Technological Transition and Price Discrimination*

Max R. P. Grossmann

June 6, 2017

Abstract

We use microeconomic analysis, elementary game theory and simulation–based inference to derive optimal pricing strategies under five different regimes of price discrimination. Using principles of quasilinear utility, we present the three classic degrees of price discrimination and two more recent variants: intertemporal price discrimination and price discrimination by purchase history. We give examples for applications of price discrimination and show how modern technology enables firms to price discriminate. Our use of simulation–based inference demonstrates that it is an attractive method for approaching problems in the field of managerial economics. Finally, we conclude.

1 Introduction

In a low-interest environment, German banks are resorting to an extraordinary measure: Charging *some* customers for drawing cash at an automated teller machine. Volksbanken in the south of Hesse charge a fee for drawing money after closing time and in Offenbach, customers even have to pay from 1 p.m. until 2 p.m.[1] Other customers pay no fee.

The Offenbach case is especially interesting. The bank's costs for ATM operation are the same at lunchtime, but why do they charge more? Clearly a different type of customer, one with a higher willingness to pay, usually

*Submitted in partial fulfillment of the requirements for the degree of Bachelor of Science at the University of Cologne. Corrected and copyedited version, July 30, 2017.

[1]See Darmstädter Echo (2017).

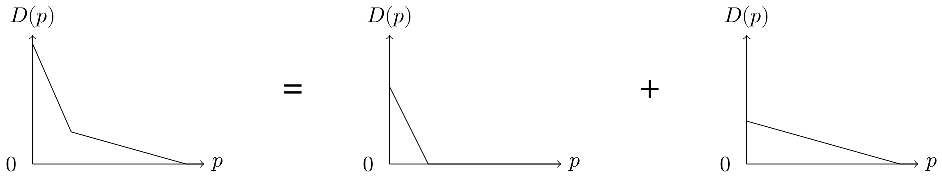


Figure 1: Consumers differ, leading to a subdivision of demand.

draws cash then—knowing this, the bank charges different prices to increase its profit. We are dealing with price discrimination.

This thesis is intended as a description and analysis of topics in price discrimination, including both classic and more recent approaches. This document is written in the spirit of being complete, yet not concluded: Complete in the sense that the topics are covered accurately and with meticulous scrutiny, where necessary; and not concluded in the sense that we plan to update it from time to time to reflect an ever growing number of new models and to perfect existing sections. However, the immediate goal of this thesis is to present five models of price discrimination, to give examples for their applicability, to analyze and discuss them both mathematically and informally, to compare them against each other and to reflect on the influence of technology on price discrimination.

This thesis consists of four main parts: Firstly, we give introductory remarks to price discrimination and we show when and how it is not only relevant in reality, but also how technology can aid the application of price discrimination. Secondly, we present the classic three degrees of price discrimination as per Pigou (1920), enriched with generalizations, simulations and graphics of our own. Thirdly, we discuss two more recent varieties of price discrimination, namely intertemporal price discrimination and price discrimination by purchase history. Finally, we conclude.

1.1 Fundamentals of price discrimination

As of June 2017, a Google Scholar search yields a ninety-eight page list of scientific papers on the topic. Despite of significant interest to economists, no general definition of price discrimination has been agreed upon. For the purpose of this thesis, we shall use the term "price discrimination" in a broad sense. *From the consumers' perspective, any price-setting behaviour that is not one price, forever, everywhere, under otherwise equal conditions is price discrimination. From the firm's perspective, price discrimination is the recognition of a subdivision of demand.* These two definitions will help

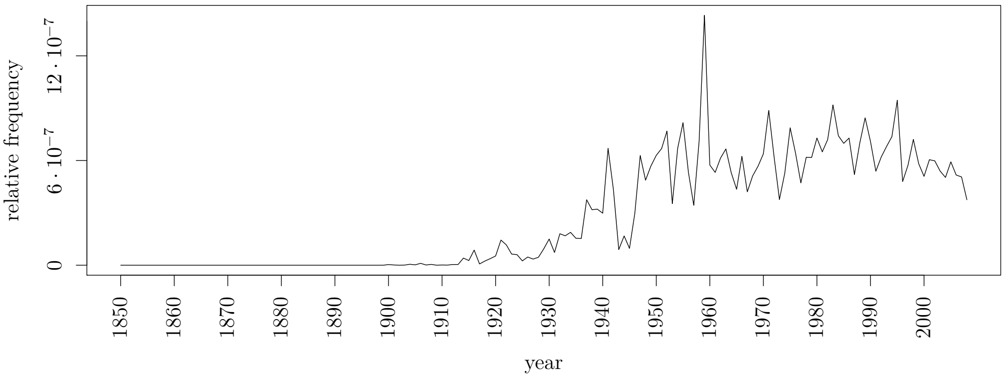


Figure 2: The popularity of the term "price discrimination" in books published from 1850 to 2008, according to Google Ngram Viewer (`https://books.google.com/ngrams`), accessed June 6, 2017.

us to encompass a larger spectrum of theories while remaining precise and vigilant.

The influence of price discrimination has been far-reaching and pervasive and since its inception, it has become a staple of literature on industrial organization. Figure 2 shows the relative frequency of the term 'price discrimination' in books published from 1850 to 2008.

For the firm, price discrimination consists of two parts. The first part is the *subdivision of demand*: The discriminating firm knows that consumers differ in some regard. For example, the firm knows that its product is sold on different continents, whose inhabitants possibly have a different willingness to pay. The firm might also know that some consumers differ in their (temporal) desire for the good. In other words, the firm is required to 'split up', or subdivide, the demand function that it faces. Figure 1 gives an illustration.

The second part is offering different prices to different groups. Once a subdivision of demand is ascertained, the firm has to calculate optimal prices. This, however, depends on the market and the type of subdivision: If we consider an electricity corporation that operates in a duopoly, it may be want to pay customers to switch. If the firm is a pharmaceutical monopolist, it can charge different prices in each country. If the firm sells cell phone contracts, it may wish to offer several tariffs and induce high-demand customers to choose a tariff with low per-unit costs, but a higher connection fee, while low-demand customers choose the tariff with higher per-unit costs, but a smaller connection fee. Theoretical and empirical models and technology aid in the determination of optimal prices, and technology aids in the process of putting up these prices.

Some of these models require that the subdivision of demand be linked to a distinguishing characteristic of consumers so that consumers can be easily recognized to be of one type or another; other models rely on self-selection. Models with self-selection are more robust as consumers cannot achieve lower prices by cheating the system.

It is generally agreed that price discrimination requires some form of market power, since a competitive firm can only charge a price that is equal to marginal cost. However, empirical studies have shown that even in markets more competitive than a monopoly, price discrimination can exist (cf. Borenstein and Rose (1994) and Escobari et al. (2016)). In this thesis, we will generally assume at least a duopolistic market.

1.2 Mechanisms and limitations

A firm that desires to implement price discrimination must find a real method to discern different types of customers (the subdivision of demand) and put up prices conditional on the type recognized.

Technological innovation eases both aspects of price discrimination: For example, when accessing a website, a web server receives certain information about the user from which inferences can be made. The information provided originates from multiple technological levels that are defined in ISO/IEC 7498-1 (E). This OSI model entails a modular architecture of networking; for example, a user application does not need to implement the bit-by-bit modulation/demodulation depending on whether datagrams are transmitted over Digital Subscriber Line (DSL) or over mobile internet. This is handled on another layer; the user application only determines what is sent and how received datagrams are processed, but not every detail of the physical transmission.

The Internet Protocol is the main standard to transmit data across networks. To identify and address any digital computer connected to the internet, any such computer receives an Internet Protocol address (IP address). As the Internet Protocol is implemented on the *Network layer* defined in the OSI model, it stems from the basic nature of the internet that if a computer receives a datagram from another computer, the originating IP address is also transmitted to enable a bilateral connection. Since IP addresses are sold in bulk to individual internet providers, it is often possible to reveal a precise region of access.[2] It is also generally possible to distinguish between mobile internet devices and DSL endpoints. Oftentimes, this information alone will

[2] Cf. Postel (1980) and for a discussion of IP geolocation, cf. Poese et al. (2011).

be enough: For example, videosharing websites determine the content that they are licensed to show based on IP geolocation.

However, on the *Application layer*, more information can be elicited. Consider HTTP, the protocol that is wrapped around websites: It is common practice that any request sent to a HTTP server includes `User-Agent`, a character string that reveals information about the device and browser of the user, cf. Berners–Lee et al. (1996), p. 46. Here are two real examples for `User-Agent` strings that are easily understood by specialized software and experts:

- `Mozilla/5.0 (Windows NT 6.1; WOW64; Trident/7.0; rv:11.0) like Gecko`
- `Mozilla/5.0 (iPhone; CPU iPhone OS 10_3_2 like Mac OS X) AppleWebKit/603.2.4 (KHTML, like Gecko) Version/10.0 Mobile/14F5089a Safari/602.1`

The HTTP protocol also allows *cookies*, small text files that are permanently stored on users' computers. These text files can be used to identify recurring users and can, in some cases, also be used to track users across websites.
Additionally, JavaScript, a scripting language that runs on the computer of the user, has classes like `navigator.plugins.*`, which allow the elicitation of active browser extensions, cf. Bewersdorff (2014). It is also possible to determine screen size and other parameters. Taken together, these technologies enable website providers to detect not only precise properties of the device being used, but also identify browsing habits. This information can be used to draw inferences about income, willingness to pay and other relevant variables that influence users' demand functions. Once a user has been identified to be of some type, putting up different prices based on information the user has—knowingly or unknowingly—provided, is trivial considering the dynamic nature of many websites.

The techniques used to draw inferences are not the main topic of this thesis. However, let it be said that *pricing experiments* can reveal the abovementioned subdivision of demand if one is willing to control for variables that were elicited as shown above. From such statistical analyses, cautious inferences about, for example, willingness to pay can be made. In essence, the information that was elicited by using technological means would be used while conducting separate pricing experiments: For example, users with a large screen size and users with smaller screen sizes could be presented with randomized prices and the firm would record sales along with the largeness of the screen size and the price that was shown to the user. Subsequently, demand functions can be estimated for each group, unveiling the subdivision of demand. For examples of pricing experiments, see Sexton et al. (1987), Lam and Small (2001), Gao et al. (2004) and many others; and for econometric methods of estimating demand functions, see Varian (1992), section 12.11.

Varian (1989) gives examples of legislation targeted against price discrimination (pp. 643 *et seq.*). Varian notes that early attempts such as the Clayton Antitrust Act of 1914 to outlaw 'predatory' price discrimination were not successful because their wording still allowed most forms of price discrimination, most importantly nonlinear pricing through quantity discounts. Varian (1989) goes on to show that subsequent efforts like the Robinson–Patman Act of 1936 were more radical and more successful. However, the federal government of the United States essentially halted the prosecution of cases of price discrimination in the 1960s, leaving private parties only with the possibility to seek damages. Blair and DePasquale (2014) call for the repeal of the Act, citing its *negative* effects on competition and arguing that recent court decisions have effectively gutted the prosecution of price discrimination under it. The negative implications of the Act are corroborated by O'Brien and Shaffer (1994), who show that not only may significant welfare losses occur under a prohibition of price discrimination, but every retailer may end up paying higher prices. Therefore, specific laws against price discrimination seem to be largely inoperative and legal action against price discrimination would only stem from other laws on competition policy (if the price discrimination undermines competition) or general statutes against discrimination based on privileged information (like race or gender).

Laws against (price) discrimination also apply to websites—it does not follow that price discrimination is legal where it is possibly easier. In fact, Article 5(3) of the ePrivacy Directive (2002) of the European Parliament specifically outlaws the use of information that was collected without the user's consent.

In the physical world, price discrimination seems to be harder to implement. Not only is it often hard to extract relevant information about customers in daily anonymous transactions, but the use of information that *is* available in such transactions is frequently outlawed. For example, gender and race are attributes that are in fact available. Laws limit the use of these attributes without regard to welfare effects.[3] Implementing price discrimination is therefore also limited. The fact that collecting information and

[3]Relevant statutes are the Civil Rights Act of 1964 in the United States and the Allgemeines Gleichbehandlungsgesetz in Germany.

changing (physical) price tags is costly amplifies this fact—but as we will see, the latter is only required by *some* forms of price discrimination. Forms that are not discriminatory on their face—or that discriminate not on individual characteristics—make legal challenges less problematic for a firm that is otherwise able to implement price discrimination. For example, in this thesis, we will present three types of price discrimination—second degree price discrimination, intertemporal price discrimination and price discrimination by purchase history—that rely on self-selection or voluntarily provided information and that are therefore not covered by such laws. Another factor may be much more relevant: As A. C. Pigou, the dean of price discrimination, once wrote: The discriminating monopolist must not "outrage the popular sense of justice" (Pigou (1920), p. 281) ...

2 The Classics

In this section, we discuss the three "classic" degrees of price discrimination. They were first introduced by Pigou (1920), pp. 240 *et seq.* We also give a brief introduction regarding the concept of quasilinear utility. The exclusion of income effects simplifies our analysis.

This presentation is largely based on Schmitz (2016), Bester (2012), Varian (1989), Varian (1992) and Varian (2011), but with a more general approach. Table (1) shows the different degrees that are discussed in this section with their respective payments for a customer of type i that buys the quantity x.

Degree	**Total payment for type i**
(None)	px
First	$p_i x + a_i$
Second	$px + a$
Third	$p_i x$

Table 1: The classic degrees of price discrimination, cf. Pigou (1920).

In this section, we assume perfect information and that consumers exhibit price-taking behaviour. Often, we will make use of the notation in Schmitz (2016).

2.1 Utility maximization

In this section, we will make use of a simplification regarding the utility function of the consumer. We will generally assume that each consumer i

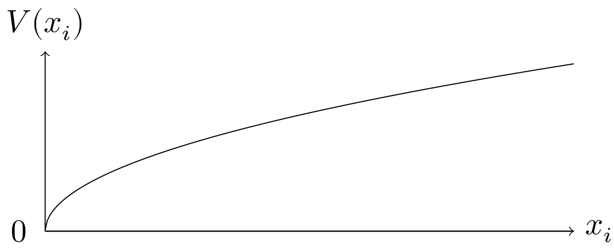


Figure 3: An exemplary utility function.

has the following utility function (see Schmitz (2016), p. 1 and cf. Varian (1992), p. 164):

$$U(x_i, z_i) = V(x_i) + z_i. \tag{1}$$

Here, x_i is the amount consumed of the good of interest and z_i is the numéraire, a fictional good with price 1 that is representative for all other consumed goods. First note that $V(x_i)$ is only dependent on x_i. We assume that $V(0) = 0$, $V(x_i) > 0$, $\frac{\partial V(x_i)}{\partial x_i} > 0$, $\frac{\partial V(x_i)}{\partial x_i}\Big|_{x_i=0} = \infty$ and $\frac{\partial^2 V(x_i)}{\partial x_i^2} < 0$.[4] The consumer maximizes $U(x_i, z_i)$ subject to him spending his income $b_i \geq 0$, leading to the following constrained optimization problem:

$$\begin{aligned} \max_{x_i, z_i} \quad & U(x_i, z_i) = V(x_i) + z_i \\ \text{subject to} \quad & px_i + z_i = b_i. \end{aligned} \tag{2}$$

z_i is immediately determined by the relation specified in the constraint if some x_i is chosen. Therefore $z_i \equiv b_i - px_i$, yielding the following unconstrained optimization problem:

$$\max_{x_i} \quad \mathfrak{U}(x_i) = V(x_i) + b_i - px_i. \tag{3}$$

$\mathfrak{U}(x_i)$ is the reduced form utility function. Importantly, $\mathfrak{U}(x_i)$ is also the net utility or 'surplus' of the consumer. It now becomes visible why this assumption is commonly referred to as "quasilinear utility": monetary units not spent on the relevant good flow back to the consumer *as utility*—and $V(x_i)$ depends only on x_i, not on other goods or, perhaps most importantly, the income b_i.

