International Journal of Industrial Organization 30 (2012) 605–614

Contents lists available at SciVerse ScienceDirect

International Journal of Industrial Organization

journal homepage: www.elsevier.com/locate/ijio

Customer-class pricing, parallel trade and the optimal number of market segments ☆

Yann Braouezec

IESEG School of Management, LEM-CNRS, Paris La Défense, France

ARTICLE INFO

Article history:
Received 14 September 2010
Received in revised form 26 June 2012
Accepted 27 June 2012
Available online 5 July 2012

JEL classifications:
D42
L11
L50

Keywords:
Monopoly
Third-degree price discrimination
Endogenous market segmentation
Parallel trade

ABSTRACT

We consider the optimal market segmentation problem of a monopolist that faces a *continuum* of customers when it is costly to prevent resale (or parallel trade) among groups. In our framework, the monopolist chooses the number $k \geq 1$ of market segments, but also their design and the discriminatory prices. All these quantities are chosen to maximize the total profit. We solve the profit maximization problem when demands are linear and parallel as a function of the cost of separating markets. We show that market segmentation and prices cannot be chosen independently, and we also show that it is optimal to create only a few market segments. We then turn to the welfare analysis and show that the socially optimal number of market segment is equal to three.

1. Introduction

Price discrimination is a common practice both in the private sector and the public sector (e.g., McAfee, 2008; Steinberg and Weisbrod, 2005). It is well-known that hotels, car manufacturers, airline companies, but also national museums and galleries charge a different price for the same good or service to different subsets of customers. For a private firm, price discrimination is seen as a natural way to increase its profit; the set of customers is segmented on the basis of some observable characteristics (e.g., age, geography, gender) related to their price sensitivity, and customers of different market segments are charged a different price. For example, the price of a train ticket can differ according to whether the purchaser is a senior citizen or not while the price of medications or cars can depend on the country where the purchase is made (McAfee, 2008; Goldberg and Verboven, 2001). This customer-class pricing is generally called third-degree price discrimination from the pioneered work of Arthur Pigou (1932). In the monopoly literature on the subject, it is generally assumed that the underlying set of customers has already been divided in a *given number of perfectly sealed* market segments (i.e., classes, groups or sub-markets), so that the monopolist problem reduces to a pure pricing one.

We argue that this rather simplistic picture of the monopolist problem is surprising not only because both the design and the number of market segments are critical variables of the overall marketing strategy, but also because this simplistic picture provides no clues to explain the observed customer-class pricing. Since the aim of the monopolist is obviously to charge as many prices as there are customers, why do most firms choose to create few market segments compared to the large number of customers? For a discriminating monopolist, a realistic explanation may be the cost associated to the prevention and detection of parallel trade (i.e., arbitrage or gray market). It is well-known that when there are large price differences, this may give a strong incentive for some agents to engage in arbitrage (e.g., McAfee, 2008; Anderson and Ginsburgh, 1999). This phenomenon occurs if customers of a high-priced market segment can easily buy the good from customers of low-priced market segments, but, and more critically, if there exist parallel traders (or gray marketers) that do business by selling the goods from the low-priced market segment to the high-priced market segment without the authorization of the monopolist, the intellectual property rights (IPR) owner. It turns out that via these unauthorized channels distribution may be legal.

Consider the case of a manufacturer (of a new patented good) that wishes to charge a different price in each country of the European Union (EU). Since the EU is considered as a single market, parallel imports are consistent with the principle of free movement of goods so

☆ I'm grateful to three anonymous referees and to the associate editor, Ulrich Doraszelski, for their insightful comments and remarks. I also want to thank Eric Avenel, Pascal Belan, Pascal Bouyaux, Cyril Grunspan, Fabien Moizeau, Thierry Penard and Régis Renault for their input on an earlier version of this paper. The usual disclaimers obviously apply.

E-mail address: y.braouezec@ieseg.fr.

0167-7187/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijindorg.2012.06.006

that the European Court of Justice has established a policy of "community exhaustion" of rights.[1] This means that once the good has been sold by the monopolist (the IPR owner) within the EU, its rights are exhausted. The buyer can thus freely resell the patented good without the authorization of the rights owner. On the contrary, reimportation from a country outside the EU is not allowed. For the particular case of medications, report that 20% of the (branded) pharmaceuticals in the U.K was sold via parallel trade while Arfwedson (2004) reports that within the European Union, parallel import of drugs were estimated to represent 3.3 billion dollars in 2001. A wide variety of industries such as personal computers, perfumes, watches etc. are actually subject to parallel trade. Antia et al. (2006) report that estimates of gray market activity range from 10 billion to 20 billion dollars in economy-wide annual gray market sales. In response to parallel trade, a manufacturer's strategy may simply be to charge a uniform price (Danzon, 1997). However, as noted by Cavusgil and Sikora (1988) in their early paper, and more recently among others by Ahmadi and Yang (2000), Antia et al. (2004) and Myers and Griffith (1999), this parallel trade problem should be anticipated by the discriminating monopolist, and the nonpricing responses to combat it should be part of the marketing strategy. It is only when these proactive strategies are too costly that it may perhaps be optimal for the monopolist to charge a uniform price. In Gerstner and Holthausen (1986), Ahmadi and Yang (2000) and more recently in Zhang and Bell (2010), they also consider a model in which isolating each market segment is costly but they assume that the design and the number of market segments are exogenous. It is thus the aim of the present paper to provide a model in which the monopolist chooses the design, the prices, and the number of market segments when it is costly to prevent parallel trade.

