Cowles Foundation
for Research in Economics at Yale University

Cowles Foundation Discussion Paper No. 1896R

THE LIMITS OF PRICE DISCRIMINATION

Dirk Bergemann, Benjamin Brooks, and Stephen Morris

May 2013
Revised July 2013

The Limits of Price Discrimination*

Dirk Bergemann† Benjamin Brooks‡ Stephen Morris§

July 2, 2013

Abstract

We analyze the welfare consequences of a monopolist having additional information about consumers' tastes, beyond the prior distribution; the additional information can be used to charge different prices to different segments of the market, i.e., carry out "third degree price discrimination".

We show that the segmentation and pricing induced by the additional information can achieve *every* combination of consumer and producer surplus such that: (i) consumer surplus is non-negative, (ii) producer surplus is at least as high as profits under the uniform monopoly price, and (iii) total surplus does not exceed the efficient gains from trade.

As well as characterizing the welfare impact of price discrimination, we examine the limits of how prices and quantities can change under price discrimination. We also examine the limits of price discrimination in richer environments with quantity discrimination and limited ability to segment the market.

KEYWORDS: First Degree Price Discrimination, Second Degree Price Discrimination, Third Degree Price Discrimination, Private Information, Privacy, Bayes Correlated Equilibrium.

JEL CLASSIFICATION: C72, D82, D83.

*We gratefully acknowledge financial support from NSF SES 0851200 and ICES 1215808. We would like to thank Nemanja Antic, Simon Cowan, Babu Nahata, Omer Reingold, Juha Tolvanen, and Ricky Vohra for informative discussions. We have also benefitted from comments by seminar participants at NYU and CalTech. Lastly, thanks to Alex Smolin for excellent research assistance.

†Department of Economics, Yale University, New Haven, U.S.A., dirk.bergemann@yale.edu.

‡Department of Economics, Princeton University, Princeton, U.S.A., babrooks@princeton.edu.

§Department of Economics, Princeton University, Princeton, U.S.A., smorris@princeton.edu.

1 Introduction

A classic and central issue in the economic analysis of monopoly is the impact of discriminatory pricing on consumer and producer surplus. A monopolist engages in *third degree price discrimination* if he uses additional information about consumer characteristics to offer different prices to different segments of the aggregate market. A large and classical literature (reviewed below) examines the impact of particular segmentations on consumer and producer surplus, as well as on output and prices.

In this paper, we characterize what could happen to consumer and producer surplus for all possible segmentations of the market. We know that at least two points will be attained. If the monopolist has no information beyond the prior distribution of valuations, there will be no segmentation. The producer charges the uniform monopoly price and gets the associated monopoly profit, which is a lower bound on producer surplus; consumers receive a positive surplus, the standard information rent. This is marked by point A in Figure 1. On the other hand, if the monopolist has complete information about the valuation of the buyers, then he can completely segment the market according to true valuations. This results in perfect or *first degree price discrimination*. The resulting allocation is efficient, but consumer surplus is zero and the producer captures all of the gains from efficient trade. This is marked by point B in Figure 1.

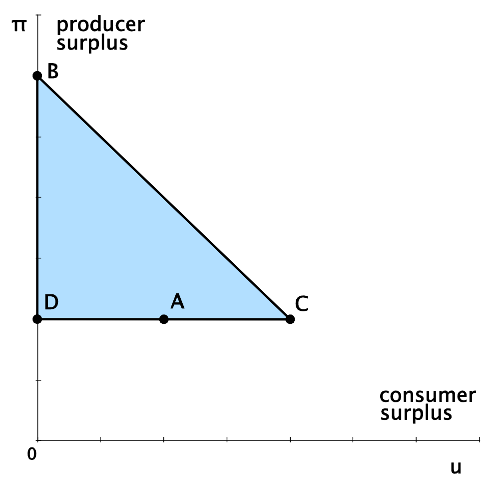

Figure 1: The Surplus Triangle of Price Discrimination

To begin with, we can identify some elementary bounds on consumer and producer surplus in any market segmentation. First, consumer surplus must be non-negative as a consequence of the participation constraint; a consumer will not buy the good at a price above his valuation. Second, the producer must get at least the surplus that he could get if there was no segmentation and he charged the uniform monopoly price. Third, the sum of consumer and producer surplus cannot exceed the total value that consumers receive from the good, when that value exceeds the marginal cost of production. The shaded right angled triangle in Figure 1 illustrates these three bounds.

Our main result is that *every* welfare outcome satisfying these constraints is attainable by some market segmentation. This is the entire shaded triangle in Figure 1. The point marked C is where consumer surplus is maximized; in particular, the producer is held down to his uniform monopoly profits, and consumers get the residual of the social surplus from an efficient allocation. At the point marked D, social surplus is minimized by holding producer surplus down to uniform monopoly profits and holding consumer surplus down to zero.

We can explain these results most easily in the case where there is a finite set of possible consumer valuations and the cost of production is zero. The latter is a normalization we will maintain throughout most of the paper. We will first explain one intuitive way to maximize consumer surplus, i.e., realize point C. The set of market prices will consist of every valuation less than or equal to the uniform monopoly price. Suppose that we can divide the market into segments corresponding to each of these prices in such a way that (i) in each segment, the consumers' valuations are always greater than or equal to the price for that segment; and (ii) in each segment, the producer is indifferent between charging the price for that segment and charging the uniform monopoly price. Then the producer is indifferent to charging the uniform monopoly price on *all* segments, so producer surplus must equal uniform monopoly profit. The allocation is also efficient, so consumers must obtain the rest of the efficient surplus. Thus, (i) and (ii) are sufficient conditions for a segmentation to maximize consumer surplus.

We now describe a way of constructing such a market segmentation iteratively. Start with a "lowest price segment" where a price equal to the lowest valuation will be charged. All consumers with the lowest valuation go into this segment. For each higher valuation, a share of consumers with that valuation also enters into the lowest price segment. While the *relative* share of each higher valuation (with respect to each other) is the same as in the prior distribution, the proportion of all of the higher valuations is lower than in the prior distribution. We can choose that proportion between zero and one such that the producer is indifferent between charging the segment price and the uniform monopoly price. We know this must be possible because if the proportion were equal to one, the uniform monopoly price would be profit maximizing for the producer (by definition); if the proportion were equal to zero—so only lowest valuation consumers were in the market—the lowest price would be profit maximizing; and, by keeping the relative proportions above the lowest valuation constant, there is no price other than these two that could be optimal. Now we have created one market segment satisfying properties (i) and (ii) above. But notice that the consumers not put in the lowest price segment are in the same relative proportions as they were in the original population. In particular, the original uniform monopoly price will be optimal on this "residual segment." We can apply the same procedure to construct a segment in which the market price is the second lowest valuation: put all the remaining consumers with the second lowest valuation into this market; for higher valuations, put a fixed proportion of remaining consumers into that segment; choose the proportion so that the producer is indifferent between charging the second highest valuation and the uniform monopoly price. This construction iterates until it reaches the uniform monopoly price at which point we have recovered the entire population and we have attained point C. An analogous construction—reported in the paper—shows how to attain point D.

We also have a deeper geometric proof of our main result. This argument establishes an even stronger result: Any point where the monopolist is held down to his uniform monopoly profits—including outcomes A, C, and D in Figure 1—can all be achieved *with the same segmentation*! In this segmentation, consumer surplus varies because the monopolist is indifferent between charging different prices. This argument gives a deeper insight into why our results are true. Consider the set of all markets where a given monopoly price is optimal. This set is convex, so any aggregate market with the given monopoly price can be decomposed as a weighted sum of markets which are extreme points of this set, which in turn defines a segmentation. These extremal markets must take a special form. In any extremal market, the monopolist will be indifferent to setting any price in the support of consumers' valuations. Thus, each subset of valuations that includes the given monopoly price generates an extreme point. If the monopolist charges the uniform monopoly price on each extreme segment, we get point A. If he charges the lowest value in the support, we get point C, and if he charges the highest value we get point D.

Thus, we are able to demonstrate that points B, C, and D can be attained. Every point in their convex hull, i.e., the shaded triangle in Figure 1, can also be attained simply by averaging the segmentations that work for each extreme point, and we have a complete characterization of all possible welfare outcomes.

While we focus on welfare implications, we can also completely characterize possible output levels and derive implications for prices. An upper bound on output is the efficient quantity, and this is realized by any segmentation along the efficient frontier. In particular, it is attained in any consumer surplus maximizing segmentation. In such segmentations, prices are always (weakly) below the uniform monopoly price. We also attain a lower bound on output. Note that the monopolist must receive at least his uniform monopoly profits, so this profit is a lower bound on social surplus. We say a segmentation is conditionally efficient if, conditional on the amount of output sold, the allocation of the good is socially efficient. Such segmentations minimize output for a given level of social surplus. In fact, we construct a social surplus minimizing segmentation that is conditionally efficient and therefore attains a lower bound on output. In this segmentation, prices are always (weakly) higher than the uniform monopoly price.

Using our result for discrete distributions, we are able to prove similar results for any market that has a well-behaved distribution of consumers' valuations. A convergence result establishes the existence of segmentations that attain points C and D for any Borel measurable distribution. When the distribution over values has a density, we can construct market segmentations analogous to those for discrete values. These segmentations involve a continuum of segments which are indexed by a suggested market price for each segment. Conditional on a given price, there is a mass point of consumers with valuation equal to the market price, with valuations above (for consumer surplus maximization) and below (for social surplus minimization) distributed according to densities. The densities are closed form solutions to differential equations.

We contribute to a large literature on third degree price discrimination, starting with Pigou (1920). This literature examines what happens to prices, quantity, consumer surplus, producer surplus and social welfare as the market is segmented. Pigou (1920) considered the case of two segments with linear demand, where both segments are served when there is a uniform price. In this special case, he showed that output does not change under price discrimination. Since different prices are charged in the two segments, this means that some high valuation consumers are replaced by low valuation consumers, and thus social welfare decreases. We can visualize the results of Pigou (1920) and other authors in Figure 1. Pigou (1920) showed that this particular segmentation resulted in a west-northwest move (i.e., move from point A to a point below the negative 45^0 line going through A). A literature since then has focused on identifying sufficient conditions on the shape of demand for social welfare to increase or decrease with price discrimination. A recent paper of Aguirre, Cowan, and Vickers (2010) unifies and extends this literature[1] and, in particular, identifies sufficient conditions for price discrimination to either increase or decrease social welfare (i.e., move above or below the negative 45° line through A). Restricting attention to market segments that have concave profit functions and an additional property

[1]Key intervening work includes Robinson (1933), Schmalensee (1981) and Varian (1985).

("increasing ratio condition") that they argue is commonly met, they show that welfare decreases if the direct demand in the higher priced market is at least as convex as that in the lower priced market; welfare is higher if prices are not too far apart and the inverse demand function in the lower priced market is locally more convex than that in the higher priced market. They note how their result ties in with an intuition of Robinson (1933): concave demand means that price changes have a small impact on quantity, while convex demand means that prices have a large impact on quantity. If the price rises in a market with concave demand and falls in a market with convex demand, the increase in output in the low-price market will outweigh the decrease in the high price market, and welfare will go up.

Our paper also gives sufficient conditions for different welfare impacts of segmentation. However, unlike most of the literature, we allow for segments with non-concave profit functions. Indeed, the segmentations giving rise to extreme points in welfare space (i.e., consumer surplus maximization at point C and social surplus minimization at point D) rely on non-concave profit functions. This ensures that the type of local conditions highlighted in the existing literature will not be relevant. Our non-local results suggest some very different intuitions. Of course, consumer surplus always increases if prices drop in all markets. We show that for any demand curves, low valuation consumers can be pooled with the right number of high valuation consumers to give the producer an incentive to offer prices below the monopoly price. Moreover, this incentive can be made arbitrarily weak, so that consumers capture the efficiency gain.

The literature also has results on the impact of segmentation on output and prices. On output, the focus is on identifying when an increase in output is necessary for an increase in welfare. Although we do not analyze the question in detail in this paper, a given output level is associated with many different levels of producer, consumer and social surplus. We do identify the highest and lowest possible output over all market segmentations. On prices, Nahata, Ostaszewski, and Sahoo (1990) offer examples with non-concave profit functions where third degree price discrimination may lead prices in all market segments to move in the same direction; it may be that all prices increase or all prices decrease. We show that one can create such segmentations *for any* demand curve. In other words, in constructing our critical market segmentations, we show that it is always possible to have all prices fall or all prices rise (with non-concave profit functions in the segments remaining a necessary condition, as shown by Nahata, Ostaszewski, and Sahoo (1990)).

If market segmentation is exogenous, one might argue that the segmentations that deliver extremal surpluses are special and might be seen as atypical. But to the extent that market segmentation is endogenous, our results can be used to offer predictions about what segmentations might arise. For example, consider an internet company with a large amount of data about the valuations of a large numbers of consumers. If the internet company sold this information to producers who would use it to price discriminate, they have an incentive to sell as much information as possible. But suppose that the internet company instead chose to release the information for free to producers in order to maximize consumer welfare (perhaps because of regulatory pressure or a longer term business model). Our results describe how such a consumer minded internet company would endogenously choose to segment the market. In particular, they would have an incentive to segment the market in such a way that profits were not concave.[2]

We also consider the extension of our results to two important environments. First, we ask what would happen if each consumer demands more than one unit of the good, so there is scope for second degree price discrimination in concert with market segmentation. Consumers vary in their marginal utility for quantity, and in each segment, the producer can screen using quantity-price bundles, as in Maskin and Riley (1984). We derive a closed form characterization of the set of attainable consumer and producer surplus pairs. Now, the earlier extreme results that efficient and zero consumer surplus segmentations holding producer surplus to the prior information profit exist no longer hold. But there continues to be a very large set of feasible welfare outcomes, and thus scope for market segmentation to be Pareto-improving or Pareto-worsening.

Second, we consider the case where there are exogenous limits on the kind of market segments that can be induced. This would be the case if the monopolist is limited to access information about particular consumer characteristics, and those characteristics are associated with characteristic-specific demand curves. The monopolist's information would induce segments that are convex combinations of the underlying demand curves. This gives rise to problems that are intermediate between our main results, where any segmentation is possible, and the classical price discrimination literature (reviewed above and summarized and extended by Aguirre, Cowan, and Vickers (2010)), where there is an exogenous division of the market. The literature compares outcomes under full discrimination and no discrimination, whereas we consider the range of outcomes possible under partial segmentation, where the monopolist imperfectly observes which division of the market he is facing. We give examples with intermediate results, in which the set of possible welfare outcomes is larger than in the classical

[2] A subtlety of this story, however, is that this could only be done by randomly allocating consumers with the same valuation to different segments with different prices. Thus consumers who knew their valuations would still have an incentive to misreport them to a benevolent intermediary, and thus they would still have an incentive (although perhaps a more subtle one) to conceal their valuations in anticipation of their later use in price discrimination, as in recent work of Taylor (2004) and Acquisti and Varian (2005).

literature but less permissive than our benchmark unrestricted model.

