Monopoly pricing, optimal rationing, and resale*

Simon Loertscher† Ellen V. Muir‡

September 19, 2019

Abstract

Non-market clearing prices that induce rationing and open scope for resale–that sellers actively attempt to prevent–are a persistent feature of reality but have proved puzzling for theory. Why, one wonders, would the seller not set market clearing prices in the first place, thereby increasing revenue and preempting resale? We first show that the phenomenon of non-market clearing pricing together with the prevention of resale is consistent with optimal behaviour by a monopoly seller. Selling a given quantity with non-market clearing prices is optimal if and only if the revenue function is convex at this quantity. The seller is always harmed by resale. Moreover, we provide conditions such that consumers are also harmed by resale, and we derive the optimal pricing and level of production for the seller when resale in the presence of non-market clearing prices cannot be avoided. Extending the model to allow for vertically differentiated goods, we show that a non-concave revenue function can give rise to goods of different qualities being bunched and sold at a uniform price and to underpricing and hence random rationing.

Keywords: events industry, ticket pricing, secondary markets, underpricing, conflation
JEL-Classification: C72, D47, D82

---

*We are grateful for conversations with and comments by Gabriel Carroll, Laura Doval, Scott Kominers, Leslie Marx, and Cédric Wasser and for feedback from seminar audiences at the University of Melbourne, the 7th Workshop on the Economic Analysis of Institutions at Xiamen University, the Queensland University of Technology, Stanford University, and the Australian National University. The paper has also benefited from discussion with the first year PhD students at University of Melbourne. Financial support through a Visiting Research Scholar grant from the Faculty of Business and Economics at the University of Melbourne and from the Samuel and June Hordern Endowment is also gratefully acknowledged.

†Department of Economics, Level 4, FBE Building, 111 Barry Street, University of Melbourne, Victoria 3010, Australia. Email: simonl@unimelb.edu.au.

‡Department of Economics, Stanford University. Email: evmuir@stanford.edu

1 Introduction

Would a profit-maximizing seller ever deliberately set non-market clearing prices? Would the same seller then be concerned about resale that reduces the inefficiency resulting from the random allocation and try to prevent it? Why, economic reasoning and intuition suggest, would the seller not set market clearing prices in the first place? Thereby it could both preempt the emergence of a resale market *and* make a larger profit because it sells at a higher price.

Yet, this is precisely what happens, for example, in the events industry. There, tickets are regularly sold at a menu of prices which induce excess demand, random rationing and scope for brokers and speculators to profit from resale, much to the chagrin of the events organizers who dislike the ensuing resale and try to prevent it. As noted by Becker (1991), this poses no small conundrum. Perhaps the sellers are not profit-maximizing? Or they care about bringing in low-income audiences, which improves the ambience and thereby increases the willingness-to-pay of high-income customers? Maybe the sellers are afraid of jacking up prices for fear of looking too greedy, or they genuinely care for lower income, lower-willingness-to-pay customers? Of course, it could also be that sellers imperfectly observe demand prior to committing to a price and have an interest in ensuring that the event is sold out, for example, because the artists that play (and perhaps the audience as well) have a preference for sold-out events. Possibilities of plausible explanations that go beyond simple, some might say simplistic, economic theory obviously abound.

A similarly puzzling feature, also frequently observed in the event industry, is that goods of vertically different qualities are bunched together and sold at a single price. For example, seats in a sports stadium are often sold in coarse tiers, with seats in the same price category exhibiting considerable differences in their qualities. As a case in point, the more than fourteen thousand seats at Rod Laver Arena at the Australian Open or sold in four (or, including court side seating, five) different categories only. This raises the question as to why the seller does not use a finer price schedule and a less coarse categorization of seats. Again, there is an abundance of hypotheses that can be put forth to explain this seemingly stark departure from optimality, perhaps the most popular one being that some kind of transaction costs prevent the seller from creating and managing different price and seating categories.

In this paper, we provide a different explanation. We show that standard theory has got it exactly right and that the seemingly compelling economic logic invoked in the introductory paragraph is simply wrong. Beyond private information, no additional transaction costs are invoked to explain coarse pricing. In particular, we show that setting non-market clearing prices and prohibiting resale is part of the optimal strategy for a monopoly under standard assumptions when one does not restrict the monopoly to set market clearing prices and one does not assume that revenue under market clearing pricing is a concave function of quantity. All our other assumptions are standard in monopoly pricing. We show that increasing marginal costs and non-concave revenue are necessary and sufficient to make rationing part of the uniquely optimal strategy for the monopoly.[1] Of course, strictly (and steeply) increasing marginal costs are the appropriate assumption for organizers of entertainment events because there typically a fixed capacity is being allocated.