Differentiating (3) with respect to x_i, we reach the following first-order condition for interior maxima (cf. Schmitz (2016), p. 2):

[4] Note that $V(x_i)$ and $U(x_i, z_i)$ are both strictly concave functions. These assumptions are due to Schmitz (2016), p. 1.

$$\frac{\partial V(x_i)}{\partial x_i} = p. \tag{4}$$

We define $x_i^{(*)}(p)$ to be the solution to this condition. Since the utility function is strictly concave and has infinite slope at $x_i = 0$, the solution to equation (4) is always a local maximum.
These same results could also be obtained through a Lagrangian of the form $\mathcal{L} = V(x_i) + z_i + \lambda(b_i - px_i - z_i)$, which is optimized for x_i and z_i.

In the context of price discrimination, the firm will sometimes try to capture some of the consumer's utility through nonlinear pricing, thereby essentially reducing b_i through a fixed payment.[5] If such a fixed payment is demanded, an optimum need not occur in the interior: The consumer will only accept this extraction of utility if consumption is still beneficial, i. e. if $\mathfrak{U}\left(x_i^{(*)}(p)\right) \geq b_i$ for *any* reduced form quasilinear utility function $\mathfrak{U}$ (cf. Varian (1992), p. 165).[6] If the net utility achieved by consuming $x_i^{(*)}(p)$ and given some utility extraction scheme were lower than that hypothetical utility derived from purely consuming other goods via the numéraire, he consumes nothing of the good that we consider.[7] Therefore, let us define $x_i^*(p)$ to be the optimal level of consumption:

$$x_i^*(p) = \begin{cases} x_i^{(*)}(p) & \text{if } \mathfrak{U}\left(x_i^{(*)}(p)\right) \geq b_i, \\ 0 & \text{if not.} \end{cases} \tag{5}$$

It is often claimed that there are no income effects because $\partial x_i^*(p)/\partial b_i = 0$, being one of the most important implications of quasilinear utility. In the standard model, where the monopolist only sets prices, the consumer will always consume a positive amount and there are no income effects (since the utility function is strictly concave, has an infinite slope at $x_i = 0$ and income is nonnegative). However, if a fixed payment is charged, as in first and second degree price discrimination, the monopolist will have to take into account that consumers only consume if their surplus from consumption $\mathfrak{U}\left(x_i^{(*)}(p)\right)$ is indeed nonnegative.

[5] "Nonlinear pricing refers to any case in which the tariff is not strictly proportional to the quantity purchased" (cf. Wilson (1993), p. 4).

[6] The assumption that the consumer will consume if $\mathfrak{U}\left(x_i^{(*)}(p)\right) = b_i$ will ease our analysis later.

[7] Note that we implicitly assume that if the consumer chooses to consume nothing, the firm will not be able to extract utility.

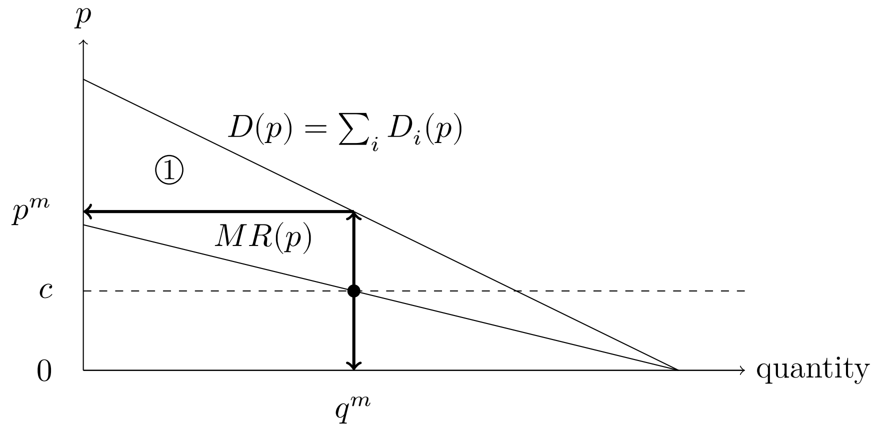


Figure 4: The monopolist's price and quantity choice, given no price discrimination. In the profit optimum, it must hold that marginal revenue $MR = \frac{\partial}{\partial p} pD(p)$ equals marginal cost. Area ① is equal to the consumer rent.

2.2 Perfect price discrimination

We now consider a market model with a discriminating monopolist with constant marginal production cost c and n types of consumers, within which the utility function is identical. All consumers have quasilinear utility. A type $i \in \{1, ..., n\}$ shall consist of m_i consumers. In this and the following sections, we will abstract from the consumers' incomes by setting $b_i = 0$ for all customers.[8]

In a market without price discrimination, the monopolist sets one price p, leaving all consumers with positive, non-zero utility (if utility functions are strictly concave). A consumer of type i with a quasi-linear utility function that buys x_i is left with the following surplus:

$$\mathfrak{U}_i(x_i) = V_i(x_i) - px_i. \tag{6}$$

However, if the monopolist is able to detect each consumer's type, there is the ability to charge different prices to different types. Additionally, the monopolist could charge a fee that is independent of the quantity consumed. This kind of price discrimination, where a consumer of type i pays $p_i x_i + a_i$ if $x_i > 0$,[9] is called first degree price discrimination (see Bester (2012), p. 65 *et seq.*) and leads, through its *de facto* reduction of $b_i = 0$, to the following surplus:

[8]This does not impede our further analysis due to quasilinear utility. In fact, any other nonnegative value of b_i could be assumed instead.

[9]If $x_i^* (p_i) = 0$, the consumer pays nothing.

$$\mathfrak{U}_i(x_i) = V_i(x_i) - p_i x_i - a_i \quad \forall i. \tag{7}$$

In section 1, we defined the consumer to (barely) buy if $\mathfrak{U}_i(x_i^{(*)}(p)) = 0$ and the discriminating monopolist will exploit this situation. The surplus will be zero if the firm uses the fixed payment a_i to extract all net utility as per equation (7), weakly inducing the consumer to buy. Therefore, let the optimal fixed payments a_i^* have the following value (cf. Bester (2012), p. 65):

$$a_i^* = V_i(x_i^*(p_i)) - p_i x_i^*(p_i) \quad \forall i. \tag{8}$$

Any further increase in a_i would lead to no consumption. Since we assumed quasilinear utility, and a_i effectively reduces b_i, the optimal level of consumption remains unaffected otherwise and $x_i^*(p_i) = x_i^{(*)}(p_i)$, where $x_i^{(*)}(p_i)$ is the solution to equation (4). Because the firm is able to discriminate perfectly and to seize all consumer surplus by offering a different tariff to each type, first degree price discrimination is also referred to as "perfect price discrimination".

This leaves the monopolist with setting the prices $p_i \quad \forall i$

$$\max_{p_1,\dots,p_n} \quad \pi = \sum_i m_i \left(p_i x_i^*(p_i) + a_i^*\right) - c\left(\sum_i \underbrace{m_i x_i^*(p_i)}_{D_i(p_i)}\right), \tag{9}$$

and therefore, after substitution and simplification:

$$\max_{p_1,\dots,p_n} \quad \pi = \sum_i m_i \left(V_i(x_i^*(p_i))\right) - c\left(\sum_i m_i x_i^*(p_i)\right).$$

The n first-order conditions simplify as follows:[10]

$$\begin{aligned} m_i \frac{\partial}{\partial x_i} V_i\left(x_i^*(p_i)\right) \frac{\partial}{\partial p_i} x_i^*(p_i) - c\, m_i \frac{\partial}{\partial p_i} x_i^*(p_i) &= 0, \\ m_i \frac{\partial}{\partial x_i} V_i\left(x_i^*(p_i)\right) \frac{\partial}{\partial p_i} x_i^*(p_i) &= c\, m_i \frac{\partial}{\partial p_i} x_i^*(p_i), \\ \frac{\partial}{\partial x_i} V_i\left(x_i^*(p_i)\right) \frac{\partial}{\partial p_i} x_i^*(p_i) &= c\, \frac{\partial}{\partial p_i} x_i^*(p_i), \\ \underbrace{\frac{\partial}{\partial x_i} V_i\left(x_i^*(p_i)\right)}_{\text{This expression must be equal to } p_i \text{, see eq. (4).}} &= c, \end{aligned} \tag{10}$$

[10]Note that these first-order conditions are for all i. Also, cf. Schmitz (2016), ch. 2, p. 38.

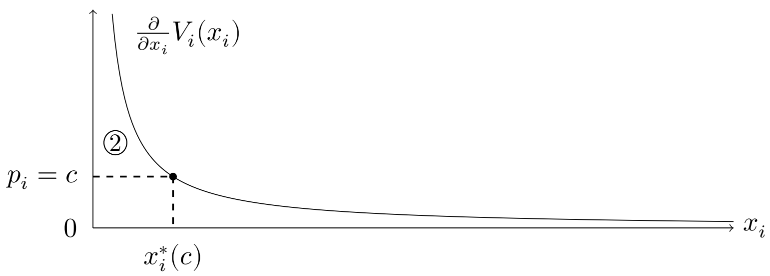


Figure 5: Consumer choice given a price of c. Each consumer's total surplus is extracted: ② $= \int_0^{x_i^*(c)} \frac{\partial}{\partial x_i} V_i(x_i)\, dx_i - cx_i^*(c) = V_i\left(x_i^*(c)\right) - cx_i^*(c) = a_i^*$.

and therefore we have

$$p_i^* = c \quad \forall i. \tag{11}$$

Summarizing, in first degree price discrimination, the monopolist sells each unit at marginal cost but extracts every consumer's total utility by charging an additional fixed payment (e. g. an entrance fee). Equation (11), which is undoubtedly quite stunning, implies that the monopolist makes no profit from the production and sale of the good alone—this stems from the assumption of constant marginal cost. Therefore, the firm profit is only due to the extraction of utitility, $\pi^* = \sum_i m_i\, a_i^* = \sum_i m_i\, (V_i(x_i^*\,(c)) - cx_i^*\,(c))$.

The sum of consumer surpluses (also known as 'consumer rent') is zero as all utility is captured through the fixed payment, see figure 5. Varian (1992), who considered a similar model of first degree price discrimination, noted that amazingly, because the monopolist is willing to sell to any consumer that is *per se* willing to buy at marginal cost, first degree price discrimination also represents a so-called welfare optimum, albeit a dismal one. All utility is redistributed to the monopolist and consumers barely want to consume at all: Were the fixed payment just infinitesimally higher, they would choose not to consume. However, note that the quantity sold is maximal.[11] Compare figures 4 and 6.

In reality, perfect price discrimination is hard to implement. Not only is the monopolist required to know each consumer's exact demand function, but he also needs to charge each type with a different fee—which is not only

[11] In this thesis, we generally disregard welfare economics. Here, welfare is defined to be the sum of consumer surpluses plus the sum of firm profits—and it could be deemed surprising that first degree price discrimination leads to the welfare optimum, which is otherwise only reached in competitive equilibria and in no other case of monopoly pricing. However, the monopolist absorbs all welfare and therefore consumers are still left worse off than in a non-monopolistic market *and* under a non-discriminating monopolist.

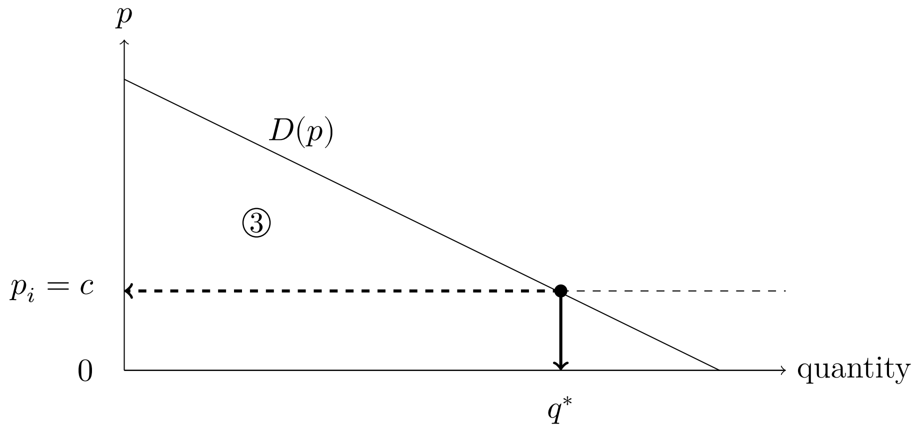


Figure 6: The monopolist's price and quantity choice, given first degree price discrimination. Area ③ is equal to the monopolist's profit as the consumer rent is completely extracted through the fixed payment a_i.

physically arduous, but requires total bargaining power. First degree price discrimination also implies that the monopolist can correctly identify each consumer's type. This assumption is not very plausible, since consumers could pretend to be of another type and achieve a lower total payment. Therefore, first degree price discrimination is usually irrelevant for reality—but it gives a hint at the maximum obtainable profit in a market.

2.3 Nonlinear pricing with constant tariffs

The setup of second degree price discrimination is identical to the setup of first degree price discrimination, with one key change: A consumer of type i does not pay $p_i x_i^* (p_i) + a_i$, but all consumers of all types pay the same price and the same fixed payment. Therefore, the payment for type i is only dependent on the level of consumption, but not on the type: $px_i^* (p) + a$ (cf. Bester (2012), p. 67). The main contrast to a non-discriminating monopolist is therefore the ability to charge a non-linear tariff.[12]

In the literature, it is often claimed that the monopolist is not informed about individual types (see Bester (2012), p. 67). However, this is misleading: The monopolist may still be able to *distinguish and recognize* the types, but he is not able to *discriminate*. The monopolist also still has full knowledge about the types' utility functions, he just cannot charge different prices or fixed payments. This inability is fundamental for second degree price dis-

[12]For alternative models on second degree price discrimination, see Varian (1992), pp. 244 *et seq.* and Oi (1971).