Our work has a methodological connection to two strands of literature. Kamenica and Gentzkow (2011)'s study of "Bayesian persuasion" considers how a sender would choose to transmit information to a receiver, if he could commit to an information revelation strategy before observing his private information. They provide a characterization of such optimal communication strategies as well as applications. If we let the receiver be the producer choosing prices, and let the sender be a planner maximizing some weighted sum of consumer and producer surplus, our problem belongs to the class of problems analyzed by Kamenica and Gentzkow (2011). They show that if one plots the utility of the "sender" as a function of the distribution of the sender's types, his highest attainable utility can be read off from the "concavification" of that function.[3] The concavification arguments are especially powerful in the case of two types. While we do not use concavification arguments in our main result at all, we use them directly in our two type analysis of second degree price discrimination and partial segmentation.

Bergemann and Morris (2013a) examine the general question, in strategic many-player settings, of what behavior could arise in an incomplete information game if players observe additional information not known to the analyst. They show that behavior that might arise is equivalent to an incomplete information version of correlated equilibrium termed "Bayes correlated equilibrium". Bergemann and Morris (2013a) explore the one-player version of Bayes correlated equilibrium, and its connection to the work of Kamenica and Gentzkow (2011) and others. In Bergemann and Morris (2013b), these insights were developed in detail in the context of linear-quadratic payoffs and normal distributed uncertainty. Using the language of Bergemann and Morris (2013a), the present paper considers the game of a producer making take-it-or-leave-it offers to consumers. Here, consumers have a dominant strategy to accept all offers strictly less than their valuation and reject all offers strictly greater than their valuation, and we select for equilibria in which consumers accept offers that make them indifferent. We characterize what could happen for any information structure that players might observe, as long as consumers know their own valuations. Thus, we identify possible payoffs of the producer and consumers in all Bayes correlated equilibria of the price setting game. Thus, our results are a striking application of the methodologies of Bergemann and Morris (2013a), (2013b) and Kamenica and Gentzkow (2011) to the general problem of price discrimination.

[3] Aumann and Maschler (1995), show that the concavification of the (stage) payoff function represents the limit payoff that an informed player can achieve in a repeated zero sum game with incomplete information. In particular, their Lemma 5.3, the "splitting lemma", derives a partial disclosure strategy on the basis of a concavified payoff function.

We present our main result in the case of discrete values in Section 2. We first give a characterization of the welfare set using the extremal segmentations described above. In addition to this abstract argument, we also provide a constructive approach that demonstrates the range of segmentations that can arise. The constructive arguments also allow us to characterize other consequences of discrimination, e.g., bounds on output. In Section 3, we extend our results to general settings with a continuum of values, so that there is a continuous demand curve. The discrete and continuum analyses are complementary: while they lead to the same substantive conclusions and economic insights, the arguments and mathematical formulations look very different, so we find it useful to report both cases independently. In Section 4, we analyze a version of the quantity discriminating monopolist with two types, and we then analyze price discrimination when there are exogenous limitations on how the market can be segmented. In Section 5, we conclude.

2 The Limits of Discrimination: The Discrete Case

A monopolist sells to a continuum of consumers, each of whom demands one unit of the good being sold. We normalize the constant marginal cost of the good to zero. In the current section, we assume that there are K possible values $v_k \in V \subseteq \mathbb{R}_+$ that the consumers might have with:

$$0 < v_1 < \cdots < v_k < \cdots < v_K.$$

A *market* is a vector $x = (x_1, ..., x_K)$ specifying the proportion of consumers with each of the K valuations. Thus market x corresponds to a step demand function, where $\sum_{j \geq k} x_j$ is the demand for the good at any price in the interval $(v_{k-1}, v_k]$ (with the convention that $v_0 = 0$). The set of possible markets X is the K-dimensional simplex,

$$X \triangleq \left\{ x \in \mathbb{R}_+^K \,\middle|\, \sum_{k=1}^{K} x_k = 1 \right\}.$$

We denote the *given aggregate market* by

$$x^* \in X. \tag{1}$$

We hold the aggregate market x^* fixed in the analysis and use stars to indicate properties of the aggregate market.

We say that price v_k is *optimal* for market x if the expected revenue from price v_k satisfies:

$$v_k \sum_{j \geq k} x_j \;\geq\; v_i \sum_{j \geq i} x_j \quad \text{for all } i = 1, ..., K. \tag{2}$$

We write X_k for the set of markets where price v_k is optimal,

$$X_k \triangleq \left\{ x \in X \,\middle|\, v_k \sum_{j \geq k} x_j \geq v_i \sum_{j \geq i} x_j \text{ for all } i = 1, ..., K \right\}.$$

Let $v^* \triangleq v_{i^*}$ be the optimal (i.e., revenue maximizing) uniform price v_{i^*} for the aggregate market x^*. For the entire analysis, it does not matter if there are multiple optimal uniform prices; any one will do. For notational convenience we shall assume that there is a unique optimal price, and hence that the inequality (2) is strict. Note that this implies that $x^* \in X^* \triangleq X_{i^*}$. The maximum feasible surplus is

$$w^* \triangleq \sum_{k=1}^{K} x_k^* v_k, \tag{3}$$

i.e., all consumers purchase the good (as they all value it above marginal cost). Uniform price producer surplus is then

$$\pi^* \triangleq \left(\sum_{k=i^*}^{K} x_k^* \right) v^* \;=\; \max_{i \in \{1,\ldots,K\}} \left(\sum_{j=i}^{K} x_j^* \right) v_i. \tag{4}$$

Uniform price consumer surplus is

$$u^* \triangleq \sum_{k=i^*}^{K} x_k^* \left(v_k - v^* \right).$$

We will also be interested in the lowest output $\underline{q}$ required to generate social surplus of at least the uniform price producer surplus π^*. This will come from selling to those with the highest valuations; thus in particular, there must be a critical valuation $v_{\underline{i}}$ such that the good is always sold to all consumers with valuations above $v_{\underline{i}}$ and never sold to consumers with valuations below $v_{\underline{i}}$. Thus letting $\underline{i}$ and $\underline{\beta} \in (0, 1]$ uniquely solve:

$$\underline{\beta} x_{\underline{i}}^* v_{\underline{i}} + \sum_{k=\underline{i}+1}^{K} x_k^* v_k = \pi^*, \tag{5}$$

we obtain a lower bound on output which is given by:

$$\underline{q} \triangleq \underline{\beta} x_{\underline{i}}^* + \sum_{k=\underline{i}+1}^{K} x_k^*. \tag{6}$$

The additional variable $\underline{\beta} \in (0, 1]$ allows us to randomize the decision to sell to the buyers at threshold value $v_{\underline{i}}$ and hence to achieve equality (5) in this finite setting.

2.1 A Simple Uniform Example

We will use a simple example to illustrate results in this section. Suppose that there are five possible valuations, $1, 2, 3, 4$ and 5, with equal proportions. Thus $K = 5$, $v_j = j$, and $x_j^* = \frac{1}{5}$ for each j. In this case, simple calculations show that feasible social surplus is $w^* = \frac{1}{5}(1 + 2 + 3 + 4 + 5) = 3$. The uniform monopoly price is $v^* = 3 = i^*$. The uniform monopoly profit is then $\pi^* = \frac{3}{5} \times 3 = \frac{9}{5}$, consumer surplus is $u^* = \frac{1}{5}(3 - 3) + \frac{1}{5}(4 - 3) + \frac{1}{5}(5 - 3) = \frac{3}{5}$, and deadweight loss is $3 - \frac{9}{5} - \frac{3}{5} = \frac{3}{5}$. The minimum output is $\underline{q} = \frac{2}{5}$. The consumer and producer surplus for this example is illustrated in Figure 2.

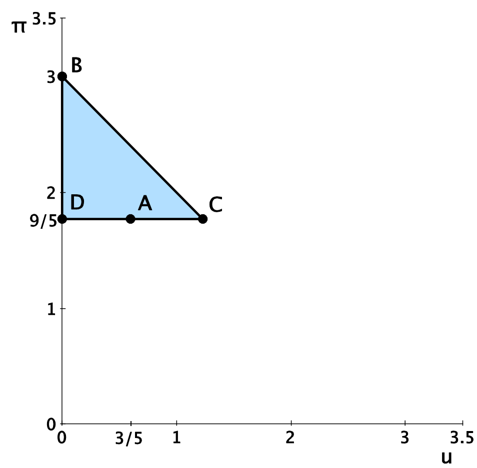

Figure 2: The Surplus Triangle of Price Discrimination: The Uniform Example

2.2 Segmentation and Pricing Strategy

A *segmentation* is a division of the social market into different markets. Thus, a segmentation σ is a simple probability distribution on X, with the interpretation that $\sigma(x)$ is the proportion of the population in market x. A segmentation can be viewed as a two stage lottery on outcomes $\{1, ..., K\}$ whose reduced lottery is x^*. Writing **supp** for the support of a distribution, the set of possible segmentations is given by

$$\left\{ \sigma \in \Delta(X) \left| \sum_{x \in \mathbf{supp}(\sigma)} \sigma(x) \cdot x = x^*, \;\; |\mathbf{supp}(\sigma)| < \infty \right. \right\}.$$

We restrict attention to finitely many segments, hence $|\mathbf{supp}(\sigma)| < \infty$, which is without loss of generality in the present environment with finitely many valuations. A *pricing strategy* for a segmentation

σ specifies a price in each market in the support of σ,

$$\phi : \mathbf{supp}\,(\sigma) \rightarrow \Delta\,\{v_1, ..., v_K\}\,, \tag{7}$$

which gives a distribution over prices for every market. A pricing strategy is *optimal* if, for each x, $v_k \in \mathbf{supp}(\phi(x))$ implies $x \in X_k$, i.e. all prices charged with positive probability must maximize profit on market x. If a pricing rule puts probability 1 on price v, we will simply write $\phi(x) = v$, and otherwise $\phi_k(x)$ is the probability of charging price v_k in market x. A segmentation σ and pricing strategy ϕ together determine the outcomes that we care about, namely the joint distribution of prices and consumers' valuations. An example of a segmentation and an associated optimal pricing rule is given by the case of perfect price discrimination. In this case the pricing strategy is deterministic in every segment, and we have five market segments with five associated prices as illustrated in the table below:

	value 1	value 2	value 3	value 4	value 5	price	weight
market 1	1	0	0	0	0	1	$\frac{1}{5}$
market 2	0	1	0	0	0	2	$\frac{1}{5}$
market 3	0	0	1	0	0	3	$\frac{1}{5}$
market 4	0	0	0	1	0	4	$\frac{1}{5}$
market 5	0	0	0	0	1	5	$\frac{1}{5}$
total	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$		

More generally, the consumer surplus with a segmentation σ and a pricing rule ϕ is given by:

$$\sum_{x \in \mathbf{supp}(\sigma)} \sigma\,(x) \sum_{k=1}^{K} \sum_{j=k}^{K} \phi_k\,(x)\, x_j\,(v_j - v_k)\,;$$

the producer surplus is

$$\sum_{x \in \mathbf{supp}(\sigma)} \sigma\,(x) \sum_{k=1}^{K} \sum_{j=k}^{K} \phi_k\,(x)\, x_j v_k;$$

and the output is

$$\sum_{x \in \mathbf{supp}(\sigma)} \sigma\,(x) \sum_{k=1}^{K} \sum_{j=k}^{K} \phi_k\,(x)\, x_j.$$

Our first result is a linear algebraic characterization of the set X_k of markets where price v_k is optimal. Write V_k for the set of non-empty subsets of $\{v_1, ..., v_K\}$ containing v_k. We will write $V^* \triangleq V_{i^*}$. A subset $S \in V_k$ represents a finite (support) set of valuations including v_k. Now for every support set $S \in V_k$, we define a market x^S : $x^S = \left(...., x_i^S, ...\right) \in X$,with the properties that (i) no consumer has a valuation outside the set S; and (ii) the monopolist is indifferent between charging any price in S. Thus, for every support set $S \in V_k$, we define the market x^S by the indifference conditions that for all $v_i \in S \in V_k$:

$$v_j \sum_{j \geq i} x_j^S = v_k \sum_{j \geq k} x_k^S, \tag{8}$$

and the inclusiveness condition:

$$\sum_{\{j | v_j \in S\}} x_j^S = 1. \tag{9}$$

Now, for every S, there exists a unique solution to the above condition (8) and (9), which can be described explicitly in terms of the element of set S. For any $S \in V_k$ write $\min S$ and $\max S$ for the smallest and largest element of S and, for each element $v_i \in S$ different from $\max S$, write $\mu(v_i, S)$ for the smallest element of S which is greater than v_i. For every $S \in V_k$, the uniquely defined market x^S given by:

$$x_i^S \triangleq \begin{cases} 0, & \text{if} \quad v_i \notin S; \\ \min S \left(\frac{1}{v_i} - \frac{1}{\mu(v_i, S)} \right), & \text{if} \quad v_i \neq \max S; \\ \frac{\min S}{\max S}, & \text{if} \quad v_i = \max S. \end{cases} \tag{10}$$

There are a finite set of such markets for every V_k. We next show that all markets $x \in X_k$ in which v_k is an optimal price are convex combinations of these extreme points x^S.

Lemma 1 (Extremal Segmentation)
X_k is the convex hull of $\left(x^S\right)_{S \in V_k}$

Proof. X_k is a finite-dimensional compact and convex set, so by the Krein-Milman theorem it is equal to the convex hull of its extreme points. We will show that every extreme point of X_k is equal to x^S for some $S \in V_k$. First observe that if v_i is an optimal price for market x, then $x_i > 0$. Otherwise the monopolist would want to deviate to a higher price if $\sum_{j=i+1}^{K} x_j > 0$ or a lower price if this quantity is zero, either of which contradicts the optimality of v_i.