In a nutshell, non-market clearing pricing and random rationing allow the monopoly to serve low-value consumers whose marginal revenue is high with the same probability as it serves higher-value consumers whose marginal revenue is low. Non-market clearing pricing and random rationing render the monopoly's revenue function concave in situations where under market clearing pricing it is not. This theory brings to light not only an explanation for consistently observed phenomena but also a novel source of inefficiency of monopoly behaviour—random allocations. According to to this theory, the reason why prices are not market clearing is not that the monopoly cares for consumers with lower willingness to pay, but simply that it maximizes profits.

In practice, a monopolist can implement this scheme by first selling the tickets at a high price before having a sale during which the remaining tickets are rationed at a low price. Alternatively, a monopolist can simultaneously sell "premium" and "mass-market" tickets. The premium tickets are sold at a high price and need not be differentiated from the mass market tickets in any other way. Our model thus captures situations in which the primary

[1] If marginal costs are constant, then rationing can be part of the optimal strategy but when it is it is neither uniquely optimal nor generic. With strictly increasing marginal costs, uniqueness and genericity obtain.

motivation for purchasing premium tickets—the very reasons they are premium and higher priced—is to guarantee access by avoiding the lottery associated with the "cheap" mass market tickets.

Importantly, our insights carry over to situations in which vertically differentiated variants of the good are sold at different prices under market clearing pricing (for example, front row or court side seats, and so on). Interestingly, in this case, profit-maximization under a non-concave revenue function may require the seller to lump together goods in vertically differentiated categories into a single category that is sold at a uniform price, which, as mentioned, is also a persistently observed phenomenon in the events industry.

Of course, because rationing is random and inefficient, it offers gains from trade and thereby scope for a resale market and for entry by profit-seeking intermediaries. Not surprisingly, rationing, or "underpricing", goes empirically hand in hand with resale. As resale transaction prices are regularly observed that are far above the initial sale price (or face value) of a ticket, such resale transactions raise, also not surprisingly, the ire of the initial sellers and the fundamental question of why rationing that gives rise to resale can possibly be in the interest of the seller. As Bhave and Budish (2018) put it (emphasis in the original), "the true puzzle is the *combination* of low prices and rent seeking by speculators due to an active secondary market."

Motivated by this, we also extend our model and analysis to account for the possibility of resale, confirming some of the preceding observations while qualifying others. In particular, we show that resale harms the initial seller, thereby corroborating the negative views regarding resale expressed by initial sellers (see e.g. Miranda, 2016). In light of the multitude of ways in which resale can be modelled, this result is remarkably general. It only requires that the resale market constitute a (Bayes Nash) equilibrium and that this equilibrium is anticipated by the seller and all the agents in the initial allocation process that the seller controls. That resale harms the seller then follows from a revealed preference argument: In the world without resale, it could choose the same allocation probabilities that obtain in the equilibrium with resale. Because it is strictly optimal for him to choose different probabilities than those that obtain with resale, it follows that resale harms the seller.

We also show that resale transaction prices are necessarily above the lower of the initial prices (simply because all initial buyers—including the lottery winners— have values above that price), and we provide simple conditions—random-proposer take-it-or-leave-it offers and a matching probability sufficiently close to zero— for the highest resale transaction prices to exceed the higher of the seller's initial prices. Although, as mentioned, resale harms the seller, if resale is not too effective, the seller is strictly better off by inducing rationing and swallowing the bitter pill of having some resale transactions occurring than by setting a uniform market clearing price. Put differently, it is perfectly consistent with our theory to have, simultaneously, rationing, resale, and sellers' complaining about resale. Moreover, we show that it is possible that resale prohibitions increase total consumer surplus. Because the seller is always harmed by resale, it is thus possible to have social and consumer surplus increasing resale prohibition.

Assuming that resale occurs with some probability and is perfectly efficient when it occurs, we are also able to derive the seller's optimal strategy when the seller anticipates that resale occurs on or off the equilibrium path. Among other things, we show that if resale is perfectly effective it will not occur on the equilibrium path.

The remainder of this paper is organized as follows. Section 2 introduces the setup. We analyze the monopoly problem without resale in Section 3. Resale is analyzed in Section 4. In Section 5, we analyze the extension of the model in which the monopolist offers a menu of vertically differentiated goods such as first-row and second-row seats and the like. Sections 6 and 7 discuss the related literature and contain the conclusions, respectively.