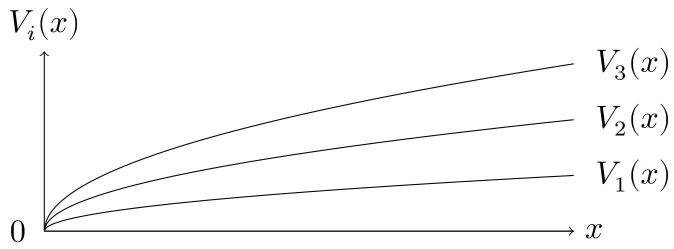


Figure 7: Three types of consumers.

crimination. The reasons may vary, but it is most likely that the recognition of group members is economically or legally prohibitive. But by setting the tariff, the monopolist still reacts to the subdivision of demand he ascertained.

Second degree price discrimination induces self-selection. In a way, the discrimination is 'outsourced' to consumers. In the variant discussed here, with only one tariff, consumers only 'self-select' into buying or not buying; but in variants with multiple tariffs, consumers must, as a rule, trade off lower per-unit prices against a higher fixed fee. Therefore, the discrimination lies in the fact that $\frac{px_j+a}{x_j} \neq \frac{px_k+a}{x_k}$ if only one tariff is considered. Essential to all variants of second degree price discrimination is that the average per unit payment is strictly decreasing in units consumed. Varian (1992) discusses the necessary self-selection constraints (pp. 245–246).

In the following, we assume that the existing types $1, \dots, n$ are ordered; that is, for any given price, type 1 has the lowest demand, type 2 has the second-lowest demand, etc. This requires the utility functions V_i to be non-intersecting, except at $x_i = 0$ (see Varian (1989), p. 612, or, more precisely, Varian (1992), p. 245). Figure 7 shows exemplary utility functions for $n = 3$ from which the ordering trivially follows, cf. equation (4). Stated mathematically:

$$x_1^*(p) < x_2^*(p) < \dots < x_n^*(p).$$

Through choosing only one tariff (p, a), the monopolist faces a critical choice, namely which customers to supply. A customer of type i will only buy if his surplus is nonnegative, or

$$\mathfrak{U}_i(x_i) = V_i(x_i) - px_i - a \geq 0. \tag{12}$$

(Cf. equation (7)). Since the monopolist chooses only one price and one fixed payment, he is able to "pick" the consumers he wants to serve. Choosing a high a will alienate some customers with low utility; they set $x_i^* = 0$. It is easily proven that the monopolist will supply all types starting from some $f \in \{1, \dots, n\}$ and extract f's full surplus by setting equation (12) equal to zero and solving for a.[13] The consumers with higher utility than this marginal type f will still consume, but those with a lower index will not. Therefore, we shall refer to f as the marginal type. The monopolist is subsequently left with setting prices.

Therefore, let the optimal fixed payment—depending on the marginal type f—have the following value:

$$a = V_f(x_f^*(p)) - px_f^*(p)\,. \tag{13}$$

The optimization problem is then stated as follows:

$$\max_f \quad \max_p \quad \pi = \sum_{i=f}^{n} m_i\left(px_i^*(p) + a\right) - c\left(\sum_{i=f}^{n} m_i x_i^*(p)\right), \tag{14}$$

(Please note that this is not an ordinary optimization problem. It is generally necessary to choose the optimal f by means of trial and error.)

For example, should the monopolist choose to serve all types ($f = 1$) the following optimization problem emerges:

$$\max_p \quad \pi = \sum_i m_i\left(px_i^*(p) + V_1(x_1^*(p)) - px_1^*(p)\right) - c\left(\sum_i m_i x_i^*(p)\right). \tag{15}$$

We omit the first-order condition for reasons of legibility. However, let it be said that $p \geq c$ (with equality if $f = n$) and therefore, a welfare optimum is generally not reached (cf. Bester (2012), pp. 68 *et seq.*).

We used R (see R Core Team (2017)) to devise a simulation with $n = 100$ types to examine how the choice of f, i. e. the first type to serve, influences profits. In our simulation, profits are highest if the 46 types with the lowest utility are excluded; but excluding types beyond the optimum leads to a sharp reduction in profits. This insight is very important: It is *possible* to serve all types (by setting $f = 1$), but it is not necessarily *optimal.* In the literature, this is sometimes boldly disregarded (cf. Bester (2012), p. 68, footnote 34). Figure 8 shows optimal profits as a function of only f.[14]

In figure 9, we show optimal prices as a function of f. It can be seen that prices are highest at $f = 1$, a compensation for the initially low fixed payments a. As soon as f (and therefore a) rise, prices fall, with their lowest value at $f = 100$, where $p = c = 1$. This result, too, is intuitive: If only the

[13] See section A.1 for details; an alternative proof is given in Varian (1992), p. 246.

[14] p and a are optimized for any given choice of f, see equation (14).

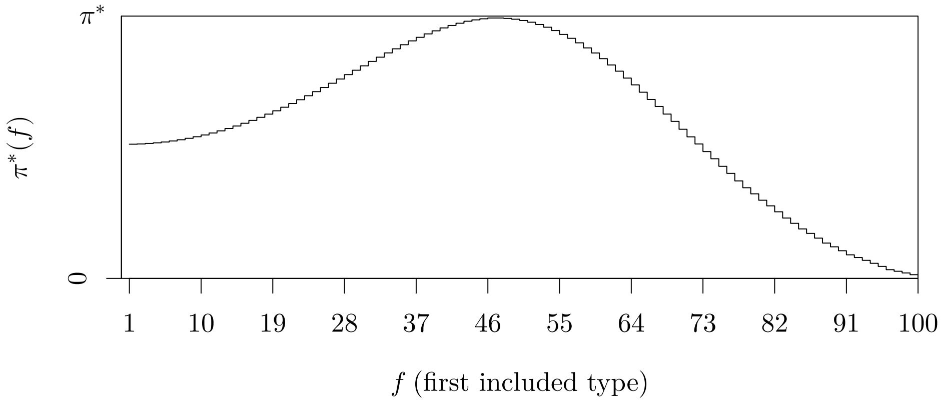


Figure 8: Optimal profit as a function of the marginal type. For any point on the x-axis, it holds that all types left from it are excluded (their demand is zero).

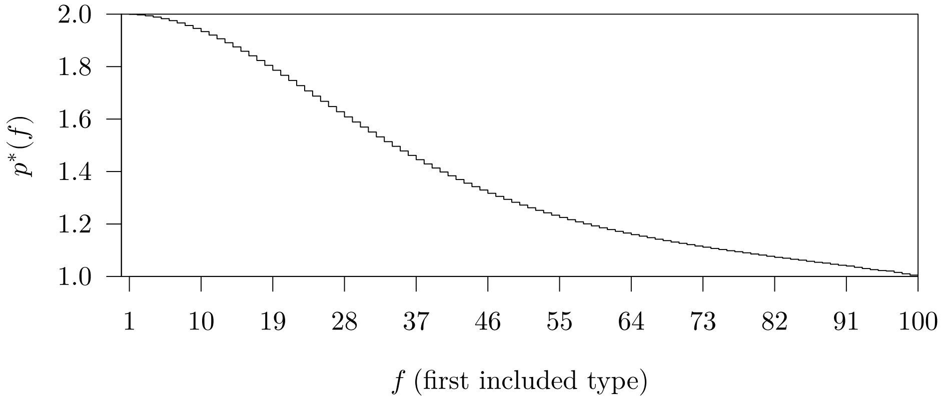


Figure 9: Optimal price as a function of the marginal type.

type with the highest utility, and therefore the highest willingness to pay is supplied, a will be set as in first degree price discrimination (to extract all remaining utility) and $p = c$, cf. equations (10).

In reality, payments of the form $px + a$ are often observed, especially for subscription-type purchases: Providers of gas and electricity charge a monthly connection fee and so do cell phone contracts and DSL providers. Not all of these examples are cases of pure second degree price discrimination, but some degree of it may be involved. Cases where multiple such tariffs are offered are especially interesting as they induce consumers to choose the preferred tariff from a menu of tariffs. Consider a cell phone contract: It follows logically that if two tariffs $(p_1, a_1), (p_2, a_2)$ are offered, and $a_2 > a_1$, then $p_2 < p_1$.[15] Ultimately, large-scale consumers will choose the second tariff, and lower-demand consumers will choose the first tariff. Self-selection is therefore a defining feature of second degree price discrimination.

If consumers can trade with each other, bulk consumers can profitably sell to lower-demand consumers. However, due to the fact that not all consumer's surplus is seized—in contrast with first degree price discrimination—the incentive to resell is comparably lower. Also note that the subscription-type purchases mentioned above often have physical limitations regarding their ability to be resold.

Without reselling, as second degree price discrimination induces self-selection, consumers are best-off when choosing the tariff for which their surplus is highest.[16] Therefore, second degree price discrimination is an *incentive compatible* and *strategyproof* mechanism that reveals the true preferences of the consumer.

2.4 Third degree price discrimination

In the third and final degree of 'classic' price discrimination, the monopolist is no longer able to charge a fixed payment, but he can still set different prices for different types of consumers. Therefore, the payment demanded by the monopolist for x units of the good is $p_i x$ (see Varian (1992), p. 248). Consumer surplus is equivalent to equation (6), but with p_i instead of p.

The profit maximization problem is as follows:

$$\max_{p_1,\ldots,p_n} \quad \pi = \sum_i m_i p_i x_i^* (p_i) - c \left(\sum_i m_i x_i^* (p_i) \right). \tag{16}$$

[15] Otherwise, no consumer would choose the second tariff.

[16] However, in our case, there is only one tariff and consumers can only self-select into buying or not buying.

This leads to the following first-order conditions, which are subsequently simplified:[17]

$$m_i\left(p_i\frac{\partial}{\partial p_i}x_i^*\left(p_i\right)+x_i^*\left(p_i\right)\right)-c\,m_i\frac{\partial}{\partial p_i}x_i^*\left(p_i\right)=0, \tag{17}$$

$$m_i\left(p_i\frac{\partial}{\partial p_i}x_i^*\left(p_i\right)+x_i^*\left(p_i\right)\right)=c\,m_i\frac{\partial}{\partial p_i}x_i^*\left(p_i\right),$$

$$p_i\frac{\partial}{\partial p_i}x_i^*\left(p_i\right)+x_i^*\left(p_i\right)=c\,\frac{\partial}{\partial p_i}x_i^*\left(p_i\right),$$

and therefore it holds that

$$p_i=c-\frac{x_i^*\left(p_i\right)}{\frac{\partial}{\partial p_i}x_i^*\left(p_i\right)}\qquad\forall i. \tag{18}$$

Note that $p_i > c \quad \forall i$, because $\frac{\partial}{\partial p_i}x_i^*\left(p_i\right) < 0$—we are dealing with a non-Giffen good, but this also follows from the strict concavity of the utility function—and $x_i^*\left(p_i\right) \overset{\text{here}}{\neq} 0$.[18]

We found that affine transformations of some classes of utility functions are insufficient to induce third degree price discrimination. See section A.2 for an illustration.

In this section, profit was—despite many variations—always generally written as a function of price. Writing profit as a function of quantity, we reach the conclusion that in the profit optimum of third degree price discrimination, the marginal revenue of all consumer types are equal, and equal to marginal cost:

The profit maximization problem in third degree price discrimination follows from equation (16). Now, let us use the demand function from equation (9), $D_i\left(p_i\right)=m_i x_i^*\left(p_i\right)$. The maximization problem is then

$$\max_{p_1,\ldots,p_n}\quad \pi=\sum_i p_i D_i\left(p_i\right)-c\sum_i D_i\left(p_i\right). \tag{19}$$

However, it is possible to formulate profit as a function of quantity per type. Let $p_i\left(q_i\right)$ be the inverse demand function, i. e. $p_i\left(q_i\right)=D_i^{-1}\left(p_i\right)$ and let us define revenue per type to be $R_i\left(q_i\right)=q_i p_i\left(q_i\right)$. We reach the following new maximization problem:

[17] Note again that there are n first-order conditions and they all simplify like that.

[18] No fixed payment is charged; all consumers will want to consume a positive amount since $\left.\frac{\partial V(x_i)}{\partial x_i}\right|_{x_i=0}=\infty$.

$$\max_{q_1,\dots,q_n} \quad \pi = \sum_i R_i(q_i) - c \sum_i q_i. \tag{20}$$

Note that the cost side of above equation simplifies greatly. We derive the first-order conditions by taking the derivative with respect to $q_i \quad \forall i$. In the profit optimum it holds that $\frac{\partial R_i(q_i)}{\partial q_i} = c \quad \forall i$: In the profit optimum, marginal revenues are all equal to marginal cost.

There is an immediate intuition for this result: Were marginal revenue of one type higher than marginal revenue of another type, it would strictly increase the profit to sell more to the type with higher marginal revenue and less to the type with lower marginal revenue. This hypothetical 'swapping' of units takes place until all marginal revenues are equal to marginal cost: Profit maximization in third degree price discrimination is therefore identical to an ordinary monopolist that sells the same good on n mutually independent markets.

Third degree price discrimination seems to be the most popular variant of price discrimination since no fixed payment is demanded. In reality, third degree price discrimination is often observed, even in more competitive markets. For example, the Cologne Zoological Garden has different prices for students, the elderly and the disabled.[19] Such examples are ubiquitous. It is truly no wonder that Pigou once called it "of chief practical importance" (Pigou (1920), p. 246).

Varian (1992) discusses what occurs if the price charged on one 'market' influences another, i. e. if the discrimination is imperfect (pp. 249–250). Arbitrage concerns are sometimes mitigated by law. For example, local public transport corporations sell one day travel passes that are cheaper than the individual tickets. Thereby, the corporations are involved in third degree price discrimination. However, these day passes are often valid even after the travel is concluded. It would therefore seem sensible to resell them to fellow travellers. However, as per the terms and conditions of public transport corporations, such resold tickets are invalid and price discrimination is at least somewhat upheld.

Pharmaceuticals are also exemplary. Emtricitabine/tenofovir—a potent HIV medication—costs about \$ 6.08 per month in the developing world while it is \$ 1,644.59 per month in the United States.[20] It is illegal to import or export drugs without the approval of the United States Food and Drug Administration, and therefore, due to a differing willingness to pay, third degree

[19] Cf. `http://www.koelnerzoo.de/index.php/besuch#preise`, accessed June 6, 2017.

[20] Cf. `http://apps.who.int/medicinedocs/documents/s21982en/s21982en.pdf`, p. 222 and `https://www.drugs.com/price-guide/truvada`, both accessed June 6, 2017.