Now, the set X_k is characterized by the linear constraints that for any $x \in X_k$:

$$\sum_{i=1}^{K} x_i = 1,$$

the nonnegativity constraints

$$x_i \geq 0, \text{ for all } i,$$

and the optimality (of price v_k) constraint:

$$\left(\sum_{i=j}^{K} x_i\right) v_j \leq \left(\sum_{i=k}^{K} x_i\right) v_k \text{ for } i \neq k.$$

Any extreme point of X_k must lie at the intersection of exactly K of these constraints. One active constraint is always $\sum_{i=1}^{K} x_i = 1$, and since v_k is an optimal price, the non-negativity constraint $x_k \geq 0$ is never active. Thus, there are exactly $K - 1$ active non-negativity and pricing constraints for $i \neq k$.

But as we have argued, we cannot have both the optimality and non-negativity constraints bind for a given i, so for each $i \neq k$ precisely one of the non-negativity and optimality constraints is binding. This profile of constraints defines x^S, where S is the set valuations which are optimal prices for the seller. ■

Thus for the given aggregate market $x^* \in X^*$ there are segmentations of x^* which have support on the markets x^S for $S \in V^*$ as defined above in (10) only. We refer to any market x^S as an *extremal market*, and to any segmentation consisting only of extremal markets as an *extremal segmentation*. In general, there will be many such segmentations. Our main result using extremal segmentations does not depend on which one we choose. In our uniform example, the segmentation of the uniform market x^* described in the table below represents one such extremal segmentation. This segmentation is the solution to particular segmentation algorithm, a "greedy" like algorithm formally described in the next subsection.

	value 1	value 2	value 3	value 4	value 5	weight
market $\{1,2,3,4,5\}$	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{20}$	$\frac{1}{5}$	$\frac{2}{5}$
market $\{2,3,4,5\}$	0	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{10}$	$\frac{2}{5}$	$\frac{3}{10}$
market $\{2,3,4\}$	0	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{2}$	0	$\frac{1}{10}$
market $\{3,4\}$	0	0	$\frac{1}{4}$	$\frac{3}{4}$	0	$\frac{2}{15}$
market $\{3\}$	0	0	1	0	0	$\frac{1}{15}$
total	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	

. (11)

2.3 Limits of Price Discrimination on Welfare

For a given market x, we define the minimum pricing rule $\underline{\phi}(x)$ to deterministically charge $\min(\mathbf{supp}(x))$ and, similarly, we define the maximum pricing rule $\overline{\phi}(x)$ to deterministically charge $\max(\mathbf{supp}(x))$. We observe that the minimum pricing rule always implies an efficient allocation in the market x and the maximum pricing rule implies an allocation in the market x where there is zero consumer surplus.

Theorem 1 (Minimum and Maximum Pricing)

In every extremal segmentation, minimum and maximum pricing strategies are optimal; producer surplus is π^* under every optimal pricing strategy; consumer surplus is zero under the maximum pricing strategy and consumer surplus is $w^* - \pi^*$ under the minimal pricing strategy.

Proof. By construction of the extremal markets, any price in S is an optimal price in market x^S. This implies that minimum and maximum pricing rules are both optimal. Since always setting the price equal to v^* is optimal, producer surplus must be exactly π^* in any extremal segmentation. Consumer surplus is always zero under the maximum pricing strategy. Since the minimal pricing rule always gives social surplus w^* and producer surplus is π^*, consumer surplus must be the difference $w^* - \pi^*$. ■

The above result only refers to aggregate consumer surplus over all valuations. But in fact, the minimum and maximum pricing strategies under every extremal segmentation allow the same predictions to hold pointwise, i.e. for every valuation of the consumer. That is, in the minimum pricing strategy, the expected net utility for every valuation type of the buyer is (weakly) larger than with uniform pricing in the aggregate market. Conversely, in the maximum pricing strategy, the expected net utility for every valuation type of the buyer is (weakly) smaller than with uniform pricing in the aggregate market. With the maximum pricing rule $\overline{\phi}(x)$, this follows directly from the construction of the maximum pricing rule. After all, only the buyer with the highest value in the segment x purchases the product under the maximum pricing rule but has to pay exactly his valuation. Hence, the expected net utility conditional on a purchase is zero, but so is the expected net utility without a purchase. All valuations are weakly worse off relative to the uniform price in the aggregate market. There, every buyer with a valuation $v_i > v^*$ received a strictly positive information rent. As for the minimum pricing rule $\underline{\phi}(x)$, first we observe that all efficient trades are realized as opposed to only those with a value equal or above the uniform price $v_i \geq v^*$; second by construction of the minimum pricing rule $\underline{\phi}(x)$, all sales are realized at prices below or equal to v^*. So we have:

Corollary 1 (Pointwise Consumer Surplus)

In every extremal segmentation, for every valuation v_i, the expected net utility is (weakly) larger in the minimum pricing strategy; and (weakly) smaller in the maximum pricing strategy than under the uniform price in the aggregate market.

Now, if we consider segmentations different from the extremal segmentation, then it still remains true that for any segmentation and optimal pricing rule, producer surplus must be at least π^*, consumer surplus must be at least zero, and the sum of producer surplus and consumer surplus must be at most w^*. And the set of attainable producer surplus and consumer surplus pairs must be convex. So we have:

Corollary 2 (Surplus Triangle)
For every (π, u) satisfying $\pi \geq \pi^*$, $u \geq 0$ and $\pi + u \leq w^*$, there exists a segmentation and an optimal pricing rule with producer surplus π and consumer surplus u.

There is a large multiplicity of segmentations and pricing rules that attain the maximal consumer surplus and minimal social surplus. We now provide a construction for a canonical "greedy" extremal segmentation. First put as much mass as we can on the market $x^{\mathrm{supp}(x^*)}$, i.e., the extremal market in which the monopolist is indifferent to charging all prices in the support of x^*. At some point, we will run out of mass for some valuation in $\mathbf{supp}\,(x^*)$. We then proceed with a new segment that puts as much mass as possible on the extremal market corresponding to all remaining valuations; and so on. More formally, we can describe the greedy algorithm as follows. Let F be the distribution function of the aggregate market with support V. We shall construct a sequence of sets, $S^0, ..., S^G \in V^*$ with $G \leq K-1$, which are initialized at $S^0 = V$ and satisfy strict set inclusion: $S^{g+1} \subsetneq S^g$. Suppose we "run" the greedy algorithm from time 0 to time 1. Write $H(v,t)$ for the cumulative probability mass left at time t under the greedy algorithm. Thus $H(v,t)$ is weakly increasing in v for all t and we set

$$\begin{aligned} H(v,0) &= F(v), \quad \text{for all } v \in V; \\ H(v_K,t) &= 1-t, \quad \text{for all } t \in [0,1]. \end{aligned} \tag{12}$$

We write $S(t)$ for the subset of values $v \in V$ where probability mass remains at time t. Thus

$$S(t) \triangleq \{v_k \in V | \, H(v_k,t) - H(v_{k-1},t)\} > 0. \tag{13}$$

By extension, we define for every support set $S(t)$, the distribution function $F^{S(t)}(v)$ associated with the probability distribution of the extremal market $x^{S(t)}$ as defined earlier by (10). And now let

$$\frac{dH(v,t)}{dt} \triangleq -F^{S(t)}(v). \tag{14}$$

Now, clearly, if we start at $t=0$, then $S(0) = S^0 = V$. By construction of the greedy algorithm $H(v,t)$, there must exist a first time $\tau_1 \leq 1$, where $S^0 = S(t) \neq S(\tau_1)$ for all $0 \leq t < \tau_1$. We set $S^1 \triangleq S(\tau_1)$, and by Lemma 1, $S^1 \in S^*$. Now, clearly, $S^1 \subsetneq S^0$, and we continue to eat into the distribution $H(v,t)$, but now removing probability only on the smaller support set S^1. Clearly, there are at most $K-1$ reductions of the support set S^g until we reach a singleton set. At each step g of the induction, we can appeal to Lemma 1 to observe that the remaining support S^g has the inclusion property $S^g \in V^*$. But since $F^{S(t)}(v_K) = 1$ for all t, it follows that at time $t=1$, there is zero residual probability left, and hence we have achieved a complete segmentation of the aggregate market.

The greedy algorithm uses the insight of Lemma 1 by constructing a sequence of segments, such that along the sequence, the number of active pricing constraints is strictly decreasing, and the number of active nonnegativity constraints is strictly increasing. More precisely, each segment g has a distinct number of non-negativity constraints active, namely at least g, and conversely each segment g, has a distinct number of pricing constraints active, namely at most $(K-1)-g$. Generically, at each stopping time only a single non-negativity constraint switches from being inactive to active, and then the above statement involving "at most" are exact statements. In our uniform example, the greedy algorithm gives rise to the segmentation (11) displayed above. If we apply either the minimum pricing rule $\underline{\phi}(x)$ or the maximum pricing rule $\overline{\phi}(x)$ to the above segmentation, then we get the following prices, displayed in the second to last and last column, respectively:

	value 1	value 2	value 3	value 4	value 5	weight	$\underline{\phi}(x)$	$\overline{\phi}(x)$
market $\{1,2,3,4,5\}$	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{20}$	$\frac{1}{5}$	$\frac{2}{5}$	1	5
market $\{2,3,4,5\}$	0	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{10}$	$\frac{2}{5}$	$\frac{3}{10}$	2	5
market $\{2,3,4\}$	0	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{2}$	0	$\frac{1}{10}$	2	4
market $\{3,4\}$	0	0	$\frac{1}{4}$	$\frac{3}{4}$	0	$\frac{2}{15}$	3	4
market $\{3\}$	0	0	1	0	0	$\frac{1}{15}$	3	3
total	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	1		

2.4 Direct Segmentations

We observe that in the greedy segmentation of the uniform examples above, there are multiple segments in which the same price is charged under either the minimum or the maximum pricing rule. In fact, if the price v_k is optimal in markets x and x', then v_k is optimal in the merged market as well:

$$\frac{\sigma(x)}{\sigma(x)+\sigma(x')}x + \frac{\sigma(x')}{\sigma(x)+\sigma(x')}x'$$

Thus, we could merge markets so that a given price is charged in only one segment. More generally, a *direct segmentation* has at most K segments, one for each possible price v_k, where $x^k \in X_k$ and

$\sum_k \sigma(x^k)x^k = x^*$. The *direct pricing strategy* is the identity mapping, i.e. $\phi_k(x^k) = 1$. (In contrast to the extremal markets where the upper case superscript S in x^S referred to the support, here the lower case superscript k in x^k refers to price v_k charged in the direct segment x^k.) It should be clear that the direct pricing strategy is optimal for direct segmentations constructed in this way, and whenever we refer to a direct segmentation in the subsequent discussion, it is assumed that the monopolist will use direct pricing.

Extremal segmentations and direct segmentations are both rich enough classes to achieve any equilibrium outcome, where again an outcome is a joint distribution of prices and valuations. In particular, if a segmentation and optimal pricing rule (σ, ϕ) induce a given outcome, then there is both an extremal segmentation and optimal pricing strategy (σ', ϕ') and a direct segmentation σ'' (and associated direct pricing strategy ϕ'') that achieve the same outcome. To find an extremal segmentation, each market $x \in \mathbf{supp}(\sigma)$ can itself be decomposed using extremal markets with a segmentation σ_x, using only those indifference sets S which contain $\mathbf{supp}(\phi(x))$. The extremal segmentation of (σ, ϕ) is then defined by:

$$\sigma'(x^S) \triangleq \sum_{x \in \mathbf{supp}(\sigma)} \sigma(x)\sigma_x(x^S),$$

and the corresponding pricing rule is

$$\phi'_k(x^S) \triangleq \frac{1}{\sigma'(x^S)} \sum_{x \in \mathbf{supp}(\sigma)} \sigma(x)\sigma_x(x^S)\phi_k(x).$$

Similarly, the direct segmentation x^k can be defined by

$$\sigma''(x^k) \triangleq \sum_{x \in \mathbf{supp}(\sigma)} \sigma(x)\phi_k(x),$$

and the composition of each direct segment x^k is given by

$$x^k \triangleq \frac{1}{\sigma''(x^k)} \sum_{x \in \mathbf{supp}(\sigma)} \sigma(x)\phi_k(x) \cdot x.$$

In the uniform example, the direct segmentation corresponding to the consumer surplus maximizing greedy extremal segmentation is:

	value 1	value 2	value 3	value 4	value 5	price	weight
market 1	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{20}$	$\frac{1}{5}$	1	$\frac{2}{5}$
market 2	0	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{3}{10}$	2	$\frac{2}{5}$
market 3	0	0	$\frac{1}{2}$	$\frac{1}{2}$	0	3	$\frac{1}{5}$
total	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$		1

,

where the markets for prices 4 and 5 are degenerate. The direct segmentation corresponding to the social surplus minimizing greedy extremal segmentation is:

	value 1	value 2	value 3	value 4	value 5	price	weight
market 3	0	0	1	0	0	3	$\frac{1}{15}$
market 4	0	$\frac{1}{7}$	$\frac{3}{14}$	$\frac{9}{14}$	0	4	$\frac{7}{30}$
market 5	$\frac{2}{7}$	$\frac{5}{21}$	$\frac{5}{42}$	$\frac{1}{14}$	$\frac{2}{7}$	5	$\frac{7}{10}$
total	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$		1

, (15)

where the markets for prices 1 and 2 are degenerate.

Direct segmentations are well suited to the exploration of some of the alternative segmentations that attain the welfare bounds. Let us give a formal description of the first segmentation described in the introduction attaining maximum consumer surplus. For each $k \leq i^*$, let market x^k have the features that (i) the lowest valuation in the support is v_k; (ii) all values of v_{k+1} and above appear in the same relative proportion as in the aggregate population:

$$x_i^k \triangleq \begin{cases} 0, & \text{if } \; i < k; \\ 1 - \gamma_k \sum\limits_{i \geq k+1} x_i^*, & \text{if } \; i = k; \\ \gamma_k x_i^*, & \text{if } \; i > k; \end{cases} \tag{16}$$

where $\gamma_k \in [0,1]$ uniquely solves

$$\left(x_k^* + \gamma_k \left(\sum_{i=k+1}^{K} x_i^* \right) \right) v_k = \gamma_k \left(\sum_{i=i^*}^{K} x_i^* \right) v^*.$$

By construction of the above equality, both v_k and v^* are optimal prices for segment x^k. We can always construct a segmentation of the aggregate market x^* that uses only $\left(x^k\right)_{k=1}^{i^*}$. We establish the construction inductively, letting

$$\sigma\left(x^1\right) \triangleq \frac{x_1^*}{x_1^1} \tag{17}$$

and

$$\sigma\left(x^k\right) \triangleq \frac{x_k^* - \sum\limits_{i<k} \sigma\left(x^i\right) x_k^i}{x_k^k}. \tag{18}$$

We can verify that this segmentation generates maximum consumer surplus by charging in segment x^k the price v_k. The direct pricing rule is optimal and gives rise to an efficient allocation. Because the monopolist is always indifferent to charging v^*, producer surplus is π^*.