2 Setup

We assume that there is a continuum of consumers each with demand for one unit of the good and denote by $P(Q)$ the willingness to pay of the consumer with Q-th highest valuation. We assume that $P(0)$ is positive and finite, $P(Q)$ is decreasing in Q, and that there is a $\overline{Q} < \infty$ such that $P(\overline{Q}) = 0$. While mechanically the model and many results extend beyond this setup in a straightfoward manner, to fix ideas, we further assume that each buyer's valuation v is an independent draw from an absolutely continuous distribution $F(v)$ with support $[0, P(0)]$ with positive density which we denote by $f(v)$. Letting $\mu = \overline{Q}$ denote the total mass of consumers, we thus have $D(p) = \mu(1 - F(p))$ as the demand function for

$p \in [0, P(0)]$, and the inverse demand function is $P(Q) = F^{-1}(1 - Q/\mu)$ for any $Q \in [0, \overline{Q}]$. Denote by

$$R(Q) = P(Q)Q$$

the revenue of a seller who sells the quantity Q at the market clearing price $P(Q)$.

The standard assumption, which is almost universally maintained in economics, is that R is concave. The typical justification for this assumption, other than it being standard, is that it is deemed an analytic simplification that permits one to focus on the key economic insights without cluttering the analysis with case distinctions and multiplicity of local maxima. We have never seen it justified on the basis of empirical evidence, and we will not impose it. With this in mind, a key take-away from this paper is that the assumption that revenue is concave obscures important economic insights and phenomena.

Our analysis is unaffected if we allow for non-identical distributions, provided the seller cannot distinguish consumers. All subsequent arguments then apply to a representative consumer whose value is drawn from the aggregate distribution. Note that even if each consumers draws their value from a distribution that gives rise to a concave revenue function, if these distributions have non-identical supports then the aggregate distribution does not necessarily give rise to a concave revenue function. A formal argument illustrating this point is provided in Appendix A.1.

3 Optimal rationing

We now analyze the optimal selling mechanism and determine when rationing is optimal. Throughout this section we will maintain the assumption that even when there is rationing, there is no resale. Resale is analyzed in Section 4.

3.1 Selling a given quantity optimally

To accommodate the possibility of non-market clearing pricing and rationing, we assume that the monopoly sets two prices, denoted p_1 and p_2, satisfying $p_1 \geq p_2$, such that consumers who buy at price p_1 are served with probability 1 while consumers who opt to buy at price p_2 are served with strictly lower probability if $p_2 < p_1$. This is not only a simple way of incorporating the possibility of non-market clearing prices but as we shall see also without loss of generality. Let Q_1 be the mass of buyers who buy at price p_1, Q_2 be all the buyers who are willing to buy at price p_2 (which includes those who buy at p_1), and Q be the quantity the monopoly sells. Making the participation constraint for the marginal consumer bind, we have

$$p_2 = P(Q_2),$$

or equivalently $Q_2 = D(p_2)$. The incentive compatibility constraint for the consumer with value $P(Q_1)$ who is indifferent between buying at the high price and being served with probability 1 and participating in the random rationing lottery, where the price is $p_2 = P(Q_2)$, is

$$P(Q_1) - p_1 = \frac{Q - Q_1}{Q_2 - Q_1}(P(Q_1) - P(Q_2)),$$

where $Q_2 - Q_1$ is the mass of consumers participating in this lottery and $Q - Q_1$ is the supply allocated to these consumers. Solving for p_1 yields

$$p_1(Q, Q_1, Q_2) = \frac{Q_2 - Q}{Q_2 - Q_1}P(Q_1) + \frac{Q - Q_1}{Q_2 - Q_1}P(Q_2). \quad (1)$$

Henceforth, we shall refer to such a selling mechanism as a *lottery mechanism*. We will refer to selling the quantity Q at the market clearing price $P(Q)$ as a *posted price mechanism*. Figure 1 illustrates the equilibrium construction for a lottery mechanism.

The revenue of a firm who sells the quantity Q_1 at price p_1 and the quantity $Q - Q_1$ at price $P(Q_2)$ with $Q_1 \leq Q \leq Q_2$ is

$$Q_1 p_1(Q, Q_1, Q_2) + (Q - Q_1)P(Q_2).$$

Substituting in the expression for $p_1(Q, Q_1, Q_2)$ and simplifying reveals that this revenue is simply

$$R_\alpha(Q_1, Q_2) := \alpha R(Q_1) + (1 - \alpha)R(Q_2),$$

where $\alpha = \frac{Q_2 - Q}{Q_2 - Q_1}$. That is, when choosing Q_1 and Q_2, the monopoly obtains a convex combination of the revenue it would get if it only sold Q_1 at the market clearing price $P(Q_1)$ and the revenue it would get if it sold Q_2 at $P(Q_2)$.