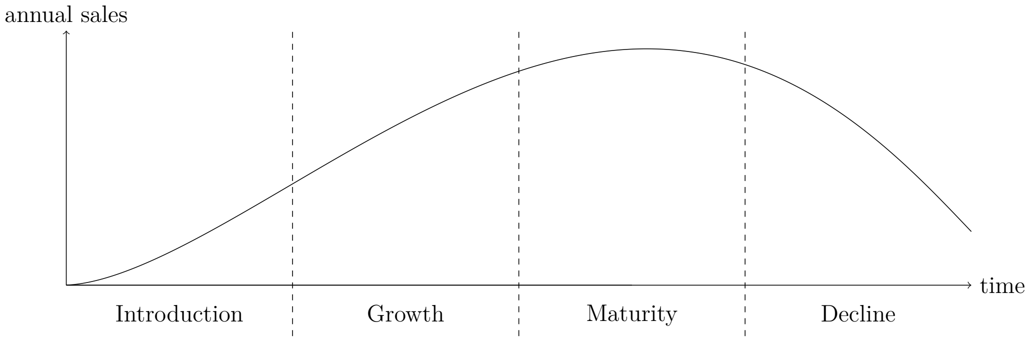


Figure 10: The product life-cycle, adapted from Malakooti (2013), p. 26.

price discrimination can persist. In a hypothetical market without such governmental restrictions, emtricitabine/tenofovir could simply be transshipped to the United States from developing countries and prices would lower.

3 New Concepts of Price Discrimination

3.1 Intertemporal price discrimination

New products are often sold at a high price that eventually lowers. One possible explanation for this change of price could be price discrimination: Early adopters often have a high time preference and income (cf. Pew Research Center (2016)), two factors that could be exploited by the firm. However, making the product available to a bigger group by lowering prices may be profitable if growth is desired. The famous product life-cycle curve, depicted in figure 10, shows four different stages of product existence.

It seems like price discrimination by product stage, or "intertemporal price discrimination", could complement the natural evolution of product sales. If intertemporal price discrimination is implemented, the firm can extract the early adopters' utility and only subsequently make the product available to consumers whose demand for the product is less immediate. Naturally, any such discrimination is only conceivable if consumers differ in that regard. In such a model, self-selection would be important since consumers have to decide at which point to buy. The discussion in this section is largely based on Stokey (1979), who considers a model of a price-setting monopolist. As there are no fixed payments, intertemporal price discrimination is clearly a variant of third degree price discrimination.

In this section, we will discuss the model by Stokey (1979) and its implications, both formally and informally. In the mathematical section, we retrace many of Stokey's analytical steps. We will also reflect on some assumptions of Stokey's model. In the formal and informal analysis, we will slightly modify the notation of Stokey (1979) to prevent confusion. The present refers to t_0—the point in time in which the monopolist sets his pricing schedule.

3.1.1 Stokey's model setup

The model of Stokey (1979) is formulated in continuous time. Some new product can be sold from t_0 until t_1 to a fixed continuum of consumers with perfect foresight to which these conditions are common knowledge. Sales can occur at any point between t_0 and t_1; these dates are exogenously given. The monopolist has to select a pricing strategy at t_0 to maximize the discounted value of profits.

Importantly, the product is durable, i. e. during the time period that we consider, each consumer will consume at most one unit of the good. For the moment, the good can be produced costlessly. Both consumers and firm discount profits and utility with a constant interest rate r.[21]

Each consumer is described by $v \in [0, V]$, the willingness to pay (or valuation) for the good at t_0 should it be received immediately. Each consumer's exact v is private information—the monopolist only knows the distribution of v in the population and can therefore set only one price per period (making Stokey's model a version of third degree price discrimination). Let $f(v)$ be the probability density function and $\Finv(v)$ the right-tailed cumulative distribution function; that is, $\Finv(v) = \int_v^V f(v)\, dv = 1 - F(v)$.[22]

v describes the utility at t_0; the *present* (or discounted) valuation for the good should it be received at time t is characterized by an utility function $U(t, v)$.[23] From the above line of reasoning, it follows that $U(0, v) = v$. For any fixed point of delivery t, an increase in v leads to an increase in U; and for any fixed valuation, a delay in delivery time leads to a decrease in U. Stokey proposes that those with a higher valuation suffer more from the delay. Now, therefore, we have $\frac{\partial}{\partial t}U(t, v) < 0$, $\frac{\partial}{\partial v}U(t, v) > 0$ and $\frac{\partial^2}{\partial t \partial v}U(t, v) < 0$.

Consumers do not make 'preorders': They pay as soon as the good is received, but not earlier. Therefore, a customer v derives the following *present* surplus $\mathfrak{U}$ (cf. the definitions of surplus in section 2) from buying the product at t:

[21] Discounting is continuous; i. e. the discount factor takes the value e^{-rT}.

[22] We introduce the notation for the right-tailed cumulative distribution function so that it is not confused with the (ordinarily used) left-tailed cumulative distribution function $F(v) = \int_0^v f(v)\, dv$. The use of a right-tailed c. d. f. is due to Stokey (1979).

[23] $U(t, v)$ is the valuation at t_0: The consumer would be willing to pay $U(t, v)$ at t_0 if promised delivery at t.

$$\mathfrak{U}(t,v) = \underbrace{U(t,v)}_{\text{present valuation for delivery at } t} - \underbrace{e^{-rt} \overbrace{p(t)}^{\text{price at } t}}_{\text{discounted payment}}. \tag{21}$$

$p(t)$ is the common knowledge price schedule of the firm; it relates the point in time to the price charged by the monopolist at that point in time. In the spirit of Stokey (1979), we assume $p(t)$ to be a continuous function and twice differentiable. Now consider the following: Each v has a point in time at which consumption is optimal (if that v consumes at all). Consumers maximize equation (21), and we shall refer to the optimal point of consumption for some consumer v as $\tau(v)$. For this $\tau(v)$, equation (21) is maximal with regard to t. For interior maxima, $\tau(v)$ is implicitly defined by the following first-order condition:

$$\frac{\partial}{\partial t}\mathfrak{U}(t,v)\Big|_{t=\tau(v)} = re^{-r\tau(v)}p\left(\tau\left(v\right)\right) - e^{-r\tau(v)}\left(\frac{\partial}{\partial t}p(t)\Big|_{t=\tau(v)}\right) + \frac{\partial}{\partial t}U(t,v)\Big|_{t=\tau(v)} = 0. \tag{22}$$

The (omitted) second-order condition also has to hold. However, an optimum need not occur in the interior of equation (21). It could very well be that a corner solution is optimal, i. e. at $t = t_0$ or $t = t_1$. Stokey (1979) considers three marginal consumers that aid this and further analysis:

1. There is a consumer v_0 that is indifferent between consuming at t_0 and $t_0 + \varepsilon$, with ε being a sufficiently small positive real number. Consumers with a higher valuation also consume at t_0, but any consumer with a marginally lower valuation consumes later. It follows that $\frac{\partial}{\partial t}\mathfrak{U}(t,v)\big|_{t=t_0,v=v_0} = 0$.

2. Similarly, there is a consumer v_1 that is indifferent between consuming at t_1 and $t_1 - \varepsilon$. For this consumer, it follows that $\frac{\partial}{\partial t}\mathfrak{U}(t,v)\big|_{t=t_1,v=v_1} = 0$.

3. Finally, there is a consumer v_2 that is indifferent between consuming at t_1 and not consuming at all. It follows that $\mathfrak{U}(t_1, v_2) = 0$.

Note that it follows trivially from the above line of reasoning that only those consume whose $v \geq v_2$. Additionally, it can be shown using implicit differentiation that $\frac{\partial}{\partial v}\tau(v) < 0$, implying that those with higher valuations buy earlier than those with lower valuations (cf. Stokey (1979), p. 359): Sorting ascendingly by valuations is congruent with sorting descendingly by optimal purchase time.

By rearranging equation (22), Stokey reaches the following intermediate result:

$$\underbrace{\frac{\partial}{\partial t}U(t,v)}_{<0} = \underbrace{e^{-r\tau(v)}}_{>0} \underbrace{\left(\frac{\partial}{\partial t}p(t) - \overbrace{rp\left(\tau(v)\right)}^{>0}\right)}_{\text{Ⓐ}}.$$

What can be said about Ⓐ? First of all, note that $\frac{\partial}{\partial t}U(t,v) < 0$ is by definition, cf. p. 21. For this relation to hold, it is necessary that $\frac{\partial}{\partial t}p(t) < rp\left(\tau(v)\right)$, so that Ⓐ becomes negative. The implication is that if sales are to occur at all times, the rate of change of price must not exceed $rp\left(\tau(v)\right)$. Since r is small, it is an acceptable simplification to say that only continuously falling prices $\left(\frac{\partial}{\partial t}p(t) < 0\right)$ induce continuous sales with certainty.

Let us now consider the firm's maximization problem at t_0. We already ascertained that the firm will want to choose a price strategy that maximizes the discounted profit. First note that for any $v \geq v_2$, there is an optimal purchase time $\tau(v)$. Also note that $f(v)$ is the density of v. For any pricing strategy that the firm considers, the firm can anticipate the behaviour of consumers—the price that is extracted at v's optimal purchase time is $p(\tau(v))$. The exact functional form of $p(t)$ remains to be chosen *given the behaviour of consumers* $\tau(v)$. Now, therefore, we can write profit as a function of $p(t)$:

$$\max_{p(t)} \quad \pi = \int_{v_2}^{V} \underbrace{e^{-r\tau(v)}}_{\text{discount factor}} \quad p(\tau(v)) \quad \underbrace{f(v)}_{\text{density/`quantity' of } v} \quad dv. \tag{23}$$

The firm aims to choose a $p(t)$ that maximizes its discounted profit: In equation (23), we do not optimize for a variable, but we search for a function $p(t)$ that maximizes the expression on the right-hand side. Such an optimization problem is by no means trivial, and it requires dynamic optimization to find a suitable $p(t)$. Stokey (1979) proceeds to solve this optimization problem using the *calculus of variations*. Unfortunately, the methods used are too advanced to be discussed in this thesis. The interested reader is referred to Chiang (2000). We will proceed by discussing the results of Stokey (1979) less formally.

Stokey defines the following function that gives the valuation for any purchase time:

$$v(t) = \begin{cases} v_0 & \text{if } t = t_0, \\ \tau^{-1}(t) & \text{if } t_1 > t > t_0, \\ v_1 & \text{if } t = t_1. \end{cases} \tag{24}$$

In essence, $v(t)$ is the inverse of $\tau(v)$ (including corner solutions). It follows that equation (23) can also be written in terms of $v(t)$, cf. Stokey (1979), pp. 360 *et seq.* Using the calculus of variations, Stokey (1979) finds the two conditions that need to hold in a profit optimum (p. 361, (12) and (13)):

$$\Finv(v_2)\left(\frac{\partial}{\partial v}U(t,v)\bigg|_{t=t_1,v=v_2}\right) - f(v_2)U(t_1,v_2) = 0, \tag{25}$$

$$f(x(t))\left(\frac{\partial}{\partial t}U(t,v)\bigg|_{v=v(t)}\right) - \Finv(v(t))\left(\frac{\partial^2}{\partial t \partial x}U(t,v)\bigg|_{v=v(t)}\right) = 0. \tag{26}$$

Now consider the following class of utility functions: $U(t,v) = g(t) \cdot v$, with $g(t) > 0$ and the other assumptions about the utility function fulfilled. This class of utility functions is interesting because v is independent of the time of delivery t, and vice versa. The 't effect' and the 'v effect' are perfectly separable. Stokey goes on to show that for this class of utility functions, *price discrimination will not be implemented*. By setting $U(t,v) = g(t) \cdot v$ and applying (25) and (26), we reach the following:

$$\Finv(v_2)g(t_1) - f(v_2)g(t_1)v_2 = 0,$$
$$f(v(t))v(t)\frac{\partial}{\partial t}g(t) - \Finv(v(t))\frac{\partial}{\partial t}g(t) = 0.$$

We simplify, multiply the first equation by (-1) and rearrange (cf. Stokey (1979), p. 363, (16) and (17)):

$$f(v_2)v_2 - \Finv(v_2) = 0, \tag{27}$$
$$f(v(t))v(t) - \Finv(v(t)) = 0. \tag{28}$$

From the above system of equations it follows that $v(t) \equiv v_2$, and therefore, due to the definition of $v(t)$, all sales occur at $t_0 = t_1 = t_2$, and none thereafter. This is the main result of Stokey's basic model: That given a large, common class of utility functions, no price discrimination will be implemented as no sales occur except at exactly one point in time.

3.1.2 Further results

The monopolist in the previous section produced costlessly. Stokey shows that if marginal costs are constant over time, the results of above hold true.

If we consider marginal costs as a function of time, price discrimination may be profitable. The assumption of time-dependent marginal costs seems odd at first glance, since one usually assumes marginal costs to vary with output (for example, through learning effects). Nevertheless, the assumption of time-dependent marginal costs is an useful one since it allows the study of implications without information about the production function. It is therefore a mere ‘phenotypical’ simplification of the production process. If we accept this assumption, we can easily modify our optimization problem (23) by introducing $c(t)$:

$$\max_{p(t)} \quad \pi = \int_{v_2}^{V} \underbrace{e^{-r\tau(v)}}_{\text{discount factor}} \left(p(\tau(v)) - c(\tau(v))\right) \underbrace{f(v)}_{\text{density/‘quantity’ of } v} dv. \tag{29}$$

From the solutions in the original paper, it follows that price discrimination is easily profitable. If we turn once again to utility function of the form $U(t, v) = g(t) \cdot v$, we find that price discrimination will be induced if costs fall ‘fast enough’; that is, if “discounted unit costs fall at a faster proportionate rate [...] than consumers’ discounted reservation prices” (Stokey (1979), p. 366).

It is also noteworthy that if $U(0, v) = v$ is positively correlated with high time preference (i. e. $-\frac{\partial}{\partial t}U(t, v)$—how much present surplus would the consumer lose if consumption were to take place infinitesimally later?), price discrimination can be induced—but not if all consumers have the same time preference. It follows that the firm may want to exploit a consumer that has a high valuation for the good at t_0, but who also loses much of his valuation if he purchases at a later date. Stokey also notes that if there is no correlation between $U(0, v)$ and time preference, price discrimination is only sometimes profitable, even if time preference rates vary.

Furthermore, if capacity constraints or strictly convex marginal costs exist, there may be a natural incentive for the firm to spread out production over a period of time. In this instance, prices may vary, but the variation does not stem from discrimination, but from realities of production.

3.1.3 A simpler model

Using R, we commissioned a simulation to simulate consumer behaviour for different given prices. For this simulation, we simplified the model of Stokey

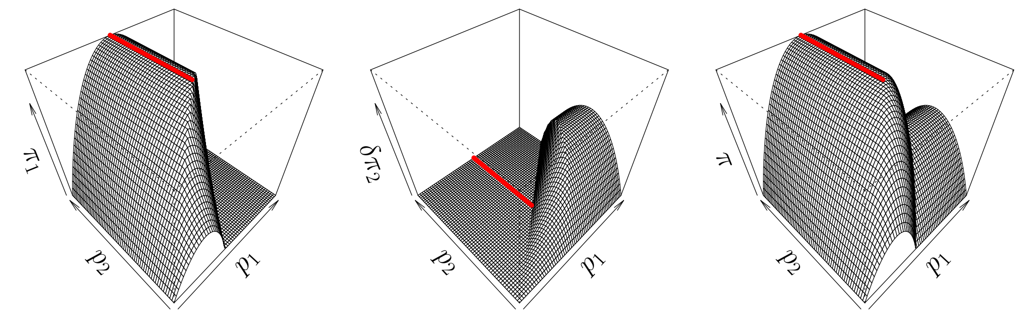


Figure 11: Given equal discount factors, no production costs and trivial valuations, no price discrimination will be induced. Note that all axes are scaled identically in all graphs so that direct comparisons are possible.