We analyze the simplest framework in which there is a *continuum* of customers, each endowed with a linear demand function. Since it does not make sense in practice to charge a *continuum* of prices, the monopolist must segment the set of potential customers into classes, but the number of these classes may be very large. Our results reveal several interesting properties. We show that when the monopolist decides to create $k+1$ market segments rather than k, this changes both the existing market segments and the prices. This shows that market segmentation and prices cannot be chosen separately. When the cost associated to the prevention of parallel trade is low, we show that it is never optimal to create more than a small number of market segments. This result may explain why in practice the number of market segments is low compared to the high number of customers. We then turn to the welfare analysis assuming that the regulator can costlessly eradicate parallel trade by making it illegal. We show that the socially optimal number of market segments is equal to three (the normative benchmark), which means that the monopolist should be allowed to charge three discriminatory prices. This result then allows us to compare this socially optimal number with the actual one chosen by the monopolist as a function of the isolation cost. When this cost is high, so that it is optimal for the monopolist to choose a single price, we show that everything is as if a fraction of customers were subsidized by another one; approximately 40% of customers with intermediate willingness-to-pay (henceforth WTP) subsidize 30% of those with the highest WTP. On the other hand, when the cost is low so that it is optimal for the monopolist to create more than three market segments, customers with lower WTP are now served and everything is as if they were subsidized by those with the highest WTP. This shows that deviations from the normative benchmark are not symmetric; there is a greater regulatory concern when one market segment is created rather than more than three.

The second section is devoted to the presentation of the model while the third one is devoted to the presentation of the results. All the proofs are relegated to Appendix A.

2. The model

2.1. Assumptions and discussion

We consider a simple model in which the set of potential customers Ω has the cardinality of the *continuum*. We focus more specifically on the case in which $\Omega := [0, 1]$, where the distribution of the customers in [0, 1] is uniform. Using the notation $x^+ = \max\{0, x\}$, we further assume that the demand function of each customer $\omega \in [0, 1]$ is linear[2]

$$q(\omega, P) = (\omega - P)^+, \tag{1}$$

where the choke price ω is interpreted as a willingness-to-pay (WTP). We consider the case of a fully informed monopolist that produces a good at a constant marginal cost $C_m \geq 0$. In what follows, we assume that $C_m = 0$ to simplify the analysis, but all the results of this paper are unchanged if we assume that ω is uniformly distributed in $[C_m, 1 + C_m]$ when C_m is positive (see Appendix A). From Eq. (1), it thus follows that when the monopolist charges the price P, its revenue from a customer (or household) ω is equal to

$$\pi(\omega, P) = P(\omega - P)^+ \forall \omega \in [0, 1]. \tag{2}$$

Following the early paper of Malueg and Schwartz (1994), we call *complete discrimination* the case in which the monopolist may charge a customized price $P(\omega)$ to *each* customer $\omega \in [0, 1]$. Discrimination is said to be incomplete otherwise. Incomplete discrimination is thus equivalent to customer-class pricing.

Remark 1. It is common to call perfect discrimination the case in which the monopolist is able to extract all the consumer surplus (e.g., Phlips, 1988). If ω is a reservation price, that is, if each customer ω buys one unit as long as $P \leq \omega$ and nothing otherwise, the demand function is rectangular and complete discrimination is also perfect. This idealized situation has been called first-degree price discrimination by Pigou (1932) and is analyzed in Appendix B. When the demand function of each customer ω is not rectangular but is a decreasing function of the price, complete discrimination is not perfect anymore. This more realistic situation has been called third-degree price discrimination by Pigou (1932).

To use a more modern (and perhaps more meaningful) terminology (e.g., McAfee, 2008 or Stole, 2007), first and third-degree price discrimination are examples of direct price discrimination since each customer is assigned by the monopolist in a group as a function of its observable characteristic ω. When price discrimination is indirect (second-degree price discrimination to use the popular Pigovian terminology) customers are presented a menu of $n \geq 2$ options (e.g., n versions of the good differentiated by quality, n two-part tariffs) and customers choose the option which is best for them, i.e., customers "self-select" into different groups.

As is well-known, an important condition for price discrimination to be feasible is that arbitrage (i.e., parallel trade or gray market) must be difficult or even impossible. In a world without arbitrage, the monopolist can completely price discriminate and will charge the price $P^*(\omega) = \frac{\omega}{2}$ (i.e., that maximizes Eq. (2)) to each customer ω.

[1] Maskus (2000) offers a summary of the IPR exhaustion regimes for various regions (EU, USA). See also chapter 19 of Sugden (2009).

[2] Eq. (1) defines the "linear parallel demand" case, as opposed to the "linear rotating demand" considered in Malueg and Schwartz (1994) and more recently in Szymanski and Valleti (2006), in which $q(\omega, P) = 1 - \frac{1}{\omega}P$, where ω is uniformly distributed in $\Omega =]1 - x, 1 + x[$ and where $x \in]0, 1]$ is interpreted as the "dispersion" of the *continuum*.