Intuitively, rationing—that is, choosing $Q_1 < Q < Q_2$—will pay off only if $p_1 > P(Q)$ for otherwise it would mean selling *all* units at or below the market clearing price. Obviously,

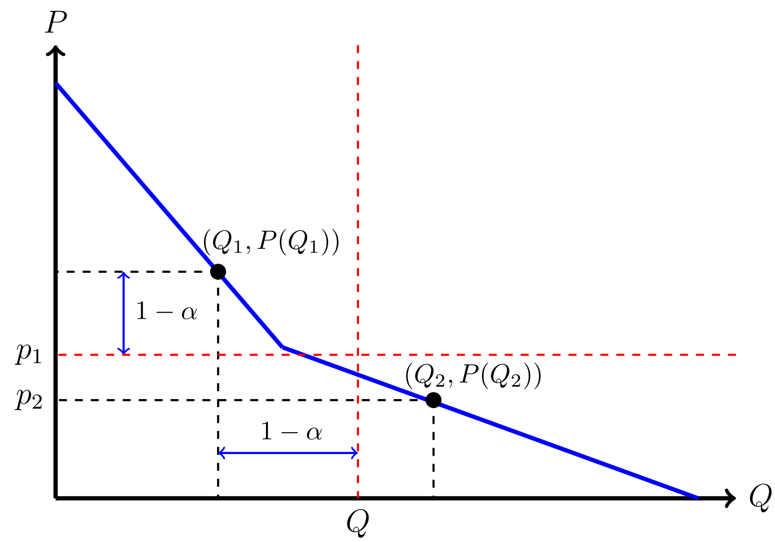

Figure 1: The equilibrium construction, where the probability of winning the lottery is denoted by $1-\alpha$. The position of Q within the interval $[Q_1, Q_2]$ is such that the distance $Q - Q_1$ is proportional to $1-\alpha$. Similarly, the position of p_1 within the interval $[p_2, P(Q_1)]$ is such that $P(Q_1) - p_1$ is proportional to $1-\alpha$.

this is dominated by selling all units at the market clearing price. To see that $p_1 > P(Q)$ is indeed the case, recall that selling Q units using rationing pays off if and only if

$$R(Q) < \alpha R(Q_1) + (1-\alpha)R(Q_2).$$

Substituting for the revenue function and dividing by Q yields

$$P(Q) < \overbrace{\frac{Q_1}{Q}}^{<1} \alpha P(Q_1) + \overbrace{\frac{Q_2}{Q}}^{>1} (1-\alpha)P(Q_2) < \alpha P(Q_1) + (1-\alpha)P(Q_2) = p_1.$$

Before we can provide a formal result that characterizes when lottery mechanisms outperform posted price mechanisms, we need to introduce some terminology. We say that R is *concave at* Q if for any $t \in (0,1)$ and any Q_1, Q_2 such that (i) $Q_1 < Q < Q_2$ and (ii) $Q = tQ_1 + (1-t)Q_2$, we have $R(Q) \geq tR(Q_1) + (1-t)R(Q_2)$. In other words, letting $\overline{R}$ denote the convex hull of R, R is concave at Q if $\overline{R}(Q) = R(Q)$. Otherwise, we say that R is *convex at* Q. From our previous expression for revenue, it follows immediately that there is no point in choosing $Q_i \neq Q$ for $i = 1,2$ if R is concave at Q because then R is everywhere above the line segment connecting any two points on R. Conversely, and by the same argument, choosing $Q_1 < Q < Q_2$ is beneficial whenever R is convex at Q. We thus have the following proposition.

Proposition 1. *Given a quantity Q, the monopoly strictly prefers a lottery mechanism to a posted price mechanism if and only if R is convex at the point Q. Moreover, revenue under the optimal lottery mechanism is given by $\overline{R}(Q)$.*

Combining the mechanism design arguments developed by Myerson (1981) with the equivalence of monopoly pricing problems and optimal auctions first observed by Bulow and Roberts (1989), we obtain an even stronger result, which we state in Theorem 1 below. Here as elsewhere, we say that a mechanism is optimal if it is the profit-maximizing mechanism for the monopoly subject to agents' incentive compatibility and individual rationality constraints. Theorem 1 implies that our restriction to selling mechanisms with at most two prices is without loss of generality because whenever the monopoly prefers price posting to a lottery (or a lottery to price posting), its preferred mechanism is actually the best mechanism available among all incentive compatible and individually rational mechanisms

Theorem 1. *Given a quantity Q, a lottery mechanism is optimal if and only if R is convex at the point Q. Otherwise, a posted price mechanism is optimal.*