(1979) into a two-period model. A related two-period model was considered by Varian (1989), but with more restrictions and only arithmetically. We assume 'many' consumers that have valuations $v(t)$ for our good, depending on the period t in which the good is received. The firm has to decide which prices to charge in the two periods; anticipating that consumers have perfect foresight, the firm maximizes profits. Both the firm and the consumers discount profits, utility and costs of the second period with some discount factor δ. Once again, consumers maximize surplus: Presented with a price schedule (p_1, p_2), a consumer will decide to buy in the first period if, for his v,

$$v(1) - p_1 \geq \delta(v(2) - p_2) \text{ and } v(1) - p_1 \geq 0. \tag{30}$$

If $\delta(v(2) - p_2) > v(1) - p_1$ and $\delta(v(2) - p_2) \geq 0$, the consumer will decide to buy in the second period. If both $v(1) - p_1$ and $\delta(v(2) - p_2)$ are lower than 0, the consumer will buy in neither period. Note that we again assume perfect information and foresight, as does Stokey (1979).
The firm's technology is characterized by marginal costs c_t that are assumed to be constant inside each period. Total firm profits are defined as $\pi = \pi_1 + \delta\pi_2$, where π_t is the profit anticipated for period t. If profits would be negative for any price, the firm decides not to produce in this period.

We executed a simulation with 500,000 consumers and we considered the following settings that are akin to some situations contained in Stokey (1979). For each consumer, consumption choice is simulated for all prices that are considered and profits are calculated subsequently. Our use of simulation enables us to study a wide range of alternative assumptions regarding cost structures and consumer preferences.

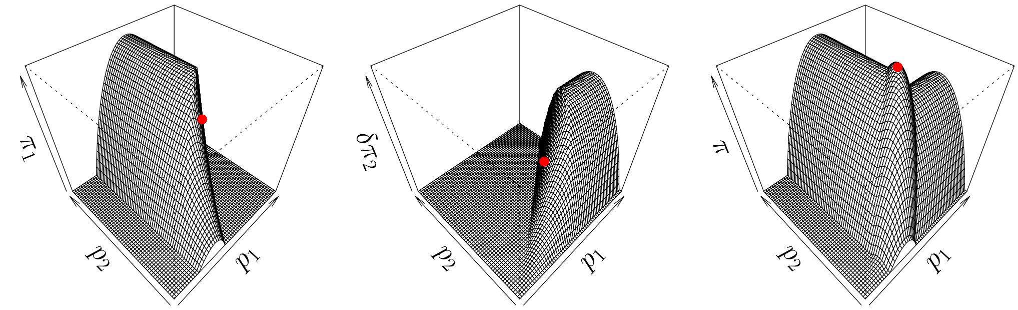


Figure 12: If we introduce falling production costs, price variation is observed.

Firstly, we considered a situation similar to Stokey's standard model with no production cost. We assumed that both consumers and the firm have a common discount factor $\delta = 3/4$. Stokey demonstrated that if time and valuation are trivially linked, no price discrimination will be induced. We assumed that period 1 valuations are $v(1) \in [0, 10]$ (uniformly distributed) and period 2 valuations are $v(2) = v(1)/1.5$.[24] Profits for each period and total anticipated profits are depicted in figure 11. In the rightmost graph, we highlighted (in red) optimal total profits; the contribution of each period to these optimal profits are highlighted in the other graphs. Note that in general, total profit is only rarely composed of both periods. It is largely only one period that is responsible for the total profit. If for any combination of (p_1, p_2), profits are not generated in some period, it follows that no sales occured in this period, confirming the result of Stokey (1979) that no price discrimination will be profitable for these prices. We found that for some values of p_1 and p_2, sales *do occur* in both periods (near $p_2 = 0$); but these instances are small and have no bearing at all on optimal profits. The multiple optima emerge from sales made in the first period and beyond some threshold, p_2 is irrelevant—at the highlighted price combinations for which total profits are optimal, simply no sales occur in the second period.

Secondly, we incorporated Stokey's implication that falling costs may induce price discrimination into our model. We modified our simulation so that $c_1 = 2$, $c_2 = 0.2$. Profits are shown in figure 12. In this case, there is only one optimum, one to which both periods contribute. It follows that price discrimination is induced. However, we confirmed the intuition given

[24] All numeric values are chosen out of convenience; other values could be assumed instead. We verified the results of our simulations with other values, and they continue to hold. Due to the boundaries implied, we only simulate prices between 0 and 10.

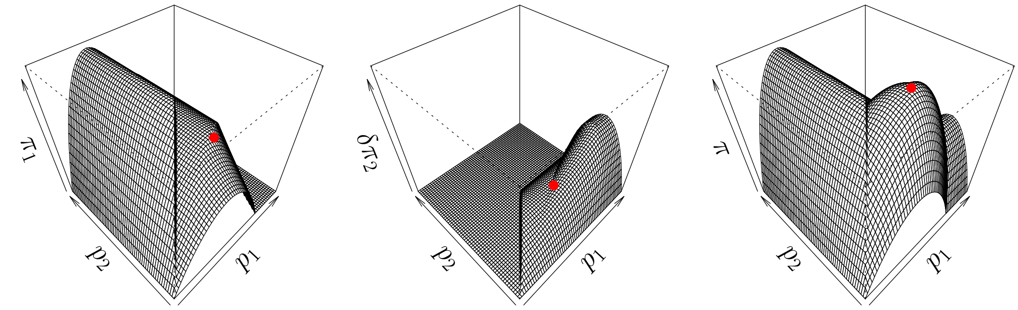


Figure 13: If we correlate period 1 valuations with consumer discount factors (that are themselves varying), we find that price discrimination is induced.

by Stokey (1979), i. e. that costs need to fall rapidly. Not all combinations of c_1 and c_2 induce price discrimination; only those for which it holds that the reduction in marginal costs is high enough induce said variation in prices (and hence sales in both periods).

Finally, figure 13 shows profits given correlated valuations. In this setting, there are again no production costs and we start out with the standard model presented in the first simulation. The analysis in Stokey (1979) yields the result that for price discrimination to be profitable, period 1 valuations need to be positively correlated with time preference. We considered two equally large groups of customers, one for whom we set $\delta = 1/2$ and one for whom we set $\delta = 1$. The firm δ remains set at 3/4. It follows that the latter group of consumers (that is infinitely patient and has a low time preference) should, in the spirit of Stokey (1979), also have a comparably lower period 1 valuation: Those who have $\delta = 1/2$ have $v(1) \in [0, 10]$, again uniformly distributed, and $v(2) = v(1)/2$; those who have $\delta = 1$ have $v(1) \in [0, 5]$ and $v(2) = v(1)$. Once again, price discrimination was induced. This case is considerably different from the previous one: Here, the variation in prices is purely because of discrimination, as the firm exploits high-time preference consumers. In the previous case, falling costs are explanatory as well. (If valuations are uncorrelated with (differing) discount factors, price discrimination may or not be profitable, depending on the precise values chosen. This, too, is corroborating a result of Stokey (1979). We did not include a figure on that simulation.)

The results of our simulation were found to be stable. Varian (1989), in his somewhat related model, found that price discrimination can be profitable even in his standard model if the firm is more patient than consumers. We found that for some combinations of firm and consumer discount factors, this result holds true in our simulation. As this setting is not considered by Stokey (1979), we did not include a specific report on it; however, readers are encouraged to tweak our simulation to investigate this and similar scenarios. The original code is available on request.

3.1.4 Summary

Stokey (1979) is the seminal paper on intertemporal price discrimination. Her main contribution lies in the fact that her arithmetically complex model allows the consideration of many cases that go beyond the standard model.

As we have seen, given a commonly used class of utility functions, intertemporal price discrimination is not profitable, resulting in a monopolist that chooses to sell at only one point in time. However, given falling prices or varying degrees of time preference, the monopolist may conclude to implement intertemporal price discrimination.

We verified the implications of the original model using a simpler model with two periods. In our simulation, we found that not only can all of Stokey's results be verified, but the use of a simulation adds flexibility and enables us to consider more cases. As consumers have perfect foresight, they anticipate the price schedule of the firm and self-select into their optimal time of purchase. From a privacy standpoint, intertemporal price discrimination is legally not problematic.

An important assumption of Stokey's model is that the firm can credibly commit to the price schedule. This assumption is by no means natural: It may be profitable *at the beginning* to choose the pricing strategy outlined above, but in the subsequent periods, the game changes and it may no longer be optimal to adhere to the strategy. If a strategy is optimal in any stage of the game, it is called *subgame perfect* (cf. Selten (1975)), and it follows that Stokey's model is lacking in subgame perfection. However, it is still conceivable that the monopolist could credibly commit: If the firm has stuck to the pricing strategy in the previous 100 iterations of the game, such credibility could be established. But in other cases, the monopolist might be tempted to switch to other pricing strategies down the road.

An alternative assumption regarding commitment credibility was made by Coase (1972). There, the monopolist is not able to make a credible commitment. If consumers are patient and have a fixed total demand, they can foresee the fall in prices induced by the sale of the good in previous periods: The durable goods monopolist has no choice but to sell at marginal cost in the first period.

Perfect foresight is another important assumption. If consumers are myopic, i. e. they do not anticipate changes in prices or they 'live from period to period', it is easily shown that a standard monopoly model emerges in which consumers buy naively in all periods if their surplus is positive (like Coase (1972) but with no strategic buying). Our simulation can be used to show that if the firm is very patient, while customers are not, price discrimination may be induced in the standard model without production costs or varying degrees of time preference, a result confirmed by the model in Varian (1989).

For goods where consumer entry timing is relevant, Bayer (2010) considered the economics of intertemporal price discrimination using a laboratory experiment. Contrastingly, in the model of Stokey (1979), all consumers entered the market simultaneously. Competition and monopoly pricing were studied under two different settings: Under the *early-bird* regime (in which consumers with a low valuation for the good start searching for low prices), price discrimination can be induced by selling to these low-valuation customers earlier for a lower price, while charging consumers with a higher valuation (that enter the market later) more—the price increases over time. In the *last-minute* situation, consumers with the high valuation enter the market first and therefore the price decreases. Experimentally, Bayer (2010) finds that under competition, the ability to price discriminate under *early-bird* is strengthened; under *last-minute*, charging the high-valuation consumers more than those with the low valuation is generally not possible. For the latter case, even the discriminatory power of the monopolist is questionable as the ability to price discriminate is eroded over multiple periods. Under Bayer (2010)'s experiment, many offers that are discriminatory are rejected by 'buyers', even when these offers would still be beneficial to the buyers. The empirical evidence is construed by Bayer (2010) to be an expression of inequity aversion towards discriminatory pricing (cf. Fehr and Schmidt (1999)). Price discrimination is not completely removed in the presence of a duopoly, but consumer welfare is enhanced. Under competition, prices are lower and the frequency of rejecting beneficial, but discriminatory offers is reduced.

Empirically, intertemporal price discrimination is often claimed. However, an empirical study needs to take into account other factors, like an intertemporal change in demand, changing costs or capacity constraints before intertemporal price discrimination can be assumed. Borenstein and Rose (1994) found that in the airline market, substantial price variation exists. The degree of price discrimination varies greatly between carriers and is indicative of price discrimination in a monopolistically competitive market, i. e. a market in which product differentiation alleviates competition. Interestingly, Borenstein and Rose (1994) show that airlines with a computer-based reservation systems engage in more price discrimination, hinting at an important role of technology. These results are corroborated by Escobari et al.

(2016) who claim that advance purchases of airline tickets are a device for price discrimination and tacit collusion due to focal points of consumer demand for tickets: 7 and 14 days before a flight, price variation is frequently observed, both on these routes that are served by multiple carriers and those that are not (although in the monopolistic case, more price discrimination occurs at the 7-day mark). In their sophisticated regression-discontinuity design, Escobari et al. (2016) find that price variation in the last month before a flight can be largely attributed to intertemporal price discrimination and not to other factors.

Empirical studies—both those that claim intertemporal price discrimination and those that do not—must not fail to account for changes in demand or capacity constraints. Stokey (1979) showed that an intertemporal price variation need not be because of pure price discrimination along the subdivision of reservation prices, but that other realities may contribute to such price variation. Among those realities considered were falling costs: In these situations, it would not be completely appropriate to claim price discrimination because the change in prices occurs (partly or mainly) for other reasons. As we do not live in a one-variable world, the price variation observed in reality will always be because of numerous factors, not all of which should be deemed discriminatory—but it is without question that a discriminating monopolist will gladly extract early adopters' utility, if at all possible.

3.2 Price discrimination by purchase history

In many markets, customers of other firms cannot easily switch to a competitor because switching incurs costs, monetary or otherwise. It has long been agreed that the presence of such switching costs may hinder competition and induce higher prices (cf. Klemperer (1987a) and Klemperer (1987b)) as well as lock-in and therefore reduced efficiency (cf. Farrell and Klemperer (2007) and Ewerhart and Schmitz (1997)). However, it is conceivable that by enticing the competitor's customers to switch with a monetary payment, the negative effects of being locked in can be alleviated. Such enticements are observed frequently: For example, German electricity and gas providers pay customers to switch, see figure 14.

The nature of switching costs is discussed in Klemperer (1995). Switching costs can take many forms, for example the time needed to learn how to use a new product. Another form of switching costs that seem relevant are the transaction costs arising from changing suppliers: Subscription-type purchases like those from an utility company are often automatically renewed at the end of a period, but both choosing not to renew and finding a new supplier are irrevocably tainted with transaction costs.

Figure 14: We are being offered € 280 to switch our gas provider.

In this section, we mainly consider a model of (competitive) price discrimination by purchase history that was initially developed by Chen (1997). It is a form of third degree price discrimination as customers pay different prices based on their purchase history—and this is where the subdivision of demand occurs.

3.2.1 Literature review

Fudenberg and Tirole (2000) is an important paper on poaching customers. They study a duopoly two-period model with consumers with brand preferences, no (exogenous) switching costs exist. Firms can offer long-term and short-term contracts; this is in contrast to Chen (1997) where only short-term contracts can be made. The preferences are either fixed or variable. Fudenberg and Tirole (2000) find that if preferences are constant, short-term contracts lead to much switching, thereby worsening welfare. In this case, allowing long-term contracts yields less switching and both types of contracts are sold. Under variable preferences, short-term contracts are optimal, but long-term contracts lead to too little switching.

