Versioning Information Goods

Hal R. Varian*
University of California, Berkeley

March 13, 1997

1 Introduction

One prominent feature of information goods is that they have large fixed costs of *production,* and small variable costs of *reproduction.* Cost-based pricing makes little sense in this context; value-based pricing is much more appropriate. Different consumers may have radically different values for a particular information good, so techniques for differential pricing become very important.

There are many forms of differential pricing; a recent survey can be found in Varian [1989]. In this paper we will focus on a particular aspect of differential pricing known as *quality discrimination* or *versioning*. These terms describe situations in which the producer provides different qualities/versions of a good which sell at different prices.

As we will see below, the point of versioning is to get the consumers to sort themselves into different groups according to their willingness to pay. Consumers with high willingness to pay choose one version, while consumers with lower willingnesses to pay choose a different version. The producer chooses the versions so as to induce the consumers to "self select" into appropriate categories.

*Research support from NSF grant SBR-9320481 is gratefully acknowledged. Prepared for *Digital Information and Intellectual Property*, Harvard University, January 23-25, 1997. I would like to thank Fu-Chuan Lai and Char-Ming Yu for catching an error in an earlier version.

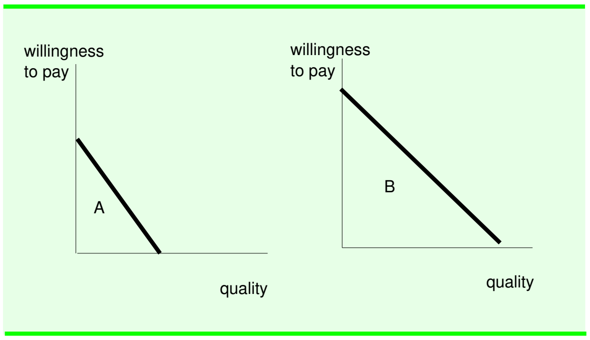

Figure 1: Pricing with observable characteristics

2 Observable characteristics

The fundamental problem in any form of differential pricing is to set prices so that purchasers who are able and willing to pay high prices do so. If willingness to pay is correlated with observable characteristics, such as membership in certain social or demographic groups, prices can be keyed to these observable characteristics. Senior citizen discounts, student discounts, AAA discounts, and so on, are examples of this sort of differential pricing.

The basic theory of this *third-degree price discrimination* is well known. Let's examine a very simple case. Suppose that there are two groups, one with high willingness to pay for "quality," the other with low willingness to pay. In this discussion, quality simply refers to some characteristic of the good that is desirable to consumers; in the case of information goods this could be resolution of a digital image, timelessness of financial news, or speed of operation of a software package.

We assume that the producer can perfectly discern the type of the consumer by observing some exogenous characteristic of the consumer: zip code, age, gender, etc.

Figure 1 depicts the demand curves for quality by the two types of consumers. Note that the demand curve in Figure 1A depicts a lower willingness to pay for the good than the demand curve in Figure 1B for each different quality at which the good may be offered.

For simplicity we will assume that the marginal cost of producing incremental quality is zero. It is easy to modify the analysis for any non-constant cost of providing quality.

The profit-maximizing policy for this simple problem is immediate. Since the producer, by assumption, can perfectly identify the type of the consumer, it will price the good so as to extract the entire consumer's surplus. But if the producer is able to extract the entire surplus, it will choose the quality so that total surplus is maximized. Using the notation in Figure 1, the firm would set the quality intended for type 1 to be x_1^0 and charge $r_1 = \text{area}A$; the quality intended for type 2 would be x_2^0 and the price would be $r_2 = \text{area}B$.

It is worthwhile noting that this pricing solution is Pareto efficient in that there is no way to make the consumers better off without making the producer worse off. In this simple case differential pricing has allowed the producer to capture the entire consumers' surplus.

3 Differential pricing infeasible

What happens in our simple example if price discrimination is not feasible? To be explicit, let us suppose that a fraction π of the population is the high willingness to pay type, and a fraction $1 - \pi$ is of the low willingness to pay type. In this case, the producer can set the price and quality so that only the high willingness to pay type buys the good or so that both types buy the good.