Combining Proposition 1 and Theorem 1 we have that revenue under the optimal mechanism is given by $\overline{R}(Q)$. This concavification procedure is equivalent to *ironing* the marginal revenue function (see e.g. Myerson, 1981). As an illustration, consider the inverse demand function

$$P(Q) = \begin{cases} 10 - Q & Q \in [0, 6] \\ 5.5 - Q/4 & Q \in (6, 32], \end{cases} \tag{2}$$

which we will use as a leading example. This function has a kink at $Q = 6$.[2] Consequently, the revenue function $R(Q)$ is not globally concave. Figure 2 illustrates the revenue function as well as its convex hull and the ironed marginal revenue function is shown in Figure 3. Although our leading example features a kink point this is of course not necessary for there to be a region in which lotteries are optimal. As Proposition 1 shows, this is determined by the curvature of the revenue function.

[2] It can (but of course need not) be thought of as arising from the integration of two markets, call them A and B. In each market, demand is linear with $P_A(Q) = 10 - Q$ and $P_B(Q) = 5.5 - Q/4$. Consequently, revenue in each stand alone market is concave. Nevertheless, revenue in the integrated market is not.

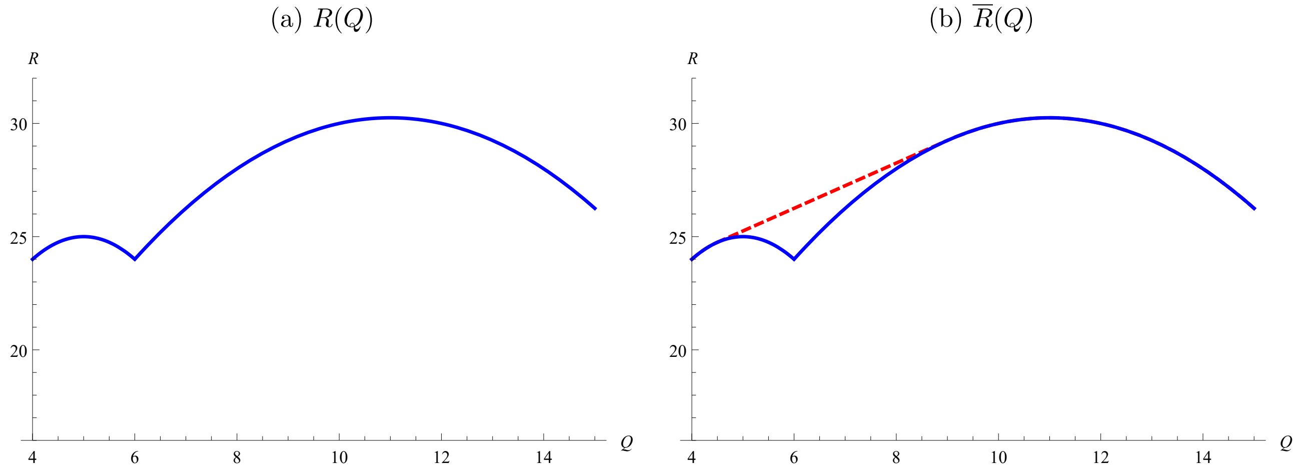

Figure 2: Panel (a): The function $R(Q)$ is not concave. Panel (b): The function $\overline{R}(Q)$ (dashed) is.

We conclude this section by explicitly showing how the optimal selling mechanism can be computed when the revenue function R has two local maxima as is the case for our leading example and Q lies in the convex interval between the local maxima. For $\alpha \in (0,1)$, which must be the case under a non-degenerate lottery mechanism, the first-order conditions for $\max_{Q_1,Q_2} R_\alpha(Q_1,Q_2)$ can be written as[3]

$$R'(Q_1) = \frac{R(Q_2) - R(Q_1)}{Q_2 - Q_1} = R'(Q_2). \tag{3}$$

Observe that (3) can never be satisfied for $Q_2 > Q_1$ if R is a strictly concave function since this implies $R'(Q_2) < R'(Q_1)$. However, since R is convex at Q, the revenue when selling the quantity Q using the optimal lottery mechanism is

$$R_{\alpha^*}(Q_1^*, Q_2^*) = R(Q_1^*) + (Q - Q_1^*)\frac{R(Q_2^*) - R(Q_1^*)}{Q_2^* - Q_1^*} > R(Q),$$

where Q_1^* and Q_2^* solve (3) and $\alpha^* = \frac{Q_2^* - Q}{Q_2^* - Q_1^*}$. This shows that a lottery mechanism strictly outperforms price posting. Evaluated at a point where the first-order conditions are satisfied,

[3]Making use of the facts that $\frac{\partial \alpha}{\partial Q_1} = \frac{\alpha}{Q_2 - Q_1}$ and $\frac{\partial \alpha}{\partial Q_2} = \frac{1-\alpha}{Q_2 - Q_1}$, the first-order conditions for $\max_{Q_1,Q_2} R_\alpha(Q_1,Q_2)$ can be written as

$$\alpha\left[R'(Q_1) + \frac{R(Q_1) - R(Q_2)}{Q_2 - Q_1}\right] = 0 = \left[R'(Q_2) + \frac{R(Q_1) - R(Q_2)}{Q_2 - Q_1}\right](1-\alpha).$$

When $\alpha \in (0,1)$, the last equation can equivalently be written as (3).