Acquisti and Varian (2005) consider a two-period model in which a monopolistic seller can charge a different price to returning customers and new customers. Additionally, consumers are able to costly conceal their purchase history. They find that in the standard model in which consumers with perfect foresight are only differing in 'tastes' for the good, no price discrimination will be optimal. This result is in line with Stokey (1979). However, if the fraction of consumers that are myopic is large or if customers cannot effortlessly prevent the detection of their purchase history, conditioning prices based on purchase history may be profitable. Now consider the competitive case. If the firm is able to offer a personalized service package to consumers with a high valuation for the good that is more valuable than the one offered to low-valuation consumers, endogenous switching costs are generated, because switching to a competitor would require to rebuild the long-standing relationship between seller and customer. These switching costs once again lead to effective price discrimination. Acquisti and Varian (2005) conclude that in industries where purchasing history is easily ascertained or where personalized services are valuable, the firm wants to implement price discrimination. It is also noteworthy that the loyalty induced by endogenous switching costs may improve total welfare, ultimately leading to a change in attitude towards techniques that lower perceived privacy.

Chen et al. (2005) consider rebates as a means of price discrimination *within customers*.[25] It is a well-known fact that not all consumers redeem rebates, even though the redemption would lower the price paid. Chen et al. (2005) argue that rebates act as 'state-dependent discounts': As redeeming rebates is costly in itself (even though these costs may be nonmonetary), customers will only choose redemption if they are in a state in which the extra income derived through the rebate is highly useful. From this, it follows that a subdivision occurs along those who redeem and those who not redeem, leading not only to price discrimination within the consumer, but effectively also between consumers. How can it be explained that customers only redeem rebates in a state with a high marginal utility of income? Consider a product whose merit is unknown at the time of purchase. If the product performs well, there is little incentive to redeem the rebate. However, if the product performs badly, the consumer may be upset and want his money back, leading to a higher marginal utility of income. In this latter state, redemption will occur. Chen et al. (2005) note that for the consumer, the rebate acts like an insurance; and its mere presence leads to an increase in the consumer's willingness to pay because the consumer is protected against a potential downside of a product. This higher willingness to pay can subsequently be used to extract higher prices.

Price discrimination can also be induced through price-matching guarantees. Png and Hirshleifer (1987) discuss a model in which there are two groups of consumers—'tourists' and 'locals'—who are differently informed about the prices of competitors: Locals know all prices, while tourists know none. When one firm charges a low price, other firms can profitably raise prices: Locals that are better informed make use of price-matching guarantees and tourists are overcharged due to their ignorance of other prices. In essence, locals pay the competitive price while tourists do not. The price-matching guarantee acts as a mechanism for distinguishing between tourists and locals and therefore as an instrument of price discrimination. Png and Hirshleifer (1987) show that given information asymmetries, price discrimi-

[25] In the nomenclature by Chen et al. (2005), rebates are discounts that can be redeemed after a purchase, in contrast to coupons that are redeemed during the purchase.

nation can exist even in comparatively competitive markets and that price-matching guarantees provide a subdivision of demand along search costs.

3.2.2 Chen's model

Chen (1997) considers two models: A system under which firms can pay customers to switch, called PCTS, and a system under which firms are not allowed to pay customers to switch, called UNIF. Note that no price discrimination occurs under UNIF, but the firms can still make use of the switching cost borne by consumers. Under both regimes, firms are not able to contract for longer than one period. We will retrace nearly all steps of PCTS pricing in detail—even those not given in the original paper—and subsequently compare these results with those of Chen (1997) with regard to UNIF.

The model of Chen (1997) consists of a continuum of risk-neutral customers with mass 1 that can be distinguished by their switching cost s that is uniformly distributed on $[0, \Theta]$ and two firms A, B with constant marginal cost c that compete in two periods. In each period, each consumer will consume exactly one unit of the good. In the first period, both firms set prices, resulting in a market share α for firm A and $1-\alpha$ for firm B. At the beginning of the second period, consumers learn their s and firms again set prices, but under PCTS, a firm i can also offer a monetary payment m_i should the consumer switch to i, thereby artificially reducing the switching cost borne by the consumers. Firms and consumers discount future profits with δ. We assume that the concrete s is private (or hidden) information that cannot be ascertained by either firms; the firms only know the distribution of s amongst the population and each consumer's purchase history. The model of Chen (1997), which will be discussed in this section, does not distinguish between the different types of costs considered by Klemperer (1995)—it is only relevant that each consumer's switching costs are revealed at the beginning at the second period.

Paying customers to switch Let us first consider the PCTS regime. We will use backward induction to solve for subgame perfect Nash equilibria. The additional notation by Chen (1997) is as follows: q_{ij} is the quantity sold to consumers that bought from j in the first period, but buy from i in the second period; p_{i2} is the price charged by i in the second period and R is the willingness to pay for the product.[26] Consider a customer that previously bought from A. This customer will want to switch to B if his surplus from switching is higher than that from staying:

[26] Like Chen (1997), we always assume that R is high enough so that all prospective customers will want to buy from at least one firm.

$$R - p_{B2} - s + m_B > R - p_{A2}. \tag{31}$$

By rearranging the inequality, we see that if $s < p_{A2} - p_{B2} + m_B$, the consumer will switch to B. Otherwise, they stay with A. The quantity sold to consumers switching to B is, due to the uniform distribution of s,

$$\begin{aligned} q_{BA} &= \alpha \int_0^{p_{A2}-p_{B2}+m_B} \frac{1}{\Theta}\, ds, \\ &= \frac{\alpha m_B + \alpha p_{A2} - \alpha p_{B2}}{\Theta}. \end{aligned} \tag{32}$$

For those that do not switch, it holds that

$$\begin{aligned} q_{AA} &= \alpha \int_{p_{A2}-p_{B2}+m_B}^{\Theta} \frac{1}{\Theta}\, ds, \\ &= \frac{\Theta\alpha - \alpha m_B - \alpha p_{A2} + \alpha p_{B2}}{\Theta}. \end{aligned} \tag{33}$$

A consumer belonging to B's market share, $(1-\alpha)$, will switch to A if $R - p_{A2} - s + m_A > R - p_{B2}$, or, similarly as above, $s < p_{B2} - p_{A2} + m_A$. Now, therefore,

$$\begin{aligned} q_{AB} &= (1-\alpha) \int_0^{p_{B2}-p_{A2}+m_A} \frac{1}{\Theta}\, ds, \\ &= -\frac{(\alpha-1)m_A - (\alpha-1)p_{A2} + (\alpha-1)p_{B2}}{\Theta}. \end{aligned} \tag{34}$$

For B's loyal customers it necessarily holds that

$$\begin{aligned} q_{BB} &= (1-\alpha) \int_{p_{B2}-p_{A2}+m_A}^{\Theta} \frac{1}{\Theta}\, ds, \\ &= -\frac{\Theta\alpha - (\alpha-1)m_A + (\alpha-1)p_{A2} - (\alpha-1)p_{B2} - \Theta}{\Theta}. \end{aligned} \tag{35}$$

Using above quantities and paying attention to enticements, profits in the second period are $\pi_{A2} = (p_{A2} - c - m_A)q_{AB} + (p_{A2} - c)q_{AA}$ and $\pi_{B2} = (p_{B2} - c - m_B)q_{BA} + (p_{B2} - c)q_{BB}$, or, written extensively:

$$\pi_{A2} = (p_{A2} - c - m_A)\left(-\frac{(\alpha-1)m_A - (\alpha-1)p_{A2} + (\alpha-1)p_{B2}}{\Theta}\right) + (p_{A2} - c)\left(\frac{\Theta\alpha - \alpha m_B - \alpha p_{A2} + \alpha p_{B2}}{\Theta}\right),$$
$$\pi_{B2} = (p_{B2} - c - m_B)\left(\frac{\alpha m_B + \alpha p_{A2} - \alpha p_{B2}}{\Theta}\right) + (p_{B2} - c)\left(-\frac{\Theta\alpha - (\alpha-1)m_A + (\alpha-1)p_{A2} - (\alpha-1)p_{B2} - \Theta}{\Theta}\right).$$

By differentiating, we find, for each firm, the following first-order conditions:

$$\frac{\partial \pi_{A2}}{\partial p_{A2}} = \frac{\Theta\alpha - 2(\alpha-1)m_A - \alpha m_B + c - 2p_{A2} + p_{B2}}{\Theta} = 0, \quad (36)$$
$$\frac{\partial \pi_{A2}}{\partial m_A} = \frac{(\alpha-1)c + 2(\alpha-1)m_A - 2(\alpha-1)p_{A2} + (\alpha-1)p_{B2}}{\Theta} = 0, \quad (37)$$
$$\frac{\partial \pi_{B2}}{\partial p_{B2}} = -\frac{\Theta\alpha - (\alpha-1)m_A - 2\alpha m_B - \Theta - c - p_{A2} + 2p_{B2}}{\Theta} = 0, \quad (38)$$
$$\frac{\partial \pi_{B2}}{\partial m_B} = -\frac{\alpha c + 2\alpha m_B + \alpha p_{A2} - 2\alpha p_{B2}}{\Theta} = 0. \quad (39)$$

For a Nash equilibrium, it has to hold that the optimal p_{A2}, p_{B2}, m_A and m_B are *mutually best responses*. Therefore, the system of equations (36)–(39) needs to be solved simultaneously. As s is private information, only one tariff (p_{i2}, m_i) can be set per firm. The simultaneous solution, confirmed by Chen (1997), is $p^*_{A2} = p^*_{B2} = c + 2\Theta/3$ and $m^*_A = m^*_B = \Theta/3$.

Let us briefly reflect on that solution. First note that the optimal solution is independent of α. Secondly, note the discriminatory effect of this price: A customer that does not switch pays $c + 2\Theta/3$ and a customer that switches pays $c + 2\Theta/3 - \Theta/3 + s$. Since only those will switch that have $s < \Theta/3$, one third of customers switch; and those that do pay effectively *less* than those that do not switch. In the model of Chen (1997), new customers pay less than those that remain with a seller. Thirdly, note that the extracted prices increase in Θ, i. e. the higher switching costs are, the higher are second-period prices: The firms actively exploit lock-in. Finally, note that both prices are above marginal costs.
By inserting these results into the formulæ for the profits given above, and simplifying, we reach

$$\pi^*_{A2} = \frac{1}{3}\Theta\alpha + \frac{1}{9}\Theta, \quad (40)$$
$$\pi^*_{B2} = -\frac{1}{3}\Theta\alpha + \frac{4}{9}\Theta. \quad (41)$$

Since we have now calculated the optimal strategies in the second period, let us turn to the first period. In Chen (1997), in the first period, no consumer is attached to any firm and firms only set prices. Chen (1997) assumes that consumers want to buy from the lower-price firm; and if both firms have equal prices, they randomize. Therefore, note that under PCTS, α can only take one of the following values: $\{0, \frac{1}{2}, 1\}$. Again, we borrow Chen's notation: p_{i1} is the price of firm i in the first period.

Chen proposes the following subgame-perfect Nash equilibrium: $p^*_{A1} = p^*_{B1} = c - (\delta/3)\Theta$. From the original paper, Chen's derivation of the equilibrium is not too clear, but let us retrace how he likely found it. Consider the case of either firm, where $\alpha = \frac{1}{2}$. In this case, each firm's discounted period 2 profits are equal to $\delta\frac{5}{18}\Theta$ and each firm's period 1 profits are $\frac{1}{2}(p-c)$. Now consider the case that the firm we look at charges a marginally lower price, $p - \varepsilon$.[27] It then attracts the whole market share, leading to a discounted period 2 profit of $\delta\frac{4}{9}\Theta$ and a period 1 profit of $p - \varepsilon - c$. For the equilibrium price p in the case where $\alpha = \frac{1}{2}$, it must hold that the profit achieved with this price (weakly) dominates the total discounted profits in case the price $p - \varepsilon$ were chosen:

$$\frac{1}{2}(p-c) + \delta\frac{5}{18}\Theta \geq (p - \varepsilon - c) + \delta\frac{4}{9}\Theta. \tag{42}$$

It follows trivially that $p \leq c - (\delta/3)\Theta + 2\varepsilon$. Similarly, it has to hold that the profits on the left side (weakly) dominate the profits in case the price $p + \varepsilon$ were chosen:[28]

$$\frac{1}{2}(p-c) + \delta\frac{5}{18}\Theta \geq 0 + \delta\frac{1}{9}\Theta. \tag{43}$$

It follows that $p \geq c - (\delta/3)\Theta$. Inequalities (42) and (43) are only fulfilled simultaneously if $p = c - (\delta/3)\Theta$ and $\varepsilon = 0$. The above line of reasoning holds true for both firms and therefore, $p^*_{A1} = p^*_{B1} = c - (\delta/3)\Theta$ is indeed a mutually best response. $\alpha = \frac{1}{2}$ is induced, as demanded. But is it necessary that $\alpha = \frac{1}{2}$ or are there equilibria with $\alpha \neq \frac{1}{2}$? Chen shows that it is indeed the unique equilibrium (cf. Chen (1997), p. 888).

$p^*_{A1} = p^*_{B1} = c - (\delta/3)\Theta$ induce $\alpha = \frac{1}{2}$ and equal second-period profits, i. e. $\pi^*_{A2} = \pi^*_{B2} = \frac{5}{18}\Theta$. The optimal total discounted profits for period 1 and 2 are therefore

$$\pi^*_A = \pi^*_B = \frac{1}{2}\left(c - (\delta/3)\Theta - c\right) + \delta\left(\frac{5}{18}\Theta\right) = \delta\frac{1}{9}\Theta. \tag{44}$$

[27] ε is a small positive real number.

[28] Which would lead to no sales in the first period and leave the other firm with the total market share.

In many ways, this result of Chen (1997) is interesting. Firstly, the optimal prices in the second period are equal and not dependent on market share, in contrast to UNIF. Secondly, the prices charged in the first period are below marginal costs: Through low prices in the first period, the firms lure consumers so that they are locked in later; an effect that is amplified the higher switching costs are. Thirdly, despite constant marginal cost, both firms make a profit; but if Θ approaches zero, prices approach marginal cost and profits go to zero. Therefore, the presence of switching costs allows firms to charge collusive prices in the second period and make a profit *despite enticements and prices below marginal costs in the first period*, as opposed to the standard Bertrand game.[29]

No enticements Under UNIF, firms are not allowed to pay customers to switch. We will again use backward induction to solve for the optimal strategy, but given space constraints and the main interest of this thesis, we will discuss the results of the first period without derivation. The additional notation by Chen (1997) is as follows: p^u_{i2} is the second-period price of firm i, q^u_i is the second-period quantity sold by firm i and π^u_{i2} is the second-period profit.