The profit from the former strategy is πr_2 and the profit from the latter strategy is $\pi r_1 + (1 - \pi) r_1 = r_1$. The producer will choose whichever strategy yields larger profit. Note that selling to both types is Pareto efficient, while restricting the quality to x_1 and selling only to the high willingness to pay market is Pareto inefficient—it would be possible to make the low-WTP type better off at zero cost. The seller refuses to do this since selling to that market would reduce the profit it makes on the high-WTP type.

4 Unobservable characteristics

We continue to investigate the simple example described above in the more interesting case in which the seller knows something about the *distribution* of willingness to pay in the population, but cannot identify the

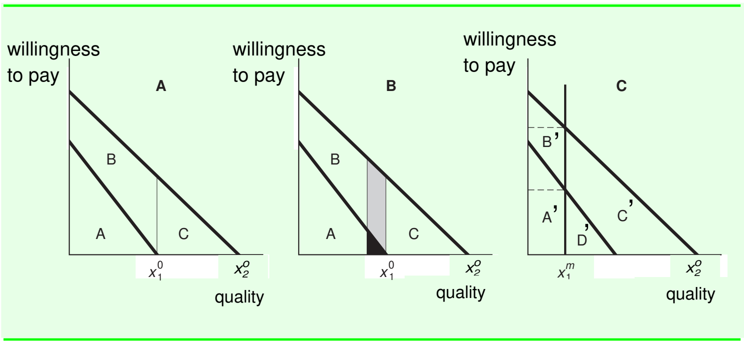

Figure 2: Self-selection problem

willingness to pay of a given consumer.[1]

In this case the seller cannot base its price on an *exogenous* observable characteristic such as membership in some group, but can base its price on an *endogenous* characteristic such as the quality of the choice the consumer purchases.

The appropriate strategy for the seller in this situation is to choose two qualities and associated prices and offer them to the consumers. Each of the different consumer types will one of the two quality/price pairs. The seller wants to choose the qualities and prices of the packages offered so as to maximize profit.

The intention is to get the consumers to *self-select* into the high- and low-WTP groups by setting price and quality appropriately. That is, the seller wants to choose price/quality packages so that the consumers with high WTP choose the high-price/high-quality package, and the consumers with low WTP choose the low-price/low-quality package.

Figure 2 depicts a possible strategy for this self-selection problem. In this figure we have superimposed the two demand curves. We start in panel A with the situation examined in the previous section: the seller produces qualities (x_1^0, x_2^0) and sells these at prices $r_1^0 = A$ and $r_2^0 = A + B + C$. respectively. By construction, each consumer gets zero surplus if he chooses the package intended for him.

[1]The basic insights in this section were first spelled out by Maskin and Riley [1984]; the geometric approach used here seems to have been first used by Varian [1992].

A careful inspection of Figure 2 shows us that the particular price/quality pairs indicated do not satisfy the self-selection constraints since the high WTP consumer can choose the package intended for the low WTP consumer and achieve a positive surplus. Specifically, if the high WTP consumer chooses the bundle (r_1^0, x_1^0) he will achieve a surplus represented by area B. The seller's profit would then be $r_1 = A$, just as if price differentiation were infeasible.

In order to prevent this arbitrage, the seller could set a price of $A + C$ for x_2^0. To see that this induces self selection, note that the high willingness to pay consumer would be indifferent to purchasing x_1^0 at a price of A, yielding a surplus of B, or purchasing x_2^0 at a price of $A + C$, yielding the same surplus. However, this pricing is definitely more profitable than the original pricing for the seller, since it yields profits of $\pi(A+C)+(1-\pi)A = A + \pi C > A$.

Using this strategy the producer is able to capture the *incremental* surplus to the high-WTP consumer represented by area C that is associated with the improving the quality from x_1^0 to x_2^0.

However, the story doesn't end here. Although this strategy is *feasible*, it is not profit-maximizing. It turns out that the seller can increase its profit even more by *reducing* the quality available to the low-WTP consumers.

This is illustrated in panel B of Figure 2. By reducing quality offered to the type 1 consumers by a small amount, the seller looses the profit denoted by the small black triangle. However, the quality reduction allows the seller to *increase* the price charged to the high-WTP consumers by the amount indicated by the grey trapezoid.

By making the low-quality bundle have even lower quality, the seller can charge more for the high-quality bundle. The seller will continue to reduce the quality of the low-quality bundle until the marginal reduction in revenues from the low-WTP consumers just equals the marginal increase in revenues from the high-WTP consumers. This is depicted in Figure 2 for the case of equal numbers of high- and low-WTP consumers ($\pi = 1/2$).