Proposition 1 (Consumer Surplus Maximizing Direct Segmentation)
There exists a direct segmentation and an optimal pricing rule where producer surplus is π^*, consumer surplus is $w^* - \pi^*$ and the output is socially efficient.

In the uniform example, this construction gives rise to the following segmentation:

	value 1	value 2	value 3	value 4	value 5	price	weight
market 1	$\frac{5}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	1	$\frac{9}{25}$
market 2	0	$\frac{1}{3}$	$\frac{2}{9}$	$\frac{2}{9}$	$\frac{2}{9}$	2	$\frac{12}{25}$
market 3	0	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	3	$\frac{4}{25}$
total	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$		1

.

2.5 Limits of Price Discrimination on Output

While our focus has been on welfare outcomes, we can also report tight results about output. The consumer surplus maximizing segmentations are efficient, and therefore maximize output among all segmentations and optimal pricing rules. To minimize output, we hold social surplus down to π^* while also ensuring that the allocation is conditionally efficient, so that the object is always sold to those who value the object the most. Note that our earlier segmentation (15) attaining minimum social surplus had some consumers with valuation 3 facing price 3 and thus buying the good but also had some consumers with valuation 4 facing price 5, and thus not buying the good. The total proportion of consumers buying the good is $\frac{5}{12}$. But we noted earlier that we attain the minimum producer surplus $\pi^* = \frac{9}{5}$ by selling only to those with valuations 4 and 5 which implies total output $\frac{2}{5} < \frac{5}{12}$. Below is a segmentation and optimal pricing rule in the example which attains minimum social surplus while only selling to those with valuations 4 and 5:

	value 1	value 2	value 3	value 4	value 5	price	weight
market 4	$\frac{8}{45}$	$\frac{8}{45}$	$\frac{2}{15}$	$\frac{2}{5}$	0	4	$\frac{1}{2}$
market 5	$\frac{2}{9}$	$\frac{2}{9}$	$\frac{4}{15}$	0	$\frac{2}{5}$	5	$\frac{1}{2}$
total	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$		1

.

More generally, we can always construct a direct segmentation using an inductive, but more subtle algorithm as in (17)-(18) that attains the output lower bound defined by condition (6). In the resulting conditionally efficient direct segmentation there will be zero consumer surplus and producer surplus π^*. In fact, consumers with valuations strictly above $\underline{i}$ purchase the good at a price equal to their valuation, and consumers with valuation $\underline{i}$ must purchase the good with probability $\underline{\beta}$, paying their valuation if they purchase the good. Consumers with valuations below $\underline{i}$ will not purchase the good. For the consumer surplus maximizing segmentation, we started by defining the segment with the lowest price and worked our way up through prices. To minimize output and social surplus, we adopt a different construction that starts by placing the highest value consumer in a segment with price equal to his own value. We then work our way down through the values. Consumers with a given value will be apportioned out to all of the segments with weakly higher prices.

Proposition 2 (Quantity Minimizing Direct Segmentation)
There exists a segmentation and optimal pricing rule where producer surplus is π^*, consumer surplus is 0 and output is $\underline{q}$.

Proof. We construct a direct segmentation that achieves the minimum output. We defined $\underline{i} \in \{1, ..., K\}$ and $\underline{\beta} \in (0, 1]$ as the unique solution of (5), identifying the conditionally efficient and quantity minimizing allocation that attains the uniform monopoly profit. We now define a particular conditionally efficient segmentation. We denote by y_i^k the probability mass of valuation v_i consumers in segment k (which will be charged price v_k). Because of our "top-down" construction, it will be easier to work directly with the object y^i, since we will not know how large each market should be until the induction terminates. Once the y^k are defined, we can easily recover the market sizes and weights by $\sigma\left(x^k\right) \triangleq \sum_{i=1}^K y_i^k$, and $x_i^k \triangleq y_i^k / \sigma\left(x^k\right)$. For $k > \underline{i}$, let

$$y_i^k \triangleq \begin{cases} 0, & \text{if} \quad i < k; \\ x_i^*, & \text{if} \quad i = k; \\ 0, & \text{if} \quad i > k. \end{cases} \tag{19}$$

For $k = \underline{i}$, let

$$y_i^{\underline{i}} \triangleq \begin{cases} 0, & \text{if} \quad i < \underline{i}; \\ \underline{\beta} x_{\underline{i}}^*, & \text{if} \quad i = \underline{i}; \\ \frac{y_i^i \left(v_i - v_{\underline{i}}\right)}{\sum\limits_{l=\underline{i}+1}^{K} y_l^l \left(v_l - v_{\underline{i}}\right)} (1 - \underline{\beta}) x_{\underline{i}}^* & \text{if} \quad i > \underline{i}; \end{cases} \tag{20}$$

and iteratively define for $k = \underline{i} - 1, \underline{i} - 2, ..., 1$:

$$y_i^k \triangleq \begin{cases} 0, & \text{if} \quad i < \underline{i}; \\ \frac{v_k y_k^k - v_i \sum\limits_{j=i+1}^{K} y_j^k}{\sum\limits_{l=\underline{i}}^{K} \left(v_l y_l^l - v_i \sum\limits_{j=i+1}^{K} y_i^l \right)} x_i^* & \text{if} \quad i \geq \underline{i}. \end{cases} \tag{21}$$

The segmentation, defined by (19)-(21) satisfies feasibility by construction. To wit, the ratios appearing in (20) and (21) are strictly positive and define shares that sum up to one. This will follow inductively from incentive compatibility, as

$$v_k y_k^k - v_i \sum_{j=i+1}^{K} y_i^k > v_k y_k^k - v_{i+1} \sum_{j=i+1}^{K} y_i^k.$$

The right-hand side is non-negative if incentive compatibility is satisfied. Now it remains to verify incentive compatibility. The non-trivial conditions we need to check are that profits cannot be increased by deviating from price v_k for $k \geq \underline{i}$ to a lower price. Consider the case where $k > \underline{i}$. First, observe that for each $i = 1, ..., \underline{i} - 1$:

$$\begin{aligned}
\sum_{l=\underline{i}}^{K} \left(v_l y_l^l - v_i \sum_{j=i+1}^{K} y_j^l \right) &= \sum_{l=\underline{i}}^{K} v_l y_l^l - v_i \sum_{j=i+1}^{K} x_j^*, \text{ by feasibility} \\
&= v^* \sum_{j=i^*}^{K} x_j^* - v_i \sum_{j=i+1}^{K} x_j^*, \text{ by construction of } \underline{i} \\
&> v_i \sum_{j=i}^{K} x_j^* - v_i \sum_{j=i+1}^{K} x_j^*, \text{ by definition of } i^* \\
&= v_i x_i^*.
\end{aligned}$$

Note that the inequality is strict, since v^* is the unique uniform monopoly price. (We could have alternatively made it the highest monopoly price, if there are multiple.) We then have from (21) that

$$y_i^k = \frac{v_k y_k^k - v_i \sum_{j=i+1}^{K} y_j^k}{\sum_{l=\underline{i}}^{K} \left(v_l y_l^l - v_i \sum_{j=i+1}^{K} y_j^l \right)} x_i^*.$$

It now follows from the above inequality, which reads as:

$$\sum_{l=\underline{i}}^{K} \left(v_l y_l^l - v_i \sum_{j=i+1}^{K} y_j^k \right) > v_i x_i^*, \tag{22}$$

that after replacing the rhs by the lhs of (22) that, after cancelling terms that:

$$y_i^k < \frac{v_k y_k^k}{v_i} - \sum_{j=i+1}^{K} y_j^k,$$

for all $k > \underline{i} \geq i$. Re-arranging this expression, we have

$$v_i \sum_{j=i}^{K} y_j^k < v_k y_k^k,$$

verifying incentive compatibility for $k > \underline{i} > i$. The same argument goes through with $k = \underline{i}$ or $k = \underline{i}$, with suitable allowance for the fact that $y_{\underline{ii}} = \underline{\beta} x_{\underline{i}}^*$. ■

The algorithm in the proof deserves some explanation. We start with the definition of $\underline{i}$ and $\underline{\beta}$ in hand, so we already know which consumers will purchase the good at which prices. This pins down the distribution of consumers to markets for valuations higher than $\underline{i}$. According to the inductive hypothesis, we have successfully assigned consumers for valuations above i without violating the optimality of price v_k on segment y^k. The proof shows that there is always enough "room" across all markets to distribute consumers with valuation v_i to markets without violating incentive compatibility, which is the content of equation (22). In fact, when this inequality is strict, there is more slack than we need and there are multiple ways to distribute the v_i consumers. The proof adopts one particular allocation rule, which is to make y_i^k / x_i^* proportional to the amount of slack that would exist in market k *if no more consumers were assigned*, which is exactly $v_k^k y_k^k - v_i \sum_{j=i+1}^{K} y_j^k$. This has implications for the structure of the segmentation. Note that the monopolist will be indifferent between v_k and v^* in each market, so when we get to $i = i^* - 1$, the slack in market k is

$$v_k y_k^k - v_{i^*-1} \sum_{j=i^*}^{K} y_j^k = v_k y_k^k \left(1 - \frac{v^*}{v_{i^*-1}}\right)$$

The term in the parentheses is common to all k, so the slacks are proportional to $v_k y_k^k$. Inductively, this will be true for every valuation less than i^*, and the distribution of valuations below the monopoly price in each market will be proportional to the prior. We will see that these features have analogues in the continuous construction of the next section.

3 The Limits of Discrimination: The Continuum Case

We now extend the arguments to a setting with a continuum of values. We construct segmented markets that mirror those in the environment with finitely many values. Thus we consider a continuum of buyers, each buyer identified by his valuation v, which are distributed on an interval $[\underline{v}, \overline{v}] = V \subset \mathbb{R}_+$ according to a Borel probability measure $x(dv)$. The corresponding distribution function is $F(v) \triangleq x([\underline{v}, v])$.

In the finite environment, the extremal segments x^S with support $S \in S^*$, in which the seller is indifferent between all the values offered as prices in the set S, played a central role. In the continuum environment, a complete description of these segments is rather involved, as the support of any extremal segment does not necessarily have to be connected. Thus, there will be extremal segments whose distribution function have a countable number of discontinuities. As we observed in the finite case, the characterization of the surplus triangle (see Corollary 2) can be achieved with either extremal or direct segmentations, and in this section we use the direct segmentations to establish the consumer surplus maximizing and the social surplus minimizing segmentation. We first state an existence result of minimal and maximal segmentations for arbitrary Borel probability measure $x(dv)$, Theorem 3. We then narrow our analysis to aggregate markets with differentiable distribution functions, for which we explicitly construct the minimal and maximal segmentations. The resulting segmentations, given in Proposition 3 and 4, mirror those in the finite environment, Proposition 1 and 2, respectively.[4]

We shall assume (as in the discrete case, without loss of generality) that there is a unique uniform monopoly price $p^* = v^*$. The monopoly profit under the uniform price v^* is $\pi^* \triangleq v^*(1 - F(v^*))$, and the social surplus and consumer surplus (under the uniform monopoly price v^*) are given by:

$$w^* \triangleq \int_{\underline{v}}^{\overline{v}} v\, x(dv), \text{ and } \quad u^* \triangleq \int_{v^*}^{\overline{v}} (v - v^*)\, x(dv).$$

3.1 Direct Segmentations

In a direct segmentation, each segment is uniquely associated with a "suggested" price $p \in V$, and every price $p \in V$ is at most quoted once. A *pricing rule* in a direct segmentation is a Borel measurable mapping $\phi : V \to V$ that maps the prices that index the direct segments into the prices that are offered to consumers. Without loss of generality, we restrict attention to pure strategies in which each segment p is charged a possibly different price $p' = \phi(p)$. The identity mapping, in which the suggested and realized prices agree, is denoted by ϕ_I. For a given Borel measure σ on $V \times V$, the space of recommended prices and realized valuations, the profit of the firm who uses the pricing rule ϕ is

$$\pi(\sigma, \phi) \triangleq \int_{(v,p)\in V^2} \phi(p)\mathbb{I}_{v\geq\phi(p)}\sigma(dv, dp).$$

[4] The proofs for the environment with a continuum of values are by necessity more elaborate than in the finite environment, and are all collected in the appendix for online publication.

The net utility of the consumers is given by:

$$u(\sigma,\phi) \triangleq \int_{(v,p)\in V^2} (v-\phi(p))\,\mathbb{I}_{v\geq\phi(p)}\sigma(dv,dp).$$

We write $\pi(\sigma) \triangleq \pi(\sigma,\phi_I)$ and $u(\sigma) \triangleq u(\sigma,\phi_I)$. We say that σ is a *direct segmentation* of x if it satisfies the aggregation constraint:

$$\sigma(Y\times V) = x(Y), \tag{23}$$

for all measurable subsets $Y \subseteq V$, and it also satisfies the optimality of the direct pricing strategy: $\pi(\sigma) \geq \pi(\sigma,\phi)$, for every pricing rule ϕ. We say that the measure $x\,(dv)$ is *simple* if F is a step function (and thus a finite set of valuations arise with probability 1). We know that for any discrete distribution over valuations, we can construct direct segmentations that hit both the consumer surplus upper bound and welfare lower bound, see Proposition 1 and 2. These results can be restated as follows:

Corollary 3 (Simple Measures)
If the measure $x(dv)$ is simple, then there exist direct segmentations, $\overline{\sigma}(dv,dp)$ and $\underline{\sigma}(dv,dp)$, that achieve the consumer surplus upper bound and the social surplus lower bound, respectively.

Now consider any Borel measure x and associated distribution function F. We can approximate F by a sequence of step functions F_k that converge to F pointwise at all points of continuity, with associated measures x_k. For each x_k, we can find a direct segmentation, denoted by $\overline{\sigma}_k$, that maximizes the consumer surplus and a direct segmentation, denoted by $\underline{\sigma}_k$ that minimizes social surplus. As the following result shows, these sequences of direct segmentations have convergent subsequences, and importantly, the limits of the subsequences are shown to be direct segmentations as well.

Theorem 2 (Direct Segmentations)
There exist direct segmentations $\overline{\sigma}$ and $\underline{\sigma}$ of x that attain the upper bound on consumer surplus and the lower bound on social surplus, respectively.