Consider a customer that previously bought from A and suppose that $p^u_{A2} \geq p^u_{B2}$. This customer will prefer to stay with A if

$$R - p^u_{A2} < R - p^u_{B2} - s, \tag{45}$$

or, equivalently, if $s > p^u_{A2} - p^u_{B2}$. But note that no consumer that previously bought from B would want to switch to A! Not only would this consumer have to pay a higher price (since $p^u_{A2} > p^u_{B2}$), but he would also need to incur his switching cost s—which is not alleviated since no enticements may be paid by assumption. Therefore, it can *never* be profitable for a consumer in $1-\alpha$ to switch firms. However, this implies that all $1-\alpha$ customers of B will stay with B, and B will additionally receive those customers for which inequality (45) is not fulfilled. It follows that

$$q^u_B = \underbrace{\alpha \int_0^{p^u_{A2}-p^u_{B2}} \frac{1}{\Theta}\, ds}_{\text{switching to } B} + \underbrace{1-\alpha}_{\text{staying with } B} = -\frac{\Theta\alpha - \alpha p^u_{A2} + \alpha p^u_{B2} - \Theta}{\Theta}, \tag{46}$$

[29] In the standard Bertrand game, the presence of only two firms suffices to achieve prices that are equal to marginal costs, cf. Bester (2012), pp. 100 *et seq.*

$$q_A^u = \underbrace{0}_{\text{switching to } A} + \underbrace{\alpha \int_{p_{A2}^u - p_{B2}^u}^{\Theta} \frac{1}{\Theta}\, ds}_{\text{staying with } A} = \frac{\Theta\alpha - \alpha p_{A2}^u + \alpha p_{B2}^u}{\Theta}. \quad (47)$$

In this case, profits are given as $\pi_{A2}^u = (p_{A2}^u - c) q_A^u$ and $\pi_{B2}^u = (p_{B2}^u - c) q_B^u$, respectively. Maximizing them with regard to prices, and finding mutually best responses, we reach to following second-period prices: $p_{A2}^{u*} = \frac{\Theta\alpha + 3\,\alpha c + \Theta}{3\,\alpha}$ and $p_{B2}^{u*} = -\frac{\Theta\alpha - 3\,\alpha c - 2\,\Theta}{3\,\alpha}$. We find that in this case, $p_{A2}^{u*} \geq p_{B2}^{u*}$ only holds if $\alpha \geq \frac{1}{2}$.

If we consider the case where $p_{A2}^u < p_{B2}^u$, we use the same line of reasoning as above: No one would switch from A to B, but the reverse is not true. We find that quantities are then given by

$$\begin{aligned} q_A^u &= \underbrace{(1-\alpha) \int_0^{p_{B2}^u - p_{A2}^u} \frac{1}{\Theta}\, ds}_{\text{switching to } A} + \underbrace{\alpha}_{\text{staying with } A} \\ &= \frac{\Theta\alpha + (\alpha - 1) p_{A2}^u - (\alpha - 1) p_{B2}^u}{\Theta}, \quad (48) \\ q_B^u &= \underbrace{0}_{\text{switching to } B} + \underbrace{(1-\alpha) \int_{p_{B2}^u - p_{A2}^u}^{\Theta} \frac{1}{\Theta}\, ds}_{\text{staying with } B} \\ &= -\frac{\Theta\alpha + (\alpha - 1) p_{A2}^u - (\alpha - 1) p_{B2}^u - \Theta}{\Theta}, \quad (49) \end{aligned}$$

and that the best responses (resulting from bilateral profit maximization) are $p_{A2}^{u*} = -\frac{\Theta\alpha - 3\,(\alpha-1)c + \Theta}{3\,(\alpha-1)}$ and $p_{B2}^{u*} = \frac{\Theta\alpha + 3\,(\alpha-1)c - 2\,\Theta}{3\,(\alpha-1)}$. $p_{A2}^{u*} < p_{B2}^{u*}$ holds if $\alpha < \frac{1}{2}$.
Concluding, we have these optimal prices in the second period (cf. figure 15):

$$p_{A2}^{u*} = \begin{cases} \frac{\Theta\alpha + 3\,\alpha c + \Theta}{3\,\alpha} & \text{if } \alpha \geq \frac{1}{2}, \\ -\frac{\Theta\alpha - 3\,(\alpha-1)c + \Theta}{3\,(\alpha-1)} & \text{if } \alpha < \frac{1}{2}, \end{cases} \quad (50)$$

$$p_{B2}^{u*} = \begin{cases} -\frac{\Theta\alpha - 3\,\alpha c - 2\,\Theta}{3\,\alpha} & \text{if } \alpha \geq \frac{1}{2}, \\ \frac{\Theta\alpha + 3\,(\alpha-1)c - 2\,\Theta}{3\,(\alpha-1)} & \text{if } \alpha < \frac{1}{2}. \end{cases} \quad (51)$$

From that, we can calculate optimal profits and analyze the first period under UNIF using backward induction. Consumer choice in the first period will, among other things, be based on the expected utility derived in period

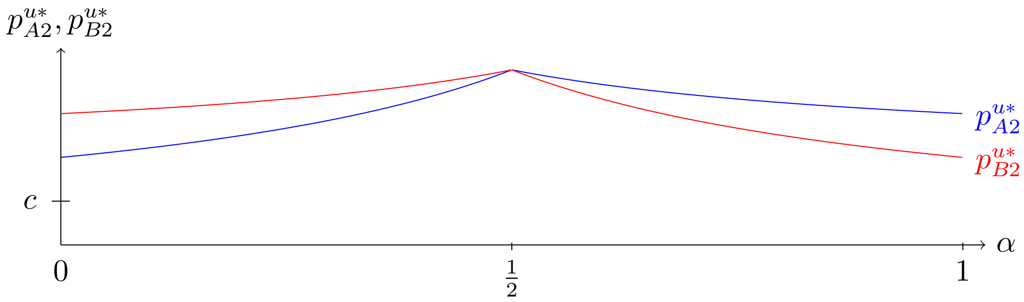


Figure 15: Optimal prices in the second period under UNIF. Note that the firm with the higher market share charges a higher price, in contrast to PCTS.

2, cf. Chen (1997), p. 889.[30] It follows that, under UNIF, α can take any value from 0 to 1.

Let us briefly summarize Chen's additional results that will not be covered extensively here: Under UNIF, there is one interesting equilibrium at $p^{u*}_{A1} = p^{u*}_{B1} = c + \frac{2}{3}\Theta\delta$, resulting in total discounted profits of $\pi^{u*}_{A} = \pi^{u*}_{B} = \frac{5}{6}\Theta\delta$. However, there are numerous subgame perfect equilibria, all of whose profits are weakly higher than under PCTS. Finally, total welfare is lower under PCTS, but consumer welfare (the sum of consumer utilities) may be higher or lower.

3.2.3 Summary

Chen (1997) presented a model of price discrimination in which firms can pay customers to switch. It has been shown that in equilibrium, firms will set prices below marginal costs in the first period and prices above marginal costs in the second period, leading to positive total profits. The presence of switching costs makes the society worse off, and firms are generally worse off under PCTS than under UNIF. Consumers can be better or worse off.

In Chen's model, the price discrimination occurs based on the purchasing history of the consumer. In the second period under PCTS, consumers that switch from one firm to the other pay $c + \frac{2}{3}\Theta - \frac{1}{3}\Theta + s = c + \frac{1}{3}\Theta + s$. Those that do not switch pay $c + \frac{2}{3}\Theta$. As only those switch for which it holds that $s < \frac{1}{3}\Theta$, switchers pay effectively less than consumers who do not switch. That is the discriminatory effect in the model of Chen (1997). There is another implication: In section 1, we showed that price discrimination often encounters legal problems. However, in Chen's PCTS model, those that switch pay a lower price and every consumer is therefore equally or

[30] Remember that consumers learn their s only at the beginning of the second period.

better off when revealing their purchase history. Therefore, it can be safely assumed that all such information will be voluntarily provided by consumers, avoiding conflict with privacy laws.

The existence of switching costs is very plausible: In Chen's model, prices are common knowledge in any stage, and therefore Chen's s is not to be understood to be 'search costs', but actual costs borne when switching. Each consumer's s is private and only revealed at the beginning of the second period. In the context of subscription-type purchases, such true switching costs exist: It is often necessary to draw up letters to cancel the subscription, take the letter to the post office, and so on. By letting firms entice customers to switch, the disutility experienced by consumers is alleviated and in the model by Chen (1997), one third of customers actually switch.

We showed that firms are worse off under PCTS than under UNIF. Why, then, bother analyzing PCTS if no rational firm would implement it? The answer lies in the interaction between competing firms. First of all, it is true that both firms could be made better off without PCTS—if they could coordinate. However, it remains to be analyzed what would occur if only one of the two firms offered enticements to lessen switching costs, but it is possible that given the other firm's inaction, implementing an enticement scheme could be a profitable strategy, depending on the realities that the firm faces. Both parties recognize this possibility and achieve a Pareto-inferior outcome, namely PCTS in the whole market. Therefore, the choice between PCTS and UNIF could be modeled akin to the original prisoner's dilemma. However, these deductions need to be taken with a grain of salt as more research into why schemes that pay customers to switch are implemented in the first place would be desirable.

Technology plays a significant role in price discrimination by purchase history. On the one hand, switching and search costs may be reduced. On the other hand, technology makes it easier to offer premiums for switching providers. The screenshot in figure 14 would not have been possible without technology, and it is hard to imagine that such tariffs could be offered without the precise information we knowingly or unknowingly provided over the internet. Technology makes paying customers to switch easier. Purchase history recognition can lead to price discrimination in other models as well: For example, the role of technology was emphasized in Acquisti and Varian (2005) who even hinted at positive welfare effects due to personalized services, which led to the creation of endogenous switching costs.

4 Conclusion

In this thesis, we discussed numerous variants and applications of price discrimination. Firstly, we gave an introduction. We also explained the rationale behind an adoption of price discrimination and gave remarks about the necessary prerequisites for price discrimination, always distinguishing between subdividing demand and physically charging different prices. We emphasized that modern technology assists the discriminating firm in both respects, but that such efforts can be curtailed by statute, irrespective of them possibly improving welfare. Pricing experiments were highlighted as a means to reveal the subdivision of demand.

Thereafter, we introduced the readers to the three classic degrees of price discrimination as they were first thought of by Pigou (1920). We adapted standard accounts by literature on industrial organization and enriched it. We initially presented the readers with principles of quasilinear utility. Quasilinear utility is an useful simplification that excludes income effects. For nonnegative incomes, the optimal consumption choice only depends on the price of the good and at the optimum, marginal utility equals price. For our subsequent analysis, we normalized incomes to zero, without loss of generality.

We then discussed first degree price discrimination, or perfect price discrimination. In it, the monopolist can charge a total payment of $p_i x + a_i$ if a consumer of type i purchases x units of the good we consider. We found that the monopolist chooses to charge prices that are all equal to marginal cost. However, a_i is set so that all otherwise remaining utility (the surplus) of the customers is seized by the firm. It is important to note that perfect price discrimination is a welfare optimum, although the monopolist absorbs all welfare.

Second degree price discrimination is characterized by a total payment of $px + a$, where (notably) p and a are identical amongst all consumers. Therefore, in this variant of price discrimination, no actual differentiation in prices takes place. The discrimination is outsourced to consumers who self-select either into buying or not buying. Under perfect information, the monopolist will choose a so that some marginal consumer's full utility is captured—those that have lower utility than this marginal consumer choose not to consume, but those with higher utility are left with positive surplus from consumption. From that, an optimal p can be calculated arithmetically. But choosing the marginal consumer itself is nontrivial. We commissioned a simulation with one hundred consumer types and found that it is generally profitable to exclude some types by charging a higher a. As more and more consumer types are excluded, p approaches marginal cost. In an extension of the model, multiple nonlinear tariffs can be offered. Second degree price discrimination is interesting from a mechanisms standpoint as well. In first and third degree price discrimination, consumers can profit from pretending to be of another type or to belong to another market; this is not the case with second degree price discrimination as all tariffs are available to anyone. Consumers are therefore best-off when revealing their true preferences, marking the strategyproofness of second degree price discrimination.

Ultimately, third degree price discrimination was discussed. This degree of price discrimination is based only on a differentiation in prices. There is no fixed payment. We found that third degree price discrimination is identical to a monopolist that sells the same good on different markets. It followed that on each such market, marginal revenue equals marginal cost, a conclusion also true for the monopolist that engages in third degree price discrimination. We also found that for some classes of utility functions, an affine transformation of such an utility function does not induce third degree price discrimination. In reality, variants of third degree price discrimination are often observed.

Thereafter, we turned to more recent concepts of price discrimination. At first, we discussed intertemporal price discrimination based on Stokey (1979). We outlined the assumptions of Stokey's model and found that its assumptions conflict with the 'Coase conjecture' (Coase (1972)) in that in Stokey's model, the monopolist can credibly commit to certain prices down the road. However, it is unclear how such commitment could be credible given the new game of price setting in later periods. This lack of subgame perfection is a weakness of Stokey's model, although mechanisms for commitment are conceivable, for example if the firm has a reputation for sticking to its pricing plans. As some methods used in Stokey (1979) are too advanced to be discussed in detail in this thesis, we discussed some of her results informally without derivation and referred to Chiang (2000) for a textbook on the calculus of variations.

The surprising result of Stokey's model is that for a given class of utility functions, no price discrimination will be implemented. It can be profitable, however, if the discount rates of consumers differ and it is generally profitable if production costs fall. However, costs need to fall rapidly. We simplified her continuous time model into a two-period model, commissioned a simulation of our model and found that all of Stokey's results can be verified. Our simulation permits the consideration of related cases. We noted that other factors like changing demand contribute to (nondiscriminatory) price variation, a notion that is not to be neglected in empirical studies. The crucial and corroborated result of Stokey (1979)—that often, no intertemporal price discrimination is profitable—combined with the seeming paradox of Coase (1972) suggests that in those empirical studies that claim intertemporal price discrimination, other factors like changing demand or falling costs might be relevant and that *ceteris paribus* may be violated. However, more sophisticated approaches like those presented in Borenstein and Rose (1994) and Escobari et al. (2016) plausibly hint at the existence of price discrimination even in competitive markets.

Then, we discussed price discrimination by purchase history. This discussion was largely based on the seminal paper by Chen (1997). In the duopoly framework considered, firms can pay customers to switch. His PCTS model is characterized by prices below marginal costs in the first period and collusive prices in the second period. In the second period, firms offer a monetary payment of one third the maximal switching cost, thereby inducing one third of customers to switch. Chen subsequently compares his PCTS regime with one in which firms are not allowed to pay customers to switch, called UNIF. We considered this regime mainly in an informal manner. Under UNIF, a multitude of equilibria emerge, all of whose firm profits are higher or equal compared to the profits achieved under PCTS.