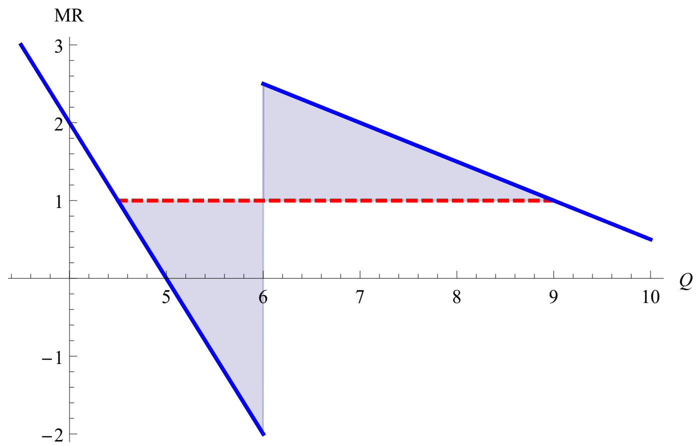

Figure 3: The original marginal revenue curve and the ironed marginal revenue curve (dashed) for our leading example. The first-order conditions in (3) are equivalent to stipulating that the two shaded regions are equal. In this example, $Q_1^* = 9/2$ and $Q_2^* = 9$.

we have

$$\frac{\partial^2 R_{\alpha^*}(Q_1^*, Q_2^*)}{\partial Q_i^2} = R''(Q_i^*) \quad \text{and} \quad \frac{\partial^2 R_{\alpha^*}(Q_1^*, Q_2^*)}{\partial Q_1 \partial Q_2} = 0.$$

So the second-order conditions are satisfied if and only if $R''(Q_i^*) \leq 0$ for $i = 1, 2$. The proof of Proposition 1 shows that Q_1^* and Q_2^* are unique and satisfy $\overline{R}(Q) = \alpha^* R(Q_1^*) + (1-\alpha^*)R(Q_2^*)$, where $\overline{R}$ is the convex hull of R.[4]

3.2 Profit maximization

Recall that given Q, Q_1^* and Q_2^* do not vary with Q and all $Q \in [Q_1^*, Q_2^*]$ give rise to the same Q_1^* and Q_2^*. We then have the following: For any revenue function $R(Q)$ there are finitely many (possibly zero) intervals, indexed by $k \in \{0, 1, \dots\}$, $[Q_1^*(k), Q_2^*(k)]$ such that the maximum revenue for selling Q is

$$\overline{R}(Q) = \begin{cases} R(Q) & Q \notin \cup_k [Q_1(k), Q_2(k)] \\ R(Q_1(k)) + (Q - Q_1(k)) \frac{R(Q_2(k)) - R(Q_1(k))}{Q_2(k) - Q_1(k)}, & Q \in (Q_1(k), Q_2(k)), \end{cases} \tag{4}$$

where $k = 0$ means that $\overline{R}(Q) = R(Q)$ for all Q. By construction, $\overline{R}(Q)$ is continuously differentiable and such that $Q \leq \hat{Q}$ implies $\overline{R}'(Q) \geq \overline{R}'(\hat{Q})$, that is, exhibits weakly decreasing marginal revenue.

[4]In general (that is, if R has more than two local maxima) if an interior solution (Q_1^*, Q_2^*) satisfying the first and second order conditions exists, it is not necessarily unique and the optimal mechanism is pinned down by the formal concavification argument provided in the appendix.