5 Examples of quality adjustment

There are many examples of this sort of quality adjustment to support self-selection. Airline pricing is an obvious example. Airlines sell discounted fares that involve various restrictions—advance purchase, Saturday night stayover, etc. Business travelers choose unrestricted fares, while tourists choose the restricted fares. Even though the cost of carrying a business traveler is essentially the same as carrying a tourist, unrestricted fares can sell for 3-5 times as much as restricted fares.

In this case reducing the "quality" of the good by imposing additional restrictions allows the producer to segment the market and induce self-selection so that prices can be based on willingness-to-pay.

Quality adjustment in transportation in order to induce self selection is a very old strategy. Consider, for example, this commentary on railroad pricing by Emile Dupuit, a nineteenth century French economist.

> It is not because of the few thousand francs which would have to be spent to put a roof over the third-class carriage or to upholster the third-class seats that some company or other has open carriages with wooden benches ... What the company is trying to do is prevent the passengers who can pay the second-class fare from traveling third class; it hits the poor, not because it wants to hurt them, but to frighten the rich ... And it is again for the same reason that the companies, having proved almost cruel to the third-class passengers and mean to the second-class ones, become lavish in dealing with first-class customers. Having refused the poor what is necessary, they give the rich what is superfluous. (Ekelund [1970])

Deneckere and McAfee [1996] contains several other examples of this sort. We describe only two: a hardware example and a service example.

- IBM LaserPrinter series E. This printer was identical to the standard LaserPrinter, but printed at 5 pages per minute rather than 10. The reason was the presence of a chip which induced wait states.

- Federal Express. Federal Express offers to deliver your package the next day, or the next day before 10 AM. They will typically make two deliveries rather than deliver standard packages before 10 AM.

If we look at (physical) information goods we can find several more examples.

- Books. Popular books are sold first in hardback; several months later they are reissued in paperback. The cost of paperback binding is only slightly less than hardback binding.

- Movies. Movies are issued first for the big screen; several months later they are reissued in video.

Other sellers of information goods and services are just beginning to implement this strategy. Here are some examples.

- PAWWS Financial Network charges $8.95 per month for a portfolio accounting system that measures stock values using 20-minute delayed quotes. Real-time quotes are available for $50 per month.

- PhotoDisc sells royalty-free stock photographs on the Web. Professional users want high-resolution images that can be printed in commercial journals; non-professionals want medium- or low-resolution images for newsletters. PhotoDisc sells different size images for different prices; at the time this chapter was written they sold 600K images (72 dots per inch resolution) for $19.95 and 10Mb images (300 dots per inch) for $49.95.

- Wolfram Research, Inc. sells a version of *Mathematica*, a computer program that does symbolic, graphical, and numerical mathematics. At one time the student version of *Mathematica* disabled the floating-point co-processor so that mathematical and graphical calculations were much slower. The student version sold for a bit over $100; the professional version sold for over $500.

- Windows NT Workstation 4.0 sells for about $260. It can be configured as a Web server, but only accepts 10 simultaneous connections. Windows NT Workstation Server can accept hundreds of simultaneous connections, and sells for $730–$1080, depending on configuration. According to an analysis by O'Reilly Software the two operating systems are essentially the same.

6 Welfare implications of versioning

Deliberate reduction of quality seems perverse to most casual observers. However, the social welfare of quality discrimination can only be judged

	Perfect PD	high-WTP	Both	Versioning
P_L	A	$A+B+C$	A	A'
P_H	$A+B+C$	$A+B+C$	A	$A'+C'+D'$
Surplus_L	0	0	0	0
Surplus_H	0	0	B	B'
Producer surplus	$A+\pi[B+C]$	$\pi[A+B+C]$	A	$A'+\pi[C'+D']$
Total surplus	$A+\pi[B+C]$	$\pi[A+B+C]$	$A+\pi B$	$A'+\pi[B'+C'+D']$

Table 1: Surplus comparisons

relative to its alternatives. In particular, if differential pricing were not allowed, the low end of the market might not be served at all.