The perspective of welfare is also interesting. Chen (1997) demonstrates that the society is worse off under PCTS, but consumers may be better or worse off compared to UNIF. But if industry profits are lower under PCTS, why is it so frequently observed? The answer may lie in a prisoners' dilemma: Market realities can make it profitable to pay customers to switch, given that the other firms have not adopted such a scheme. Ultimately, every firm has adopted PCTS, resulting in a Pareto-inferior outcome amongst firms.

Both Chen (1997) and Stokey (1979) are variants of third degree price discrimination, and third degree price discrimination is very common in reality. Additionally, both models also used quasilinear utility in that income was not a factor regarding the optimal level of consumption. The role of technology is ambiguous: On one hand, technology eases recognition, market research and putting up different prices, but on the other hand, technology enables more competition and reselling, thereby thwarting much of price discrimination. In the context of intertemporal price discrimination, technology can also be the *subject* of price discrimination.

Simulation was used substantially in this thesis, showing that it is an attractive method for approaching problems in the field of managerial economics. However, simulations cannot exist without some theory about firm and consumer behaviour: By combining theory and simulation, this thesis aimed to demonstrate that while the simple problems considered here are easily expressed in computational models, simulation as a method permits enormous flexibility—allowing for a much broader spectrum of hypotheses to be tested and their implications to be observed. Importantly, our use of simulation also showed how profits change if *sub*optimal policies are implemented.

Despite substantial effort, not all price variations observed in reality can be explained by standard models. The models presented in this thesis provide an overview and give reasonable insights into pricing behaviour, although some mysteries persist. One of these mysteries is how price discrimination can persist in markets that are more competitive than a monopoly. In imperfect markets, such as these with switching costs under a duopoly, price discrimination can persist. However, much more remains to be seen: For example, what occurs in a market with switching costs if there are not two, but n firms? The implications of competition on price discrimination remain an active and intriguing field of study to which economists are encouraged to contribute.

5 References

Acquisti, A. and Varian, H. R. (2005). Conditioning prices on purchase history. *Marketing Science*, 24(3):367–381.

Bayer, R.–C. (2010). Intertemporal price discrimination and competition. *Journal of Economic Behavior & Organization*, 73(2):273–293.

Berners–Lee, T., Fielding, R., and Frystyk, H. (1996). RFC 1945: Hypertext Transfer Protocol—HTTP/1.0. `https://tools.ietf.org/html/rfc1945`. [Online, accessed June 6, 2017].

Bester, H. (2012). *Theorie der Industrieökonomik*. Springer-Lehrbuch. Gabler Verlag, 6th edition.

Bewersdorff, J. (2014). Javascript im Browser. In *Objektorientierte Programmierung mit JavaScript*, pages 202–208. Springer.

Blair, R. D. and DePasquale, C. (2014). “Antitrust's Least Glorious Hour”: The Robinson-Patman Act. *The Journal of Law and Economics*, 57(S3):S201–S216.

Borenstein, S. and Rose, N. L. (1994). Competition and price dispersion in the us airline industry. *Journal of Political Economy*, 102(4):653–683.

Chen, Y. (1997). Paying customers to switch. *Journal of Economics & Management Strategy*, 6(4):877–897.

Chen, Y., Moorthy, S., and Zhang, Z. J. (2005). Research note—price discrimination after the purchase: Rebates as state-dependent discounts. *Management Science*, 51(7):1131–1140.

Chiang, A. C. (2000). *Elements of dynamic optimization*. Illinois: Waveland Press Inc.

Coase, R. H. (1972). Durability and monopoly. *The Journal of Law and Economics*, 15(1):143–149.

Darmstädter Echo (2017). Manche Volksbanken in Südhessen zeigen sich bei Gebühren am Geldautomaten einfallsreich. `http://www.echo-online.de/wirtschaft/wirtschaft-suedhessen/manche-volksbanken-in-suedhessen-zeigen-sich-bei-gebuehren-am-geldautomaten-einfallsreich_17816659.htm`. [Online, accessed June 6, 2017].

ePrivacy Directive (2002). Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector. Standard, European Parliament, Brussels.

Escobari, D., Rupp, N. G., and Meskey, J. (2016). Price discrimination in advance-purchase discounts with focal points for tacit collusion: Regression-discontinuity evidence from airlines.

Ewerhart, C. and Schmitz, P. W. (1997). Der Lock-in-Effekt und das Hold-up-Problem. *Wirtschaftswissenschaftliches Studium*, 26(7):360–363.

Farrell, J. and Klemperer, P. (2007). Coordination and lock-in: Competition with switching costs and network effects. *Handbook of Industrial Organization*, 3:1967–2072.

Fehr, E. and Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3):817–868.

Fudenberg, D. and Tirole, J. (2000). Customer poaching and brand switching. *RAND Journal of Economics*, pages 634–657.

Gao, M., Hyytinen, A., and Toivanen, O. (2004). Demand for mobile internet: Evidence from a real-world pricing experiment. Technical report, ETLA Discussion Papers, The Research Institute of the Finnish Economy (ETLA).

ISO/IEC 7498-1(E) (2000). Information Processing Systems, Open System Interconnection (OSI) Reference Model, The Basic Model, ITU-T Rec. X.200. Standard, International Organization for Standardization, Geneva.

Klemperer, P. (1987a). The competitiveness of markets with switching costs. *The RAND Journal of Economics*, 18(1):138–150.

Klemperer, P. (1987b). Markets with consumer switching costs. *The Quarterly Journal of Economics*, 102(2):375–394.

Klemperer, P. (1995). Competition when consumers have switching costs: An overview with applications to industrial organization, macroeconomics, and international trade. *The Review of Economic Studies*, 62(4):515–539.

Lam, T. C. and Small, K. A. (2001). The value of time and reliability: measurement from a value pricing experiment. *Transportation Research Part E: Logistics and Transportation Review*, 37(2):231–251.

Malakooti, B. (2013). *Operations and production systems with multiple objectives*. John Wiley & Sons.

O'Brien, D. P. and Shaffer, G. (1994). The welfare effects of forbidding discriminatory discounts: A secondary line analysis of Robinson-Patman. *Journal of Law, Economics, & Organization*, pages 296–318.

Oi, W. Y. (1971). A Disneyland dilemma: Two-part tariffs for a Mickey Mouse monopoly. *The Quarterly Journal of Economics*, 85(1):77–96.

Pew Research Center (2016). 28% of Americans are 'strong' early adopters of technology. `http://www.pewresearch.org/fact-tank/2016/07/12/28-of-americans-are-strong-early-adopters-of-technology/`. [Online, accessed June 6, 2017].

Pigou, A. C. (1920). *The Economics of Welfare*. McMillan&Co., London.

Png, I. P. and Hirshleifer, D. (1987). Price discrimination through offers to match price. *Journal of Business*, pages 365–383.

Poese, I., Uhlig, S., Kaafar, M. A., Donnet, B., and Gueye, B. (2011). IP geolocation databases: Unreliable? *ACM SIGCOMM Computer Communication Review*, 41(2):53–56.

Postel, J. (1980). RFC 760: DoD Standard Internet Protocol. `https://tools.ietf.org/html/rfc760`. [Online, accessed June 6, 2017].

R Core Team (2017). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

Schmitz, P. W. (2016). Managerial Economics. *Lecture material, University of Cologne.*

Selten, R. (1975). Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4(1):25–55.

Sexton, R. J., Johnson, N. B., and Konakayama, A. (1987). Consumer response to continuous-display electricity-use monitors in a time-of-use pricing experiment. *Journal of Consumer Research*, 14(1):55–62.

Stokey, N. L. (1979). Intertemporal price discrimination. *The Quarterly Journal of Economics*, pages 355–371.

Varian, H. R. (1989). Price discrimination. *Handbook of Industrial Organization*, 1:597–654.

Varian, H. R. (1992). *Microeconomic Analysis.* W. W. Norton & Company, Inc.

Varian, H. R. (2011). *Grundzüge der Mikroökonomik.* Walter de Gruyter GmbH & Co. KG.

Wilson, R. B. (1993). *Nonlinear pricing.* Oxford University Press on Demand.

A Appendix

A.1 The value of a

In section 2.3, we claimed that a and f are dependent in the following fashion:

$$a^*(f,p) = V_f(x_f^*(p)) - px_f^*(p)\,.$$

This fixed payment implies that type f's full surplus $\mathfrak{U}_f(x_f^*(p))$ is captured by the firm.

It is easily proven that this relation must hold in any profit optimum in which p is optimized for a given f. Assume that ε is a sufficiently small positive real number.

Proof by contradiction.

1. Consider the case where $a = a^*(f,p) + \varepsilon$. Type f would not consume because $a^*(f,p)$ is already the highest allowed fixed payment. Now, therefore, if f does not consume, profit is not being 'optimized for a given f', but for $f+1$, violating the above definition of profit optimization.

2. Consider the case where $a = a^*(f,p) - \varepsilon$. Type f would consume and retain some utility ε. But f would still consume if a were ε higher, and profits would be higher as well. Therefore, this cannot be a profit optimum.

This line of reasoning holds for any f, and therefore, by induction, the above is proven. □

A.2 Affine transformations

Some classes of utility functions have the property that their affine transformations $W(x) = aV(x) + b$, with $a > 0$, are insufficient to induce third degree price discrimination. One such class is the one for which it holds that

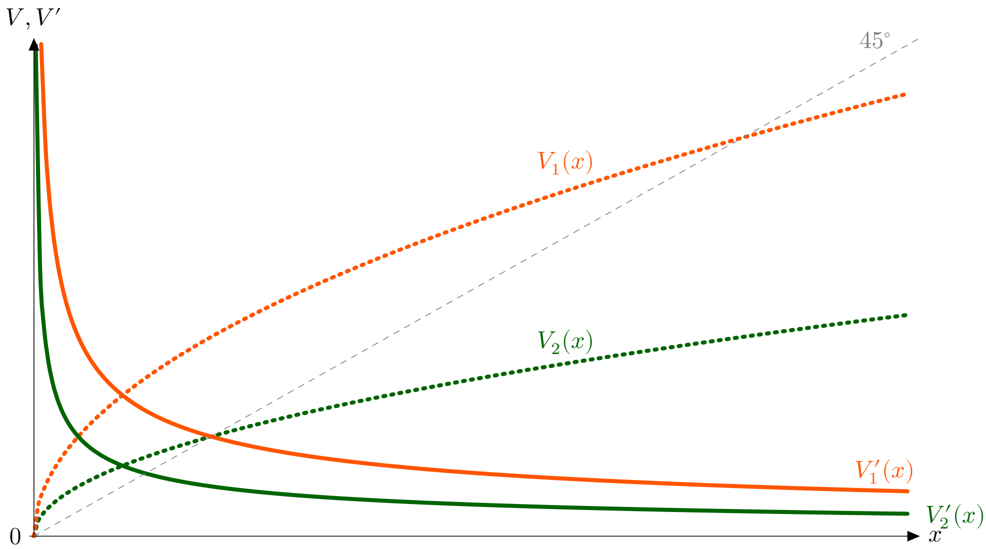
$$\frac{\partial}{\partial x}V(x) = V'(x) = x^r \quad (r < 0,\ r \in \mathbb{R}). \tag{52}$$

The proof is trivial and left to the readers. Note that equation (18) gave a necessary condition for third degree price discrimination. Let us rearrange that equation:

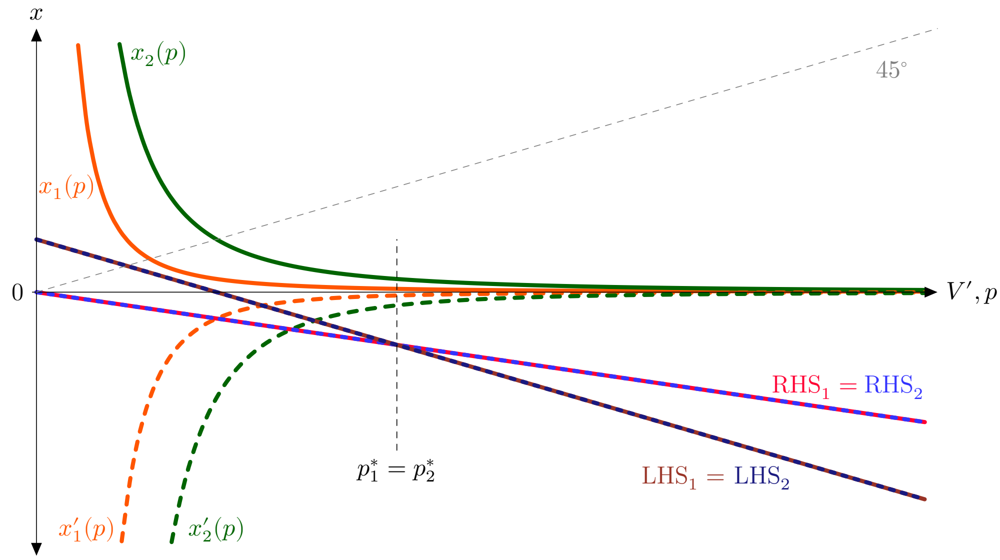
$$\underbrace{c - p_i}_{\text{LHS}_i} = \underbrace{\frac{x_i^* \left(p_i\right)}{\frac{\partial}{\partial p_i} x_i^* \left(p_i\right)}}_{\text{RHS}_i} \quad \forall i.$$

(We will now simplify the notation, dropping indices and asterisks wherever possible; in contrast to main parts of the thesis, we will use the prime notation $f'(x)$ to refer to the first derivative of a function $f(x)$.)

Now consider the following two groups of customers: Type 1 has the utility function $V_1(x) = \sqrt{x}$. Type 2 has the utility function $V_2(x) = 2\sqrt{x}$. In the following figure, their utility functions and the first derivative of their utility functions are shown:



Their optimal level of consumption follows immediately from equation (4), as no fixed payment is charged, their utiliy functions are strictly concave and it holds that $\left.\frac{\partial V_i(x)}{\partial x}\right|_{x=0} = \infty$. Therefore, for any price p, it holds that $V_i'(x) = p$, and the optimal level of consumption at any price is the inverse of $V_i'(x)$, i. e. $x_i(p) = \left(V_i'\right)^{-1}(x)$. In the next figure, we remove $V_1(x)$ and $V_2(x)$ and invert $V_1'(x)$ and $V_2'(x)$ by rotating the previous figure at the 45° line. We get the demand functions $x_1(p)$ and $x_2(p)$. We also draw LHS_i and RHS_i and, as it turns out, $\text{RHS}_1 = \text{RHS}_2$. ($\text{RHS}_i$ is the quotient of $x_i(p)$ and $x_i'(p)$, and the latter is therefore also shown.) The optimal prices are then given by the intersection of LHS_i and RHS_i, yielding the exact same price. Therefore, it is shown that affine transformations of some classes of utility functions do not induce third degree price discrimination.



x
$x_2(p)$
45°
$x_1(p)$
0
V', p
RHS$_1$ = RHS$_2$
$p_1^* = p_2^*$
LHS$_1$ = LHS$_2$
$x_1'(p)$
$x_2'(p)$