The above theorem simply asserts the existence of direct segmentations that achieve the relevant bounds for general measures $x\,(dv)$ and associated distributions $F\,(v)$. If F is differentiable, with an associated density function f, then we can describe specific algorithms to attain the lower bound on social surplus and the upper bound on consumer surplus.

3.2 Consumer Surplus (and Output) Maximizing Segmentation

We now describe a segmentation in which the consumer surplus is $w^* - \pi^*$ and output is maximal and equal to one. For every price $p \in [\underline{v}, v^*]$, there will be a market segment associated with price p, in the sense that it is revenue maximizing to offer the product at price p. Each segment p is constructed so that the distribution of valuations, denoted by $F_p(v)$, satisfies four properties: it (i) has zero mass below p, (ii) has a mass point at p, (iii) is proportional to the prior distribution above p, and (iv) the mass point at p is just large enough to make the seller indifferent between p and v^*. Thus we require that the revenue for offering price p and price v^* is the same: $p = v^*(1 - F_p(v^*))$ .The unique solution to the above four conditions is given by the direct segment $F_p(v)$:

$$F_p(v) \triangleq \begin{cases} 0, & \text{if } \underline{v} \le v < p; \\ 1 - \frac{p(1-F(p))}{v^*(1-F(v^*))}, & \text{if } v = p; \\ 1 - \frac{p(1-F(v))}{v^*(1-F(v^*))}, & \text{if } p < v \le \bar{v}. \end{cases} \tag{24}$$

To complete the description of the market segmentation, we need to specify the distribution of the buyers across the price segments. We write H for the distribution (and h for the corresponding density) over the prices $[\underline{v}, v^*]$ associated with the segments. Given the upper triangular structure of the segments, and the fact that in each segment the density of valuations is proportional to the original density, it is sufficient to insist that for all $v \in [\underline{v}, v^*]$, we have

$$\int_{\underline{v}}^{v} \frac{p}{v^*(1 - F(v^*))} f(v)\, h(p)\, dp + \left(1 - \frac{v(1 - F(v))}{v^*(1 - F(v^*))}\right) h(v) = f(v)\,. \tag{25}$$

In other words, the density $f(v)$ of every valuation v in the aggregate is recovered by integrating over the *continuous part* of the segmented markets, as v is present in every segmented market p with $p < v$, and the *discrete part*, which is due to the presence of the valuation v in the segmented market p with $p = v$.

At this point, it is not obvious that the construction of the distribution H based on the condition (25) should generally succeed. In particular, as we build up $H(p)$ by integrating from below, and thus attempt to absorb the residual density of valuation v in the market segment $p = v$, we might be confronted with two separate issues. First, it could be that we run out of probability to complement the mass point $p = v < v^*$ with higher valuations necessary to construct the indifference (24). Second, it could be that we arrive at $p = v^*$ and still have a positive residual probability $\Pr(v \ge v^*)$ to allocate, which again would not allow us to establish the specific segment $p = v^*$, $F_{v^*}(v)$.

But, in fact the condition (25) implicitly defines a separable ordinary differential equation whose unique solution, given the boundary condition $H(v^*) = 1$, is given by:

$$H(p) \triangleq 1 - \exp\left(-\int_{s=0}^{p} \frac{sf(s)}{(1-F(v^*))v^* - (1-F(s))s} ds\right). \tag{26}$$

Thus, there exists an equilibrium segmentation that attains the upper bound on the consumer surplus in the continuum model that mirrors the earlier result for the case of a finite number of valuations, see (16).

Proposition 3 (Consumer Surplus Maximizing Segmentation)
There exists a direct segmentation, represented by segments $F_p(v)$ and a distribution over segments $H(p)$, where producer surplus is π^*, consumer surplus is $w^* - \pi^*$ and output is socially efficient.

The distribution $H(p)$ implements a particular set of segmented markets which achieve the efficient allocation with the largest possible consumer surplus. Clearly, just as in the finite environment, it is not the unique market segmentation which maximizes consumer surplus and yields an efficient allocation, and extremal segmentations as described in the previous section could also be constructed in the continuum environment which would achieve the same consumer surplus. In the finite environment, we described a greedy algorithm that generated a particular extremal segmentation. The greedy algorithm has a natural translation into a continuum of values which is described in detail in the working paper version, see Bergemann, Brooks, and Morris (2013).

The explicit construction of the direct segmentation and the distribution over prices $H(p)$ allow us to confirm the results stated earlier in Corollary 1. The consumer surplus maximizing segmentation induced by $H(p)$, while maximizing the aggregate consumer surplus, also increases the expected utility of the consumers pointwise, i.e. conditional on the valuation of the consumer. In fact, the very construction of the segmentation $H(p)$ allows us to conclude that the expected sales price, conditional on the valuation v of the consumer, is increasing in the valuation of the consumer.

3.3 Social Surplus (and Output) Minimizing Segmentation

Next, we propose an explicit construction of an output minimizing segmentation for an arbitrary aggregate market $F(v)$. This construction is similar in spirit to the proof of Proposition 2. We again write $h(p)$ for the "size" of the market with price p, and $F_p(v)$ for the conditional distribution of valuations on market p. As in the finite environment, we adopt a construction that starts by placing the highest value consumer in a segment with price equal to his own value. We then work our way down through the values. Consumers with a given value will be apportioned out to all of the segments with weakly higher prices. For this reason, we do not know how large each segment will be until we reach the lowest valuation. Thus, it is convenient to work with the upper cumulative probability $1 - F_p(v)$ of every segment p, and we define the density of valuations v or higher in the segment p by:

$$G_p(v) \triangleq h(p)(1 - F_p(v)). \tag{27}$$

This quantity is analogous to the upper sums of y^k used in the discrete case. After $G_p(v)$ is obtained for all v, we can recover $h(p)$ from $h(p) = G_p(0)$, and obtain the distribution over prices as:

$$H(p) = \int_{v=0}^{p} G_v(0)\, dv,$$

and $F_p(v) = 1 - (G_p(v) / h(p))$.

Let $\widehat{v}$ denote the critical valuation v which achieves the monopoly profit π^* under perfectly discriminatory pricing:

$$\int_{v=\widehat{v}}^{1} v\ f(v)dv = \pi^*,$$

and set $\underline{q} = 1 - F(\widehat{v})$ (which is the equivalent of (6) in the finite environment). We construct the segments $G_p(v)$ for each $p \geq \widehat{v}$ as follows. We set $G_p(v) \triangleq 0$ for $v > p$. For $\widehat{v} < v \leq p$, set $G_p(v) = f(p)$. For $v^* < v < \widehat{v}$, $G_p(v)$ is defined by the differential equation:

$$g_p(v) = f(v)\frac{p\ f(p) - v\ G_p(v)}{v^*(1 - F(v^*)) - v(1 - F(v))}. \tag{28}$$

This is the analogue of the proportional allocation scheme in the previous section: we allocate a share of $f(v)$ to market p that is proportional to the "slack" in the incentive constraint $pf(p) \geq v\ G_p(v)$. The denominator is the integral of this slack over all markets. Note that when we get to v^*, the slack goes to zero and the weights are not defined. Nonetheless, the differential equation is well defined for $v > v^*$. For $v \leq v^*$, we set

$$G_p(v) = (1 - F(v))\frac{G_p(v^*)}{1 - F(v^*)}. \tag{29}$$

This also mirrors the discrete construction, in which markets are proportional to the prior below the monopoly price. Note that $G_p(p) = f(p)$ if $p \geq \widehat{v}$, and $G_p(v) = 0$ for $v > p$ or $p < \widehat{v}$, and so expected profit will be

$$\int_V \int_{v \geq p} v\ G_p(v)\, dvdp = \int_{p \geq \widehat{v}} p\ f(p)dp = \pi^*,$$

and output will be exactly $\underline{q}$. Hence, if $G_p(v)$ is incentive compatible, then it will minimize output. The revenue of the monopolist in market segment p if he sets price v is given by $vG_p(v)$, so the incentive compatibility requirement says that $v\ G_p(v) \leq p\ G_p(p) = p\ f(p)$ for all p.

Proposition 4 (Social Surplus Minimizing Segmentation)
There exists a direct segmentation, represented by segments $G_p(v)$ (and associated $F_p(v)$) and distribution over segments $H(p)$, that results in a conditionally efficient equilibrium segmentation which has zero consumer surplus, producer surplus of π^* and output $\underline{q}$.

We illustrate the preceding results with an example given by the uniform density on the unit interval $[0,1]$. In this case, the uniform monopoly price is $v^* = \frac{1}{2}$. The consumer surplus maximizing segmentation as derived in Proposition 3, leads to an associated distribution function of prices $\overline{H}(p)$ given by:

$$\overline{H}(p) = 1 - \frac{1-p}{1-2p} e^{-\frac{2p}{1-2p}}, \text{ for } p \in \left[0, \frac{1}{2}\right]. \tag{30}$$

By contrast, the segmentation of the consumer in the surplus minimizing allocation as described by Proposition 4 leads to a distribution function of prices given by:

$$\underline{H}(p) = 2p^2 - 1, \text{ for } p \in \left[\frac{1}{\sqrt{2}}, 1\right].$$

The distributions of prices induced by these distinct direct segmentations are displayed below in Figure 3a, where the upper curve represents the surplus maximizing, the lower curve the surplus minimizing distribution of prices.

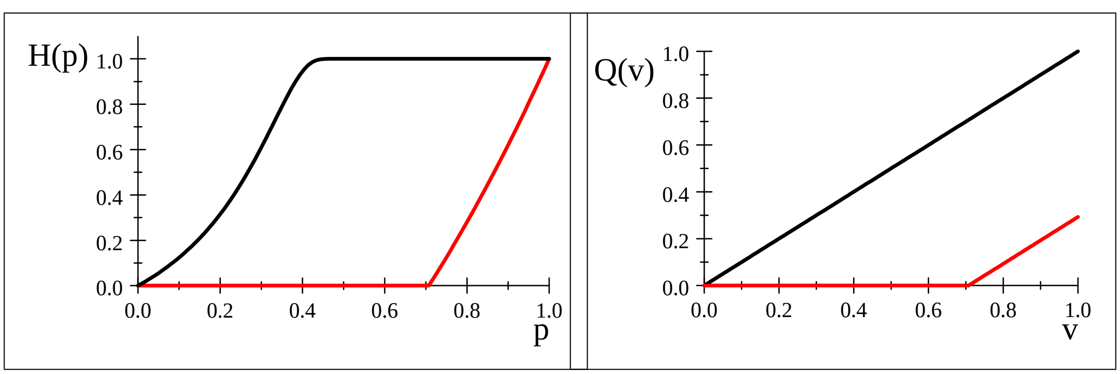

Figure 3a: Distribution of Prices Figure 3b: Distribution of Sales

The surplus minimizing and maximizing distributions represent optimal pricing policies for distinct segmentations of the *same* aggregate market. But even though they share the same aggregate market, we find that have very different structure. In fact, the supports of prices do not overlap at all. These distinct price distributions also lead to very different allocations. The surplus maximizing pricing policy generates all efficient sales, and hence the distribution of sales, $\overline{Q}(v) = v$ exactly replicates the aggregate distribution $F(v)$, and in the current example, the uniform distribution. By contrast, the surplus minimizing distribution truncates sales for values v below $1/\sqrt{2}$. As we described in Proposition 4, the allocation is conditionally efficient, and hence $\underline{Q}(v) = v - 1/\sqrt{2}$ for $v \in [1/\sqrt{2}, 1]$, and zero elsewhere. These different patterns of sales are displayed in Figure 3b, where the upper curve represents the surplus maximizing, the lower curve the surplus minimizing distribution of prices.

4 Discrimination and Segmentation: A Second Approach

So far, the construction of the extremal segmentations that generated the frontier of welfare outcomes relied on two features of the environment. First, each consumer demanded a single unit of the good; and second, the market could be segmented in an arbitrary manner consistent with the aggregation requirement. In this section, we develop a different perspective on price discrimination that does not rely on these assumptions. This permits us to investigate our original question in even broader settings: What is the set of possible welfare outcomes over a range of feasible market segmentations?

In the earlier sections, a segment represented the willingness to pay for a *single unit* of the good by different consumers. But a segment could just as well measure the willingness to pay of a *single agent* who demands more than one unit. In this case, the optimal selling mechanism is not a posted price but rather consists of a menu of quantity-price bundles to screen consumers, i.e. second degree price discrimination. Using the tools of this section, we will analyze markets in which the seller employs a combination of market segmentation and screening.

Pushing in a different direction, much of the existing literature on price discrimination has considered two (or finitely many) exogenously given market segments and asked what would happen if uniform pricing was relaxed to complete discrimination across the segments. There is an intermediate case in which the monopolist can only segment based on noisy signals about the segments, rather than the true underlying segments. The noisy signals effectively induce segments which are convex combinations of the original, exogenous segments. We refer to this as *partial segmentation* and we will give examples of the set of welfare outcomes that can be generated by partial segmentation.

In this section, we restrict attention to the case where there are only two possible types of consumer. This allows us to use a concavification argument used in Kamenica and Gentzkow (2011) to construct optimal information structures from the point of view of maximizing any weighted sum of consumer and producer surplus. Thus, this section both examines the substantive question of the robustness of our analysis, by allowing a richer class of segmentations, as well as documenting a different methodology for analyzing the problem.

4.1 Second Degree Price Discrimination

Up to now, we have considered models in which each buyer demands at most a single unit of the product. We consider a general consumption problem in which the consumer has preferences over a continuum of quantities, and in which the monopolist may use a more complicated mechanism to screen consumers. We shall therefore look at a model of quantity discrimination and allow for fully nonlinear tariffs in each segment. Importantly, the nonlinear tariffs can and will vary across segments. Thus, the results of this section explore what happens when both second and third degree price discrimination are possible.