Of course, often sellers choose the quantities they want to sell (and, of course, are typically not required to sell their whole stock as our analysis above stipulated). Trivially, profit maximization requires that the monopoly sells the quantity it produces optimally. Thus, the monopoly's profit maximization problem is

$$\max_{Q} \overline{R}(Q) - C(Q), \tag{5}$$

yielding the usual first-order condition

$$\overline{R}'(Q^*) - C'(Q^*) = 0. \tag{6}$$

If $C'' > 0$, (6) is also sufficient for a maximum. Moreover, if $C'' > 0$ and Q^* is such that $Q^* \in (Q_1(k), Q_2(k))$ for some $k \geq 1$, profit maximization necessarily involves rationing.[5]

Assuming $C'' > 0$ allows us to restrict attention, without loss of generality, to the case where $k = 1$, that is, where there is exactly one interval over which rationing will be optimal. Whether rationing occurs under profit maximization then boils down to whether the intersection of $\overline{R}'(Q)$ and $C'(Q)$ occurs on this interval or not. Assuming $C'' > 0$ (and $k = 1$) allows us also to unambiguously speak of p_1^* and p_2^* as the prices associated with the rationing interval, with p_1^* and p_2^* given by

$$p_1^* = p_1(Q^*, Q_1^*, Q_2^*) \quad \text{and} \quad p_2^* = P(Q_2^*).$$

Summarizing, we have shown the following (up to a few technical details which we relegate to the appendix).

Proposition 2. *The quantity Q^* is the profit-maximizing quantity if and only if $\overline{R}'(Q^*) = C'(Q^*)$. Profit maximization requires rationing if and only if $Q_1^* < Q^* < Q_2^*$.*

Figure 4 illustrates the solution $Q^* \in (Q_1^*, Q_2^*)$ for our leading example when the marginal costs function is $2Q/15$.

For what follows, it is useful to refer to the submarket in which rationing occurs as the *lottery* market. Assume that rationing occurs in equilibrium. After this lottery market

[5] If $C'(Q^*) = \frac{R(Q_2(k)) - R(Q_1(k))}{Q_2(k) - Q_1(k)}$ for some $k \geq 1$ and if $C'' = 0$, then the profit-maximizing quantity is not unique and the profit maximum can be implemented with and without inducing rationing as the monopoly obtains the same profit for all $Q \in [Q_1(k), Q_2(k)]$.

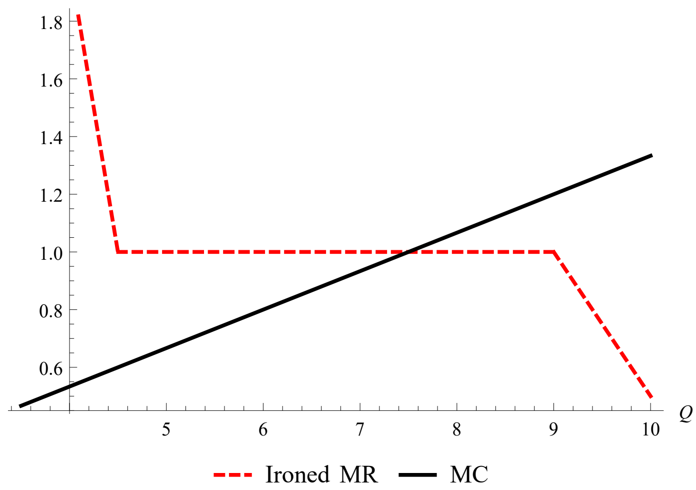

Figure 4: The optimality condition $\overline{R}'(Q^*) = C'(Q^*)$ illustrated for our leading example with $C(Q) = Q^2/15$.

closes, there will be buyers with values above p_1^* who were rationed but now might like to buy in the submarket where Q_1^* units were allocated at the price p_1^*. There are two ways to deal with this. Either one can assume that all buyers with values above $P(Q_1^*)$ immediately buy at p_1^*, so that after the lottery market closes, buyers who were rationed there cannot obtain any additional units at p_1^*. Alternatively, and in line with real-world practice, one can assume that the seller operates the two submarkets sequentially, offering the Q_1^* premium "seats" or tickets at p_1^* first, and then offers to sell the additional units $Q^* - Q_1^*$ at p_2^* only after all Q_1^* units are sold.

Interestingly, this dynamic interpretation and implementation also has a flavor of price discrimination and exploratory pricing: Observing a monopoly selling the quantity Q_1^* at p_1^* immediately and then increasing its quantity supplied to Q^*, with the remaining units $Q^* - Q_1^*$ offered at the price p_2^*, it is natural to think that the monopoly has misjudged demand and now corrects its forecast error by increasing the quantity and reducing price. Alternatively, and equivalently, the initial selling of Q_1^* at p_1^* may be interpreted as being part of an exploratory pricing strategy to gauge demand. However, in our setting, these connections are in appearance only as there is no aggregate uncertainty about demand, and the seller, as everyone else, is fully aware that it will sell the additional units at the price p_2^*

after it has sold all units at p_1^*.