Table 1 depicts the total surplus in four different pricing regimes: perfect price discrimination, a flat price at which only the high-WTP consumer purchases, a flat price at which both consumers purchase, and the optimal versioning solution.

Note that the low-WTP consumer always ends up with zero surplus in this model, and so can safely be ignored in the welfare calculations.

We are interested in the welfare effect of versioning. There are 3 natural benchmarks: perfect price discrimination and the two flat-price regimes. The surplus calculations are

$$\begin{aligned}
\text{perfect PD} - \text{versioning} &= (1-\pi)D' \\
&= \text{value of lost quality} \\
\text{versioning - sell to high-WTP only} &= (1-\pi)A' \\
&= \text{value to low-WTP consumers} \\
&\quad\ \text{of incremental consumption} \\
\text{versioning} - \text{sell to both} &= \pi C - (1-\pi)D' \\
&= \text{incremental consumption of} \\
&\quad\ \text{high-WTP consumers} \\
&\quad\ -\text{reduced quality to} \\
&\quad\ \text{low-WTP consumers}.
\end{aligned}$$

Note that social surplus always increases from versioning if only the high-WTP consumers would be served under flat pricing. If both classes of consumers would be served under flat pricing the welfare effect is ambiguous. However, the following observation gives us a clear condition under which welfare will be reduced by versioning.

Fact 1 *If it is profitable to serve both classes of consumers under flat pricing and*

$$\pi B - (1-\pi)A' > 0$$

then welfare will be reduced by versioning.

Proof. If it is profitable to serve both classes of consumers under flat pricing,

$$A > \pi[A + B + C],$$

which we can rewrite as

$$(1-\pi)A - \pi B - \pi C > 0.$$

Adding the inequality in the hypothesis gives us

$$(1-\pi)[A - A'] - \pi C > 0.$$

Substituting $D' = A - A'$ gives us

$$\pi C - (1-\pi)D' < 0.$$

The result now follows from the last surplus calculation.

Deneckere and McAfee [1996] analyze the welfare effects of "damaged goods" in a model with exogenous quality. They show that when there are two distinct markets (e.g., tourist travelers and business travelers) and only the high-WTP market would be served in the absence of price discrimination, versioning can easily result in a Pareto improvement: both class of customers and the the producer will be made better off. When there are many classes of customers with different product valuations, versioning may still result in a Pareto improvement, but the conditions required are much more stringent.

7 Policy

What should antitrust policy be towards differential pricing of this sort? I have examined the welfare effects of differential pricing elsewhere (Varian [1985, 1996]) so will limit myself to a simple summary.

The critical issue turns out to be whether differential pricing increases or decreases total output and/or quality. If the total output in a market decreases under differential pricing, then welfare (consumer plus producer surplus) definitely decreases. If total output in a market goes up, then differential pricing increases welfare.

Translated into the context of the examples given earlier, the question is whether versioning allows new markets to be served that would not be served in the absence of this pricing strategy. If the answer is "yes," then it is likely that versioning increases welfare; if the answer is "no," then versioning reduces aggregate welfare.

Unfortunately, existing antitrust law does not seem to have a very clear understanding of the issues surrounding these issues. As Huber [1993] put it:

> Almost every marketplace scheme in the information industry today could be construed as illegal under our antiquated antitrust laws. Information just doesn't obey the ordinary laws of economics, so the people who sell it can't obey ordinary antitrust laws. Judges had better get used to that. What we're talking about here is the future of our entire economy.

8 Goldilocks pricing

We have seen that the number of versions of a product offered should be equal to the number of types of consumers in the market. If there are business travelers and tourist travelers in the market, then it makes sense to have two broad "versions" of airline tickets.

But what happens if there is no obvious market segmentation? What if the number of "types" is huge? Are there any useful rules of thumb?

A common choice is to have two versions: a "standard" and an "enhanced" version. However, some recent work in marketing suggests that the optimal number of versions in this case is not two but three.[2]

The reason is due to what psychologists call *extremeness aversion:* if the only two sizes of drink that you offer are small and large, then some consumers will be on the margin between choosing one extreme or the other. *Some* of these consumers will choose the small version, thereby reducing producer revenues. But suppose the producer adds a "jumbo" version, and renamed the sizes "small," "medium," and "large," with the current medium being the same size as the previous large version. In this case, the medium size serves as a focal point for the indecisive: those who would have chosen small, end up compromising on medium, thereby increasing revenues.