We now consider a binary version of the model analyzed in the seminal paper by Maskin and Riley (1984), and we can alternatively give our analysis a quality discrimination interpretation, as in the work of Mussa and Rosen (1978). Suppose then that a good can be produced at a variety of quantities $q \in \mathbb{R}_+$. The utility function of an agent with type v is given by:

$$u(v, q, t) \triangleq v\sqrt{q} - t,$$

Hence, utility is concave in the quantity consumed. A proportion α of consumers have low willingness-to-pay, $v_l > 0$, while proportion $1 - \alpha$ have high willingness-to-pay, with $v_h > v_l$. The firm has a positive and constant marginal cost of production of $c > 0$. It follows that the socially efficient quantity to produce is given by

$$q^*(v) \triangleq \left(\frac{v}{2c}\right)^2,$$

and efficient social surplus is given by

$$v\sqrt{q^*(v)} - cq^*(v) = \frac{v^2}{4c}.$$

As before, we are interested in identifying all combinations of consumer and producer surplus that could arise as a result of some market segmentation. With complete information, the producer extracts all the surplus and gets the full gains from trade:

$$w(\alpha) = \alpha\frac{v_l^2}{4c} + (1-\alpha)\frac{v_h^2}{4c},$$

By contrast, if the producer has zero information beyond the prior distribution of the consumers, then the optimal screening solution is for the producer to "exclude" the low valuation buyers if their proportion α is sufficiently small:

$$\alpha \leq \widehat{\alpha} \triangleq 1-\frac{v_l}{v_h},$$

and to sell the socially efficient quantity $q^*(v_h)$ to the high valuation buyer while extracting the entire surplus. On the other hand, if the proportion of low valuation buyers is high, i.e., $\alpha \geq \hat{\alpha}$, then the high type is again sold the efficient quantity $q^*(v_h)$, but now the low type consumer receives quantity:

$$q_l(\alpha) \triangleq \left(\frac{v_l-(1-\alpha)\,v_h}{2\alpha c}\right)^2,$$

and pays

$$t_l(\alpha) \triangleq v_l \frac{(v_l-(1-\alpha)\,v_h)}{2\alpha c},$$

while the high type pays

$$t_h(\alpha) \triangleq \frac{(v_h-v_l)^2+\alpha v_h v_l}{2\alpha c}.$$

Thus, as a function of the composition α of the aggregate market, producer surplus is

$$\pi(\alpha) \triangleq \begin{cases} (1-\alpha)\frac{v_h^2}{4c}, & \text{if} \quad \alpha \leq 1-\frac{v_l}{v_h}, \\ \frac{1}{4\alpha c}\left((v_h-v_l)^2-\alpha v_h(v_h-2v_l)\right), & \text{if} \quad \alpha \geq 1-\frac{v_l}{v_h}; \end{cases} \tag{31}$$

and consumer surplus is

$$u(\alpha) \triangleq \begin{cases} 0, & \text{if} \quad \alpha \leq 1-\frac{v_l}{v_h}; \\ \frac{1-\alpha}{2\alpha c}(v_h-v_l)(v_l-(1-\alpha)\,v_h) & \text{if} \quad \alpha \geq 1-\frac{v_l}{v_h}. \end{cases} \tag{32}$$

In Figure 4a we illustrate the profit $\pi(\alpha)$ and its concavification $\pi^*(\alpha)$, which are the lower and upper curve, respectively. Similarly, the consumer surplus $u(\alpha)$, as well as its concavified versions $u^*(\alpha)$ are displayed in Figure 4b by the lower curve and the upper curve respectively, all for $v_l=1, v_h=2, c=1/2$.

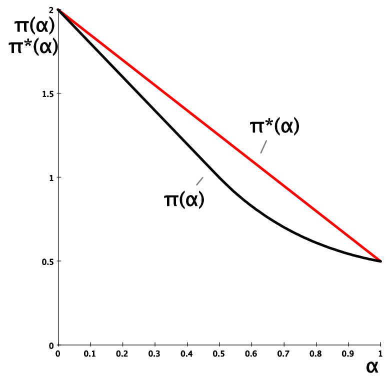

Figure 4a: π and concavified π^*

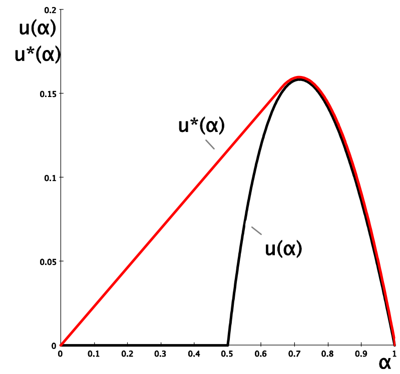

Figure 4b: u and concavified u^*

These illustrations immediately indicate some elementary properties of the profit maximizing or consumer surplus maximizing segmentations, which hold true for all values $0 < v_l < v_h$ and $c > 0$. The concavified profit function $\pi^*(\alpha)$ strictly dominates the convex profit function $\pi(\alpha)$ and hence the seller always prefers pure segmentation, i.e. segments which contain either only low or only high valuations customers. By contrast, it is indicated by the concavified consumer surplus function that the maximal consumer surplus is attained without any segmentation with a large share α of low valuation buyers, whereas a small share α of low valuation buyers requires market segmentation to achieve maximal consumer surplus. Given the binary type space, a segment is uniquely identified by the proportion of low valuation buyers. We denote by s_γ a segment with a fraction γ of low valuation buyers.

We now describe the construction of the entire equilibrium payoff set. We describe the entire frontier of the equilibrium payoff set as the solution to a weighted welfare maximization problem, where we attach the weights λ_u and λ_π to the consumer and producer surplus respectively. Thus, the objective function is $\lambda_u u(\alpha) + \lambda_\pi \pi(\alpha)$. We shall restrict attention to the case where $\lambda_u > 0$, however, the analysis extends to zero or negative weight on consumer surplus in a straightforward manner. With the restriction to $\lambda_u > 0$, it is convenient to normalize the weight of the consumer surplus: $\lambda_u \triangleq 1$, and vary the weight of the producer surplus, setting $\lambda_\pi \triangleq \lambda \in \mathbb{R}_+$, and thus the weighted sum is:

$$w_\lambda(\alpha) \triangleq u(\alpha) + \lambda\pi(\alpha). \tag{33}$$

The concavification of the weighted sum $w_\lambda(\alpha)$ is given by a linear segment that connects $w_\lambda(0)$ with an interior point of the function $w_\lambda(\alpha)$, and a tangency point α_λ, that is uniquely determined as a function of the weight λ:

$$\alpha_\lambda = \frac{(v_h - v_l)(2-\lambda)}{(2-\lambda)v_h - v_l}. \tag{34}$$

Proposition 5 (Segmentation and Second Degree Price Discrimination) *The weighted welfare sum $w_\lambda(\alpha)$ is maximized by :*

1. *for $\lambda > 1$, pure segmentation; the population is divided into segments with s_γ with $\gamma \in \{0,1\}$;*
2. *for $\lambda \leq 1$, mixed segmentation;*
   (a) *if $\alpha < \alpha_\lambda$, the population is divided into two segments s_γ with $\gamma \in \{0, \alpha_\lambda\}$;*
   (b) *if $\alpha \geq \alpha_\lambda$, the population is pooled in a single segment s_α.*

Proof. The weighted welfare sum $w_\lambda(\alpha)$ given by:

$$w_\lambda(\alpha) = \begin{cases} \frac{1}{4c}\lambda v_h^2(1-\alpha), & \text{if } \alpha \leq 1 - \frac{v_l}{v_h}; \\ \frac{1}{4c\alpha}\lambda\left((v_h - v_l)^2 - \alpha v_h(v_h - 2v_l)\right) + \frac{1}{2c\alpha}(v_h - v_l)(v_l - v_h(1-\alpha))(1-\alpha), & \text{if } \alpha > 1 - \frac{v_l}{v_h}. \end{cases}$$

The concavification of $w_\lambda(\alpha)$, denoted by $w_\lambda^*(\alpha)$, is given by a linear segment that connects $w_\lambda(0)$ with an interior point of the function $w_\lambda(\alpha)$, where the linear function has the form

$$l(\alpha) \triangleq \lambda \frac{1}{4c} v_h^2 + \gamma_\lambda \alpha,$$

and the tangency point α_λ and the slope of the linear segment γ_λ are obtained by the unique solution of the tangency condition: $l(\alpha) = w_\lambda(\alpha)$, $l'(\alpha) = w'_\lambda(\alpha)$, which uniquely determines γ_λ and α_λ as follows

$$\alpha_\lambda \triangleq \frac{((2-\lambda)(v_h - v_l))}{(v_h - v_l) + (1-\lambda)v_h}, \quad \gamma_\lambda \triangleq \frac{1}{4c}\frac{v_l^2 - (2-\lambda)\lambda v_h^2}{2-\lambda}.$$

We verify that the contact by the linear segment occurs in the interval $(0,1)$, i.e.

$$\alpha_\lambda = \frac{(v_h - v_l)(2-\lambda)}{(2-\lambda)v_h - v_l} \leq 1,$$

and thus find that

$$\frac{(v_h - v_l)(2-\lambda)}{(2-\lambda)v_h - v_l} \leq 1 \Leftrightarrow \lambda \leq 1.$$

which establishes the results. ■

Using this characterization, we can explicitly compute the set of equilibrium consumer and producer surplus pairs. For any given α, we know that the expected payoffs must be contained in a triangle as before: social surplus cannot be more than $w^*(\alpha)$, consumer surplus is at least 0 and producer surplus is at least $\pi(\alpha)$. But in contrast to the earlier analysis with linear rather than concave utility in the quantity, the set of equilibrium payoffs that can arise in some form of segmentation is given by a strictly smaller set, namely the shaded area, and hence a strict subset of the surplus triangle.

We focus on the consumer surplus maximizing segmentation and the comparison with the equilibrium surplus in the aggregate market. If there are few buyers with a low valuation, then in the aggregate market, the seller will not offer a product to the low valuation agents. We refer to this as the case of the *exclusive prior*. Here, in the equilibrium without any additional information, the seller extracts all the surplus from the high valuation buyers, and the equilibrium is socially inefficient. An important consequence of the exclusive prior is that any non-trivial segmentation will increase social surplus, and hence strictly increase the revenue of the seller and weakly increase the surplus of the buyers. Importantly, in cases where there is non-trivial screening, any attempt to increase the surplus of the buyers, and hence their information rent, leads to an inefficient allocative decision by the seller. In consequence, the efficient frontier can only be reached with perfect segmentation $s \in \{s_0, s_1\}$, as illustrated for $\alpha = 1/3$ in Figure 5a.

If the prior probability α is above the critical point $\widehat{\alpha}$, then the seller starts offering a low quantity version of the product to the low value buyers in the aggregate market, the case of the *inclusive prior*. In contrast to the case of the exclusive prior, there now exist segmentations which strictly improve the revenue of the seller while lowering consumer surplus. As before the efficient frontier can be attained only through perfect segmentation. Eventually, as α increases above $\alpha_\lambda > \widehat{\alpha}$, the equilibrium in the aggregate market leads to the largest possible consumer surplus. In fact, any segmentation now increases the revenue of the seller and strictly decreases the surplus of the buyers. We have thus arrived at an environment where segmentation (and hence additional information for the seller) unambiguously increases his revenue and decreases consumer surplus. This is illustrated for $\alpha = 0.9$ in Figure 5b.

Finally, if the share α of low value buyers is between $\widehat{\alpha}$ and α_λ, then there are segmentations of the aggregate market, in particular those involving s_0 and s_{α_λ}, that can increase both the profits of the seller and the surplus of the buyers. The resulting equilibrium set then displays features of both of the above sets. Namely, there are segmentations that increase both the consumer's and the producer's as in the exclusive prior, but the Bayes Nash equilibrium of the aggregate market already leaves the consumer with some information rent, just as in the above case of the inclusive prior with a large proportion α of low value buyers.[5]

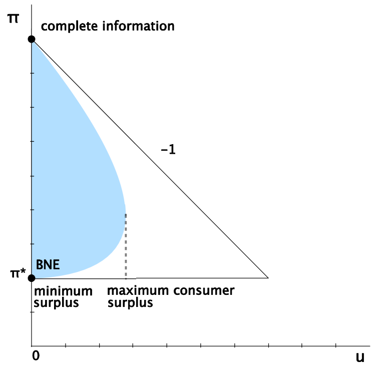

Figure 5a: Quantity discrimination, exclusive prior.

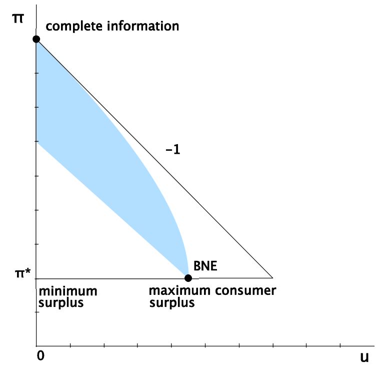

Figure 5b: Quantity discrimination, inclusive prior.

4.2 Partial Segmentation

We conclude with a distinct interpretation of the quantity discriminating monopolist, which allows us to link the present analysis more closely to the traditional analysis of third degree price discrimination. We continue with the binary type model, but now take each type to represent a separate market with a distinct demand function given by $q_i(p)$, with $i = l, h$. The aggregate demand function is given by:

$$q(p) \triangleq \alpha q_l(p) + (1 - \alpha)\, q_h(p),$$

where α represent the share of the "low" demand market l. Our analysis follows the approach taken in the previous subsection. We start by computing consumer and producer surplus as a function of α. Then, for every possible (positive) weight of consumer and producer surplus, we find the concavification of the weighted sum of these two objectives. This is a versatile technique that can be used to solve for the potential welfare consequences of partial segmentation for any demand specification. We will use the remainder of this section to exhibit the (numerical) solution of two prominent examples that have been considered in the literature.

[5]We analyzed the problem of quantity discrimination with concave utility functions and linear cost function. In fact, already in the model with a single unit demand, i.e. without quality discrimination, but increasing aggregate costs, the above qualifications regarding the set of attainable equilibrium payoffs obtain.

