Patent Litigation Settlements? Analysis Gone Astray," *University of San Francisco Law Review*, 39, 33-56.

[52] Lemley, Mark A. (2001), "Rational Ignorance at the Patent Office," *Northwestern University Law Review*, 95(4).

[53] Lemley, Marc A., and Carl Shapiro (2005), "Probabilistic Patents," *Journal of Economic Perspective*, 19(2), 75-98.

[54] Leslie, Phillip (2004), "Price Discrimination in Broadway Theater," *RAND Journal of Economics*, 35(3), 520-541.

[55] Lichtenstein, Donald R., and William O. Bearden (1989), "Contextual Influences on Perceptions of Merchant-Supplied Reference Prices," *Journal of Consumer Research*, 16, 55-66.

[56] Lichtenstein, Donald R., Peter H. Bloch, and William C. Black (1988), "Correlates of Price Acceptability," *Journal of Consumer Research*, 15, 243-252.

[57] Martins, Marielza O. B., and Kent B. Monroe (1994), "Perceived Price Fairness: A New Look at an Old Construct," *Advances in Consumer Research*, 21, 75-78.

[58] Mathewson, G. Frank, and Ralph A. Winter (1987), "The Competitive Effects of Vertical Agreements: Comment," *American Economic Review*, 77, 1057-1062.

[59] Nevo, Aviv, and Catherine Wolfram (2002), "Why Do Manufacturers Issue Coupons? An Empirical Analysis of Breakfast Cereals," *RAND Journal of Economics*, 33(2), 319-339.

[60] O'Rourke, Maureen, and Joseph F. Brodley (2003), "Antitrust Implications of Patent Settlements: An Incentives Modifying Approach," *Minnesota Law Review*, 87, 101-123.

[61] Rabin, Matthew (1993), "Incorporating Fairness into Game Theory and Economics," *American Economic Review*, 83, 1281-1302.

[62] Rasmusen, Eric B., J. Mark Ramseyer, and John S. Wiley, Jr. (1991), "Naked Exclusion," *American Economic Review*, 81, 1137-1145.

[63] Rotemberg, Julio J. (2005), "Customer Anger at Price Increases, Changes in the Frequency of Price Adjustment and Monetary Policy," *Journal of Monetary Economics*, 52, 829-852.

[64] Rotemberg, Julio J. (2011), "Fair Pricing," *Journal of the European Economic Association*, 9(5), 952-981.

[65] Posner, Richard A. (1976), *Antitrust Law: An Economic Perspective*, Chicago: University of Chicago Press.

[66] Salinger, Michael A., Pauline M. Ippolito, and Joel L. Schrag (2007), "Economics at the FTC: Pharmaceutical Patent Dispute Settlements and Behavioral Economics," *Review of Industrial Organization*, 31(2), 85-105.

[67] Schildkraut (2004), "Patent-Splitting Settlements and the Reverse Payment Fallacy," *Antitrust Law Journal*, 71(3), 1033-1068.

[68] Segal, Ilya R., and Michael D. Whinston (2000), "Naked Exclusion: Comment," *American Economic Review*, 90, 296-309.

[69] Shapiro, Carl (2003), "Antitrust Limits to Patent Settlements," *The RAND Journal of Economics*, 34(2), 391-411.

[70] Shor, Mikhael, and Richard L. Oliver (2006), "Price Discrimination Through Online Couponing: Impact on Likelihood of Purchase and Profitability," *Journal of Economic Psychology*, 27, 423-440.

[71] Simpson, John, and Abraham L. Wickelgren (2001), "The Use of Exclusive Contracts to Deter Entry," Federal Trade Comission, Bureau of Economics Working Papers: No. 241.

[72] Simpson, John, and Abraham L. Wickelgren (2007), "Naked Exclusion, Efficient Breach, and Downstream Competition," *American Economic Review*, 97, 1305-1320.

[73] Sobel, Joel (2005), "Interdependent Preferences and Reciprocity," *Journal of Economic Literature*, 43, 392-436.

[74] Spier, Kathryn E., and Michael D. Whinston (1995), "On the Efficiency of Privately Stipulated Damages for Breach of Contract: Entry Barriers, Reliance and Renegotiation," *RAND Journal of Economics*, 26, 180-202.

[75] Stefanidis, Christodoulos (1998), "Selective Contracts, Foreclosure, and the Chicago School View," *Journal of Law and Economics*, 41, 429-450.

[76] Stole, Lars A. (2007), "Price Discrimination and Competition," in: *Handbook of Industrial Organization*, 3(34), eds. M. Armstrong, and R. Porter (eds.), Elsevier, Amsterdam: North-Holland, 2221-2299.

[77] United States Patent and Trademark Office (2009), "Performance and Accountability Report," available at http://www.uspto.gov/about/stratplan/ ar/index.jsp.

[78] Urbany, Joel E., William O. Bearden, Dan C. Weilbaker (1988), "The Effect of Plausible and Exaggerated Reference Prices on Consumer Perceptions and Price Search," *Journal of Consumer Research*, 15, 95-110.

[79] Verboven, Frank (1996) "International Price Discrimination in the European Car Market," *RAND Journal of Economics*, 27(2), 240-268.

[80] Vives, Xavier (1999), *Oligopoly Pricing: Old Ideas and New Tools*, Cambridge, MA: MIT Press.

[81] Willig, Robert D., and John P. Bigelow (2004), "Antitrust Policy Toward Agreements that Settle Patent Litigation," *The Antitrust Bulletin*, 49(3), 655-698.

[82] Wright, Julian (2008), "Naked Exclusion and the Anticompetitive Accommodation of Entry," *Economics Letters*, 98, 107-112.

[83] Wright, Julian (2009), "Exclusive Dealing and Entry, when Buyers Compete: Comment," *American Economic Review*, 99, 1070-81.

[84] Xia, Lan, Kent B. Monroe, and Jennifer L. Cox (2004), "The Price Is Unfair! A Conceptual Framework of Price Fairness Perceptions," *The Journal of Marketing*, 68(4), 1-15.

[85] Yong, Jong-Say (1999), "Exclusionary Vertical Contracts and Product Market Competition," *Journal of Business*, 72, 385-406.

Lebenslauf

2008 - 2012	Doktorandenstudium, Volkswirtschaftslehre International Max Planck Research School for Competition and Innovation (IMPRS-CI)
2010	Research Fellow an der Columbia University, USA
2003 - 2008	Diplom, Volkswirtschaftslehre Ludwig-Maximilians-Universität München
2005 - 2006	Auslandsaufenthalt University of Warwick, England
1996 - 2003	Abitur Gymnasium Isernhagen
2000 - 2001	Auslandsaufenthalt Clayesmore Boarding School, England
1984	Geboren in Hannover