Simonson and Tversky [1992] describe a marketing experiment in which two groups of consumers were asked to choose microwave ovens. One group was offered a choice between two ovens, an Emerson priced at \$109.99 and a Panasonic priced at \$179.99. The second group was offered these ovens plus a high-end Panasonic priced at \$199.99.

By offering the high-end oven, Panasonic increased its market share from 43% to 73%. More remarkably, the sales of the mid-priced Panasonic oven increased from 43% to 60% apparently *because* it was now the "compromise" choice. According to Smith and Nagle [1995],

> Adding a premium product to the product line may not necessarily result in overwhelming sales of the premium product itself. It does, however, enhance buyers' perceptions of lower-priced products in the product line and influences low-end buyers to trade up to higher-priced models.

If this same psychology applies to information goods as well as physical goods, then the producer may be better off providing standard, professional, and "gold" versions than it would be if it just provided standard and professional versions even if very few consumers choose the gold version.

9 Practical implications

What are the practical implications of this analysis for the producer of information goods?

[2] See Simonson and Tversky [1992] and Smith and Nagle [1995].

The first point is the most fundamental: *make sure that you design the product so that it can be versioned.* That is, the produce should be designed in a way that it is easy to reduce its quality in order to sell to a particular market segment. In addition to being sound software engineering, modularization can be very useful from the viewpoint of marketing since it makes it easy to remove features.

The second point is that the right way to to design the product will generally be to *design for the high end of the market first, and then downgrade the product to get the versions for the other segments of the market.* Recall that our analysis implied that the produce for the high-demand consumer was chosen so that the marginal willingness to pay for additional quality equals the marginal cost of producing additional quality (which was zero in our example.)

If we think of quality as being "additional features," an admittedly dangerous equivalence, this means that the producer should add features until the willingness to pay for an additional feature by the high end of the market just equals the cost of providing that feature. Once the high end has been determined, the producer then removes features to sell to the lower segment of the market, recognizing that it feature it removes allows it to increase the price sold to the high WTP consumers.

The third point applies to information goods that need to be "viewed" using a specific piece of hardware or software. For example, think of a browser used to view on-line newspaper articles. If the owner of the content controls the browser, then it can choose the features of the browser to enhance the quality of consuming the content. For example, if the consumer is viewing page 7 of the article, it is likely that page 8 will be the next piece of content he looks at, so the browser could download page 8 in the background. Controlling the browser allows the seller of content to increase the quality of what it is selling.

But as we've seen, it is also advantageous to *reduce* the quality of a good in order to segment the market. Hence controlling the browser—or more generally, the conditions under which the information good is consumed—can be very advantageous to the content owner since it makes versioning easier. This may well be one of the factors that makes use of Java-based viewers attractive.

Finally, the discussion of "Goldilocks pricing" suggests that in the absence of any additional information having three versions rather than two may be attractive due to "extremeness aversion" on the part of consumers.

References

Raymond J. Deneckere and R. Preston McAfee. Damaged goods. *Journal of Economics and Management Strategy*, 5(2):149–174, 1996.

R. B. Ekelund. Price discrimination and product differentiation in economic theory: An early analysis. *Quarterly Journal of Economics*, 84:268–78, 1970.

Peter Huber. Two cheers for price discrimination. *Forbes*, xx:142, 1993. Available at http://khht.com/huber/forbes/092793.html.

Eric Maskin and John Riley. Monopoly with incomplete information. *Rand Journal of Economics*, 15:171–96, 1984.

Itamar Simonson and Amos Tversky. Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research*, 29:281–295, 1992.

Gerald E. Smith and Thomas T. Nagle. Frames of reference and buyers' perception of price and value. *California Management Review*, 38(1):98–116, 1995.

Hal R. Varian. Price discrimination and social welfare. *American Economic Review*, 75:870–75, 1985.

Hal R. Varian. Price discrimination. In Richard Schmalensee and Robert Willig, editors, *Handbook of Industrial Organization*. North-Holland Press, Amersterdam, 1989.

Hal R. Varian. *Microeconomic Analysis*. W. W. Norton and Company, New York, 1992.

Hal R. Varian. Differential pricing and efficiency. *First Monday*, 1(2):xx–xx, 1996. Available at http://www.firstmonday.dk.