Linear Demand The classic example of market segmentation is that of two markets with linear demand. This example was first explored in Pigou (1920), who famously concluded that uniform price and full segmentation both result in the same output, but full segmentation allocates the good inefficiently and reduces welfare. In the linear example of Pigou (1920), demand is given in market $i \in \{l, h\}$ by:

$$q_i(p) = \begin{cases} 0, & \text{if} \quad p \geq b_i, \\ b_i - p, & \text{if} \quad 0 \leq p < b_i; \end{cases}$$

and we set $b_l = 1$ and $b_h = b \geq 1 + \sqrt{2}$. If the share of the low demand market is sufficiently small, or

$$\alpha \leq \frac{b(b-2)}{(b-1)^2} \triangleq \widehat{\alpha},$$

then it is optimal to exclude the low demand segment in the aggregate market by setting the uniform price $p^* = \frac{b}{2}$. By contrast, if $\alpha \geq \widehat{\alpha}$, then in the aggregate market it is optimal to serve both segments at a price

$$p^* = \frac{b - (b-1)\alpha}{2}.$$

We can readily compute producer and consumer surplus as:

$$\pi(\alpha) = \begin{cases} (1-\alpha)\left(\frac{b}{2}\right)^2, & \text{if} \quad \alpha \leq \alpha^*, \\ \left(\frac{b-(b-1)\alpha}{2}\right)^2, & \text{if} \quad \alpha > \alpha^*, \end{cases}$$

and

$$u(\alpha) = \begin{cases} \frac{1-\alpha}{2}\left(\frac{b}{2}\right)^2, & \text{if} \quad \alpha < \widehat{\alpha} \\ \frac{1-\alpha}{2}\left(b - \frac{b-(b-1)\alpha}{2}\right)^2 + \frac{\alpha}{2}\left(1 - \frac{b-(b-1)\alpha}{2}\right)^2, & \text{if} \quad \alpha \geq \widehat{\alpha} q. \end{cases}$$

For an inclusive prior $\alpha > \widehat{\alpha}$, the frontier of welfare outcomes is generated by two families of segmentations: the first consists of perfect discrimination, where one segment has demand q_h and the other has demand q_l. The second family is indexed by:

$$\beta \in \left[0, \frac{1-\alpha}{1-\widehat{\alpha}}\right],$$

and consists of a segment of size β with demand $\widehat{\alpha} q_l + (1-\widehat{\alpha})q_h$, and a segment of size $1-\beta$ that has demand $\gamma q_l + (1-\gamma)q_h$ where $\gamma = \frac{\alpha - \beta\widehat{\alpha}q}{1-\beta}$. Note that $\beta = 0$ corresponds to a single segment which is the aggregate market, and $\beta = \frac{1-\alpha}{1-\widehat{\alpha}}$ corresponds to having one segment consist of only low demand consumers. Below we illustrate in Figure 6a the attainable equilibrium surplus set for $b = 1 + \sqrt{2}$ and $\alpha = \frac{2}{3}$.

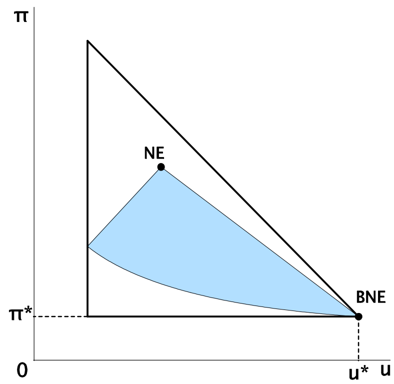

Figure 6a: Pigouvian
Surplus Set

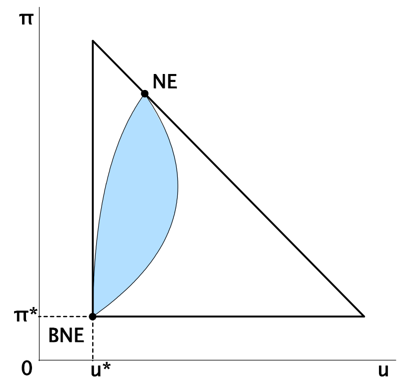

Figure 6b: Logistic Demand
Surplus Set

We see that even in this simple setting, there is a large set of possible welfare outcomes that can result from partial segmentation. Nonetheless, uniform pricing remains the best for consumers, and full segmentation is necessarily best for the producer.

Logistic Demand Our second example is drawn from Cowan (2012) and here demand follows the logistic function for $i \in \{l, h\}$:

$$q_i(p) = \frac{1}{1 + \exp(p - a_i)},$$

with $a_h > a_l > 0$. With logistic demand, both markets are always served, full discrimination always raises welfare, and under fairly general conditions full discrimination raises consumer surplus as well (see Cowan (2012), (2013)).

For this demand specification, there is no closed form expression for the optimal price as a function of α. Nonetheless, it is straightforward to compute the optimal price numerically. We illustrate in Figure 6b the attainable equilibrium surplus set for $a_h = 3$, $a_l = 1$, and $\alpha = 0.5$.

Similar to the Pigouvian example, the frontier of the welfare set is generated by two families of segmentations: A family where one of the segment has only high demand consumers, and a family where one segment has only low demand consumers. Both consumer surplus and welfare rise with full discrimination. Interestingly, consumer surplus is not maximized at full discrimination, but rather at a partial segmentation where one segment has only high demand consumers.

The takeaway from these examples is that even with restrictions on the form of segmentation, such as a convex combination of two given segments, there will generally be a large set of possible welfare outcomes due to partial segmentation. Many objectives, e.g., maximizing consumer surplus, will be achieved by segmentations that give the monopolist an intermediate level of information about demand.

5 Conclusion

It was the objective of this paper to study the impact of information on the efficiency and the distribution of surplus in a canonical setting of monopoly price discrimination. We showed that additional information above and beyond the prior distribution can have a substantial effect on producer and consumer surplus. In general, there are many directions in which welfare could move relative to the benchmark of a unified market. We showed that while additional information can never hurt the seller, it can lead social and consumer surplus to both increase, both decrease, or respectively increase and decrease. The range of these predictions is established without any restrictions on the distribution in the aggregate market, and in particular does not rely on any regularity or concavity assumption with respect to the aggregate distribution or profit function.

Exactly which form of market segmentation arises in practice is no doubt influenced by many factors, which may include technological and legal limitations on how information can be collected and used. In an age in which individuals are increasingly concerned about the preservation of privacy, it is important to understand the welfare consequences that may result from companies amassing data on consumers' preferences. Our findings indicate that the relationship between efficiency and information can only be understood in the context of how data will be used, and this crucially depends on the preferences of those who collect the information. Thus, a natural and important direction for future research is to better understand which forms of price discrimination will endogenously arise, and for whose benefit.

References

Acquisti, A., and H. Varian (2005): "Conditioning Prices on Purchase History," *Marketing Science*, 24, 367–381.

Aguirre, I., S. Cowan, and J. Vickers (2010): "Monopoly Price Discrimination and Demand Curvature," *American Economic Review*, 100, 1601–1615.

Aumann, R., and M. Maschler (1995): *Repeated Games with Incomplete Information*. MIT.

Bergemann, D., B. Brooks, and S. Morris (2013): "The Limits of Price Discrimination," Discussion Paper 1896, Cowles Foundation for Research in Economics, Yale University.

Bergemann, D., and S. Morris (2013a): "Bayes Correlated Equilibrium and Comparison of Information Structures," Discussion paper, Cowles Foundation for Research in Economics, Yale University.

——— (2013b): "Robust Predictions in Games with Incomplete Information," *Econometrica*, forthcoming.

Cowan, S. (2012): "Third-Degree Price Discrimination and Consumer Surplus," *Journal of Industrial Economics*, 60, 333–345.

——— (2013): "Welfare-Increasing Third-Degree Price Discrimination," Discussion paper, Oxford University.

Kamenica, E., and M. Gentzkow (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615.

Maskin, E., and J. Riley (1984): "Monopoly with Incomplete Information," *Rand Journal of Economics*, 15, 171–196.

Mussa, M., and S. Rosen (1978): "Monopoly and Product Quality," *Journal of Economic Theory*, 18, 301–317.

Nahata, B., K. Ostaszewski, and P. Sahoo (1990): "Direction of Price Changes in Third-Degree Price Discrimination," *American Economic Review*, 80, 1254–1258.

Pigou, A. (1920): *The Economics of Welfare*. Macmillan, London.

Robinson, J. (1933): *The Economics of Imperfect Competition*. Macmillan, London.

Schmalensee, R. (1981): "Output and Welfare Implications of Monopolistic Third-Degree Price Discrimination," *American Economic Review*, 71, 242–247.

Taylor, C. (2004): "Consumer Privacy and the Market for Customer Information," *RAND Journal of Economics*, 35, 631–651.

Varian, H. (1985): "Price Discrimination and Social Welfare," *American Economic Review*, 75, 870–875.

6 Appendix for Online Publication

We establish Theorem 2 through a sequence of lemmas that begin by establishing general results about sequences of direct segmentations. In the following, we take $\{\sigma_k\}$ to be a convergent subsequence that converges to σ.

Lemma 2 *For any sequence $\{\sigma_k\}$ of direct segmentations on $\{x_k\}$, there exists σ and a subsequence $\{\sigma_{k_l}\}$such that $\sigma_{k_l} \Rightarrow \sigma$.*

Proof. Since V^2 is a compact metric space, the space of Borel measures on V^2 is compact in the weak topology. Moreover, this topology is metrizable with the Prokhorov metric. Compact metric spaces are sequentially compact, so a subsequence of σ_k converges to some σ. ■

The limit σ is a measure on V^2, but at this point it is not clear that it is a direct segmentation. Weak convergence guarantees that the expectation of any continuous function on V^2 under σ_k converges to its expectation under σ. But profit is not a continuous function; it has a discontinuity where $v = p$. Indeed, there will be pricing rules for which $\pi(\sigma_k, \phi)$ converges to something strictly less than $\pi(\sigma, \phi)$. Nonetheless, for any ϕ, there exists a ϕ' such that $\pi(\sigma_k, \phi')$ *does* converge to $\pi(\sigma, \phi')$, and $\pi(\sigma, \phi')$ is close to $\pi(\sigma, \phi)$.

Lemma 3 *Suppose $\{\sigma_k\}$ are direct segmentations such that $\sigma_k \Rightarrow \sigma$. Then for any $\epsilon > 0$ and pricing rule ϕ, there exists a pricing rule ϕ' such that*

$$\pi(\sigma, \phi') > \pi(\sigma, \phi) - \epsilon,$$

and

$$\pi(\sigma_k, \phi') \to \pi(\sigma, \phi').$$

Proof. By Lusin's theorem, for every $\tilde{\epsilon}$, there exists a continuous function $\phi^{\tilde{\epsilon}}$ that is continuous and coincides with ϕ except on a set of measure $\tilde{\epsilon}$. Here, the measure is taken to be the marginal measure of σ on p, denoted $\sigma_p(dp)$, i.e. for any Borel set $Y \subseteq V$, $\sigma_p(Y) = \sigma(V \times Y)$. Note that $|\pi(\sigma, \phi^{\tilde{\epsilon}}) - \pi(\sigma, \phi)| < \tilde{\epsilon}\bar{v}$.

Now define $\widehat{\phi}^t \triangleq \max\{0, \phi^{\tilde{\epsilon}} - t\tilde{\epsilon}\}$ for $t \in (0, 1)$. In other words, $\widehat{\phi}^t$ is $\phi^{\tilde{\epsilon}}$ translated down by $t\tilde{\epsilon}$, with truncation at zero. Define $\text{epi}(h) \triangleq \{(v, p) \in V^2 | v \geq h(p)\}$ to be the epigraph of the Borel function $h : V \to V$.

Claim: If $\phi^{\tilde{\varepsilon}} \neq 0$, there exists a $t \geq 0$ such that $\sigma(\partial \text{epi}(\widehat{\phi}^t)) = 0$. For each $t \geq 0$, the $\partial \text{epi}(\widehat{\phi}^t)$ are disjoint sets; if $\sigma(\partial \text{epi}(\widehat{\phi}^t)) > 0$ for all t, then taking the union over all such sets we would find that the set V^2 has infinite measure.

If $\phi^{\tilde{\varepsilon}} \neq 0$, take any t such that $\sigma(\partial \text{epi}(\widehat{\phi}^t)) = 0$, and let $\phi' = \widehat{\phi}^t$. Otherwise, we can set $\phi' = \phi^{\tilde{\varepsilon}} = 0$. Let $Y = \text{epi}(\phi')$. Note that the set Y is compact (being the epigraph of a continuous function) and σ-continuous. Write $\sigma_k|Y$ and $\sigma|Y$ for the respective measures restricted to Y.

Claim: $\sigma_k|Y \Rightarrow \sigma|Y$. This is true if $\sigma_k|Y(Z) \to \sigma|Y(Z)$ for all $Z \subseteq Y$such that $\sigma|Y(\partial Z) = 0$. Since $\sigma(\partial Y) = 0$, and $\partial Z \subseteq Y$ (since Y is closed), then any $\sigma|Y$-continuous Z must also be σ-continuous, since $\sigma(\partial Z) = \sigma(\partial Z \cap \partial Y) + \sigma(\partial Z \setminus \partial Y) = \sigma|Y(\partial Z \setminus \partial Y) = \sigma|Y(\partial Z) = 0$. Thus, $\sigma_k|Y(Z) = \sigma_k(Z) \to \sigma(Z) = \sigma|Y(Z)$, and we are done.

Note that the function ϕ' is continuous when restricted to Y (since ϕ' itself is continuous). Since $\sigma_k|Y \Rightarrow \sigma|Y$ and ϕ' is zero outside of Y, we have

$$\begin{aligned}
\lim_{k\to\infty} \pi\left(\sigma_k, \phi'\right) &= \lim_{k\to\infty} \int_{V^2} \phi'(p) 1_{v \geq \phi'(p)} \sigma_k(dv, dp) \\
&= \lim_{k\to\infty} \int_Y \phi'(p) \sigma_k|Y(dv, dp) \\
&= \lim_{k\to\infty} \int_Y \phi'(p) \sigma_k(dv, dp) \\
&= \int_Y \phi'(p) \sigma(dv, dp) \\
&= \int_{V^2} \phi'(p) 1_{v \geq \phi'(p)} \sigma(dv, dp) \\
&= \pi\left(\sigma, \phi'\right).
\end{aligned}$$

And finally, observe that $\phi^{\tilde{\epsilon}} - \tilde{\epsilon} < \phi' \leq \phi^{\tilde{\epsilon}}$, so

$$\begin{aligned}
\pi(\sigma, \phi') &= \int_{V^2} \phi'(p) 1_{v \geq \phi'(p)} \sigma(dv, dp) \\
&\geq \int_{V^2} \phi'(p) 1_{v \geq \phi^{\tilde{\varepsilon}}(p)} \sigma(dv, dp) \\
&= \pi(\sigma, \phi^{\tilde{\epsilon}}) - \int_{V^2} (\phi^{\tilde{\epsilon}}(p) - \phi'(p)) 1_{v \geq \phi^{\tilde{\varepsilon}}(p)} \sigma(dv, dp) \\
&\geq \pi(\sigma, \phi^{\tilde{\epsilon}}) - \tilde{\epsilon} \int_{V^2} 1_{v \geq \phi^{\tilde{\epsilon}}(p)} \sigma(dv, dp) \\
&\geq \pi(\sigma, \phi^{\tilde{\epsilon}}) - \tilde{\epsilon}.
\end{aligned}$$

Thus,

$$\pi(\sigma, \phi') \geq \pi(\sigma, \phi) - (\bar{v} + 1)\tilde{\epsilon}$$

So taking $\tilde{\epsilon} < \frac{\epsilon}{\bar{v}+1}$, we have the desired result. ■

The first condition says that ϕ' achieves a payoff for the monopolist within ϵ of ϕ. The second condition is that the payoff from ϕ is continuous in the limit. Using this result, we can prove properties of the limit measure σ.

Lemma 4 *Suppose $\{\sigma_k\}$ are direct segmentations of $\{x_k\}$ such that $\sigma_k \Rightarrow r$ and $x_k \Rightarrow x$. Then σ is a direct segmentation of x. Moreover, $u(\sigma_k) \to u(\sigma)$ and $\pi(\sigma_k) \to \pi(\sigma)$.*

Proof. First we show that σ has x as a marginal measure. Take any continuous and bounded function $\xi(v)$ on V. Then clearly $\int_{V^2} \xi(v)\sigma_k(dv, dp) = \int_V \xi(v)x_k(dv) \to \int_V \xi(v)x(dv)$. But $\xi(v)$ is a continuous function of (v, p) as well, so $\int_{V^2} \xi(v)\sigma_k(dv, dp) \to \int_{V^2} \xi(v)\sigma(dv, dp)$. Thus, $\int_{V^2} \xi(v)\sigma(dv, dp) = \int_V \xi(v)x(dv)$ for all continuous and bounded $\xi(v)$, and we are done.

Note that $(v - p)1_{v\geq p}$ is continuous in v and p, so $u(\sigma_k) \to u(\sigma)$ follows from weak convergence. To see that $\pi(\sigma_k) \to \pi(\sigma)$, observe that $p1_{v\geq p}$ is upper semi-continuous, so $\limsup_{k\to\infty} \pi(\sigma_k) \leq \pi(\sigma)$. Suppose the inequality is strict. Then by the previous Lemma, for every $\epsilon > 0$ there exists a pricing rule ϕ' such that $\pi(\sigma_k, \phi') \to \pi(\sigma, \phi') \geq \pi(\sigma) - \epsilon$. But $\pi(\sigma_k) \geq \pi(\sigma_k, \phi')$, since σ_k is a direct segmentation, a contradiction. Hence $\pi(\sigma_k) \to \pi(\sigma)$.

Finally, we show that $\pi(\sigma) \geq \pi(\sigma, \phi)$ for all pricing rules ϕ. If not, again we can find an ϕ' such that $\pi(\sigma, \phi') > \pi(\sigma)$ and $\pi(\sigma_k, \phi') \to \pi(\sigma, \phi')$. But $\pi(\sigma_k) \geq \pi(\sigma_k, \phi')$, so $\lim_{k\to\infty} \pi(\sigma_k) > \pi(\sigma)$, a contradiction. ■

We are now able to establish Theorem 2.

Proof of Theorem 2. Take $\overline{\sigma}$ to be a limit of a subsequence of $\overline{\sigma}_k$, and $\underline{\sigma}$ to be a limit of a subsequence of $\underline{\sigma}_k$. We know that these limits exist and they are direct segmentations, and that π and u converge continuously. All that remains to show is that π and u converge to the bounds for x.

For all k, we have $\pi(\underline{\sigma}_k) = \max_p p(1 - F_k(p))$. So we can show that $\max_p p(1 - F_k(p)) \to \max_p p(1 - F(p))$. Take v^* to be the solution to $\max_p p(1 - F(p))$. Since F has countably many discontinuities, for every $\epsilon > 0$ there is a $p^\epsilon > v^*$ such that F is continuous at p^ϵ, and $v^*(1 - F(v^*)) - p^\epsilon(1 - F(p^\epsilon)) < \epsilon$. Since p^ϵ is a continuity point, by weak convergence $F_k(p^\epsilon) \to F(p^\epsilon)$, so

$$\begin{aligned}\lim_{k\to\infty} \max_p p(1 - F_k(p)) &\geq \lim_{k\to\infty} p^\epsilon(1 - F_k(p^\epsilon)) \\ &= p^\epsilon(1 - F(p^\epsilon)) \\ &\geq v^*(1 - F(v^*)) - \epsilon\end{aligned}$$

showing that $\lim_{k\to\infty} \max_p p(1-F_k(p)) \geq \max_p p(1-F(p))$. Write p^k for a solution to $\max_p p(1-F_k(p))$; the p^k live in the compact set V, so there is a subsequence that converge to some $\hat{p}$. Again, there is a $p^\epsilon > \hat{p}$ at which F is continuous and

$$\begin{aligned} p^\epsilon(1 - F(p^\epsilon)) &= \lim_{k\to\infty} p^\epsilon(1 - F_k(p^\epsilon)) \\ &\geq \lim_{k\to\infty} p^k(1 - F_k(p^k)) - \epsilon \end{aligned}$$

and we are done.

Clearly $u(\underline{\sigma}_k) = 0$ for all k, so $u(\underline{\sigma}) = 0$. Also, we know that $u(\overline{\sigma}_k) = \int_V v\ x_k(dv) - \pi(\sigma_k)$. Since $\phi(v) = v$ is a continuous function, $u(\overline{\sigma}_k) \to \int_V v\ x(dv) - \pi(\bar{\sigma})$, and we are done. ∎

Proof of Theorem 3. We verify that the solution (26) of the density:

$$h(p) = \frac{(1 - F(p^*)) f(p) p^*}{(1 - F(p^*)) p^* - (1 - F(p)) p} e^{-\int\limits_{s=0}^{p} \frac{sf(s)}{(1-F(p^*))p^*-(1-F(s))s} ds}. \tag{35}$$

solves the balancing condition:

$$\int_{\underline{v}}^{v} \frac{p}{p^*(1 - F(p^*))} f(v) h(p) dp + \left(1 - \frac{v(1 - F(v))}{p^*(1 - F(p^*))}\right) h(v) = f(v). \tag{36}$$

Thus inserting (35) into (36) we get:

$$\int_{\underline{v}}^{v} \frac{p}{p^*(1 - F(p^*))} f(v) \frac{(1 - F(p^*)) f(p) p^*}{(1 - F(p^*)) p^* - (1 - F(p)) p} \left(-\int\limits_{s=0}^{p} \frac{sf(s)}{(1 - F(p^*)) p^* - (1 - F(s)) s} ds\right) dp \qquad +$$

$$+ \left(\frac{p^*(1 - F(p^*)) - v(1 - F(v))}{p^*(1 - F(p^*))}\right) \frac{(1 - F(p^*)) f(p) p^*}{(1 - F(p^*)) p^* - (1 - F(p)) p} e^{-\int\limits_{s=0}^{p} \frac{sf(s)}{(1-F(p^*))p^*-(1-F(s))s} ds} \quad = \quad f(v),$$

or

$$\int_{\underline{v}}^{v} \frac{pf(p)}{(1 - F(p^*)) p^* - (1 - F(p)) p} e^{-\int\limits_{s=0}^{p} \frac{sf(s)}{(1-F(p^*))p^*-(1-F(s))s} ds} dp = 1 - e^{-\int\limits_{s=0}^{v} \frac{sf(s)}{(1-F(p^*))p^*-(1-F(s))s} ds}.$$

So, after integration by parts, we get:

$$\int_{\underline{v}}^{v} \frac{pf(p)}{(1 - F(p^*)) p^* - (1 - F(p)) p} e^{-\int\limits_{s=0}^{p} \frac{sf(s)}{(1-F(p^*))p^*-(1-F(s))s} ds} dp = 1 - e^{-\int\limits_{s=0}^{v} \frac{sf(s)}{(1-F(p^*))p^*-(1-F(s))s} ds}.$$

So, if we define

$$H(p) = 1 - e^{-\int\limits_{s=0}^{p} \frac{sf(s)}{(1-F(p^*))p^* - (1-F(s))s} ds}$$

then

$$h(p) = H'(p) = \frac{pf(p)}{(1 - F(p^*))p^* - (1 - F(p))p} e^{-\int\limits_{s=0}^{p} \frac{sf(s)}{(1-F(p^*))p^* - (1-F(s))s} ds}$$

and so

$$\int_0^v H'(p)\,dp = [H(p)]_0^v = H(p) - H(0).$$

The distribution function $H(p)$ is everywhere continuous, and in particular does not have a mass point at $p = p^*$ as the integral in the exponential diverges, that is

$$\lim_{p \to p^*} \int_0^p \frac{sf(s)}{(1 - F(p^*))p^* - (1 - F(s))s} ds = \infty.$$

For the divergence of the integral, it is sufficient to establish that the term inside the integral grows sufficiently fast as $p \to p^*$:

$$\frac{sf(s)}{(1 - F(p^*))p^* - (1 - F(s))s}.$$

By the $p-$test for divergence:

$$\int_0^1 \frac{1}{x^p} dx$$

is convergent if and only if $p < 1$. It thus follows that the integral always diverges, (as it relies on the square rather than the linear term, due to the first condition), and hence there is no mass point at the optimal price v^*. Namely, we can approximate the above ratio, using the quadratic polynomial:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2$$

and applying it to the function $X(s)$ as defined below:

$$X(s) \triangleq (1 - F(s))s,$$

we get

$$X'(s) = (1 - F(s)) - f(s)s, \quad X''(s) = -2f(s) - f'(s)s,$$

and thus we have the following approximation, using the fact the $0-$th term and the $1-$st term vanish, (the later due to the first order condition):

$$\frac{pf(p)}{(2f(p^*) + f'(p^*)p^*)(p - p^*)^2}.$$

The approximation rate is quadratic rather than sublinear and hence the integral diverges as $p \to p^*$.
■

Proof of Theorem 4. Write $G(v) = 1 - F(v)$, and let $G_p(v)$ be the density of consumers who are offered price p and have valuation at least v. Set $G_p(v) = f(p)$ for $p \in [\hat{v}, \bar{v}]$ and $v \in [\hat{v}, p]$ and $G_p(v) = 0$ for $v > p$. Note that for $v \geq \hat{v}$, $\int_{\hat{v}}^{\bar{v}} G_p(v)dp = \int_v^{\bar{v}} f(p)dp = G(v)$. We construct $G_p(v)$ for $v < \hat{v}$. We want to maintain $G_p(v') < \frac{v\ f(v)}{v'}$. To that end, let $g_p(v) = \frac{d}{dt}G_p(t)\big|_{t=v}$. If $v\ G(v) < v^*\ G(v^*)$, set

$$g_p(v) = g(v)\frac{p\ f(p) - v\ G_p(v)}{v^*\ G(v^*) - v\ G(v)},$$

and otherwise set $g_p(v) = \frac{p\ f(p)}{v^2}$.

Claim 1: $\int_{\hat{v}}^{\bar{v}} G_p(v)dp = G(v)$. We already argued that this is true for $v \geq \hat{v}$. For $v < \hat{v}$, note

$$\begin{aligned}\int_{\hat{v}}^{\bar{v}} G_p(v)dp &= \int_{\hat{v}}^{\bar{v}} \left[G_p(\hat{v}) + \int_v^{\hat{v}} g_p(t)dt\right] dp \\ &= G(\hat{v}) + \int_v^{\hat{v}} \int_{\hat{v}}^{\bar{v}} g_p(t)dp\ dt \\ &= G(\hat{v}) + \int_v^{\hat{v}} g(t)\frac{v^*G(v^*) - t\int_{\hat{v}}^{\bar{v}} G_p(t)dp}{v^*G(v^*) - t\ G(t)}dt.\end{aligned}$$

By induction, if $\int_{\hat{v}}^{\bar{v}} G_p(t)dp = G(t)$ for all $t < v$, then it must be true for v as well, since the weight on $g(t)$ inside the integral is 1. Since it's true for $v = \hat{v}$, we are done.

Claim 2: $G_p(v) \leq \frac{p\ f(p)}{v}$ for all p and v. Suppose not, and let v' be the largest v at which $v\ G_p(v)$ goes above $p\ f(p)$ for some p, i.e. $v\ G_p(v) > v'G_p(v') = p\ f(p)$ for all $v \in (v' - \epsilon, v')$. Since $G_p(v)$ is differentiable, it must be that

$$\frac{d}{dt}\left[t\ G_p(t)\right]\Big|_{t\uparrow v'} = -v'g_p(v') + G_p(v') < 0.$$

Note that $G_p(v') \geq 0$, since $v\ G_p(v) \leq p\ f(p)$ for $v \geq v'$, and therefore $g_p(v)$ is positive on that region. For the derivative at v' to be negative, we would then need that $v'g_p(v') > G_p(v')$. However, from the definition of $g_p(v)$ it's clear that if $v'G(v') < v^*G^*(v)$, $g_p(v') = 0$, which would be a contradiction. If $v'G(v') \geq v^*G(v^*)$, then $g_p(v') = \frac{p\ f(p)}{v'^2}$, so

$$-v'g_p(v') + G_p(v') = -\frac{p\ f(p)}{v'} + G_p(v') = 0,$$

again a contradiction. Hence, it must be that $G_p(v) \leq \frac{p\ f(p)}{v}$ for all v.

Claim 3: $v^* G_p(v^*) = p\ f(p)$. Follows easily from the previous two claims, since

$$\begin{aligned}\int_{\hat{v}}^{\bar{v}} p\ f(p)dp &= v^* G(v^*)\\ &= \int_{\hat{v}}^{\bar{v}} v^* G_p(v^*)dp,\end{aligned}$$

and $G_p(v^*) \leq \frac{p\ f(p)}{v^*}$ for every p, so in fact they must be equal (almost everywhere).

In fact, it is always the case that $G_p(v) = G(v)\frac{G_p(v^*)}{G(v^*)}$ for $v < v^*$, for then

$$\begin{aligned}g(v)\frac{v^* G_p(v^*) - vG_p(v)}{v^* G(v^*) - vG(v)} &= g(v)\frac{v^* G_p(v^*) - vG(v)\frac{G_p(v^*)}{G(v^*)}}{v^* G(v^*) - vG(v)}\\ &= g(v)\frac{G_p(v^*)}{G(v^*)} = \frac{dG_p}{dv}\end{aligned}$$

so the ODE is satisfied. The general solution for $v > v^*$ is

$$G_p(v) = e^{\int_v^1 \frac{xg(x)dx}{v^* G(v^*) - xG(x)}dx}\left(e^{-\int_{\hat{v}}^1 \frac{xg(x)}{v^* G(v^*) - xG(x)}dx} + p\int_{\hat{v}}^{v} \frac{e^{-\int_x^1 \frac{yg(y)}{v^* G(v^*) - yG(y)}dy} g(x)}{v^* G(v^*) - xG(x)}dx\right),$$

and hence, the segments are linear interpolations between $G_{\hat{v}}(v)$ and $G_1(v)$. ∎