3.3 Consumer preferences over lotteries and price posting

We now discuss how consumers' welfare depends on whether the monopolist uses a lottery or posts a price, keeping the demand function and parts of the cost function fixed as explained below. This is a useful thought experiment in itself. It is further motivated by the effects of resale that a lottery induces, which, as we show in the next section, may well be such that the monopoly prefers to post a price even when, without resale, a lottery would be optimal. For ease of exposition, we assume that the profit-maximization problem under price posting has two local maxima, denoted (Q_L, p_H) and (Q_H, p_L) with $Q_L < Q_H$ and $p_H = P(Q_L) > p_L = P(Q_H)$. For a piecewise linear demand function, Figure 5 provides an illustration of the quantities Q_L, Q_H, Q_1^* and Q_2^*. With strictly increasing marginal costs, we have

$$Q_1^* < Q_L < Q^* < Q_H < Q_2^*.$$

Observe that because of this, we have

$$p_2 = P(Q_2^*) < p_L < p_H.$$

In our thought experiment, we keep the demand function and Q_L and Q_H fixed and assume that marginal costs are strictly increasing but we allow Q^* to vary continuously between Q_1^* and Q_2^*. This corresponds to varying the marginal cost function $C'(Q)$ for $Q \in (Q_1^*, Q_2^*)$ while keeping $C'(Q_1^*)$ and $C'(Q_2^*)$ fixed. Notice that although we know $p_2 < p_L < p_H$, we cannot say in general how p_1 and p_H are ranked.

We first show that there is a potential conflict of interest among different groups of consumers regarding the desirability of lotteries. If (Q_L, p_H) is the global maximum, then all consumers with values $v \in [P(Q_2^*), p_H)$ are worse off with a lottery because they will not be able to purchase a unit of the good when the monopolist posts a price of p_H whereas they have a chance of getting one in the lottery. The welfare implications for consumers with values above p_H depend on the details, in particular because the price p_1 under the lottery mechanism may be higher or lower than the price p_H. Moreover, some of these consumers will be rationed under the lottery mechanism. If the global maximum is (Q_H, p_L), then

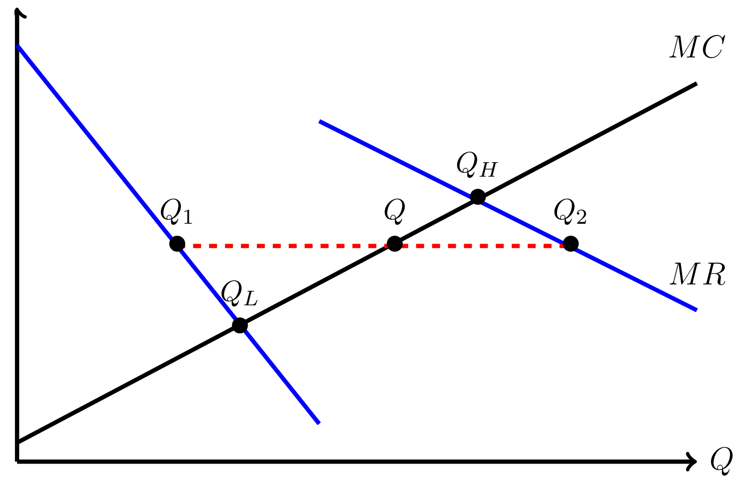

Figure 5: For the marginal revenue and marginal cost curves illustrated here, Q is the quantity sold in the absence of resale. Under a perfectly competitive resale market, the primary market equilibrium would then be either (Q_L, p_H) or (Q_H, p_L).

consumers who participate in the premium market are better off with a lottery since they will receive the good with certainty and pay a lower price under the optimal posted price. The welfare implications for consumers that participate in the mass market under the lottery cannot be determined in general. While these consumers pay a lower price $p_2 < p_L$ under the lottery, fewer units are produced in total and some of these consumers are rationed.

To complete the analysis of the conditions under which lotteries benefit consumers in the sense of increasing consumer surplus, notice that consumer surplus under the lottery that allocates Q in the revenue maximizing way, denoted $CS^L(Q)$, is

$$CS^L(Q) = \int_0^{Q_1^*} P(x)dx + (1-\alpha)\int_{Q_1^*}^{Q_2^*} P(x)dx - R_\alpha(Q_1^*, Q_2^*),$$

whereas consumer surplus under price posting given the quantity Q, denoted $CS^P(Q)$, is standard and given by

$$CS^P(Q) = \int_0^{Q} P(x)dx - R(Q).$$

Observe that, for any $Q \in [Q_1^*, Q_2^*]$,

$$CS^L(Q) \leq CS^P(Q), \tag{7}$$

with equality only if $Q = Q_1^*$ or $Q = Q_2^*$. This follows immediately from the fact that the lottery both allocates inefficiently and generates more revenue for the monopolist. Thus,

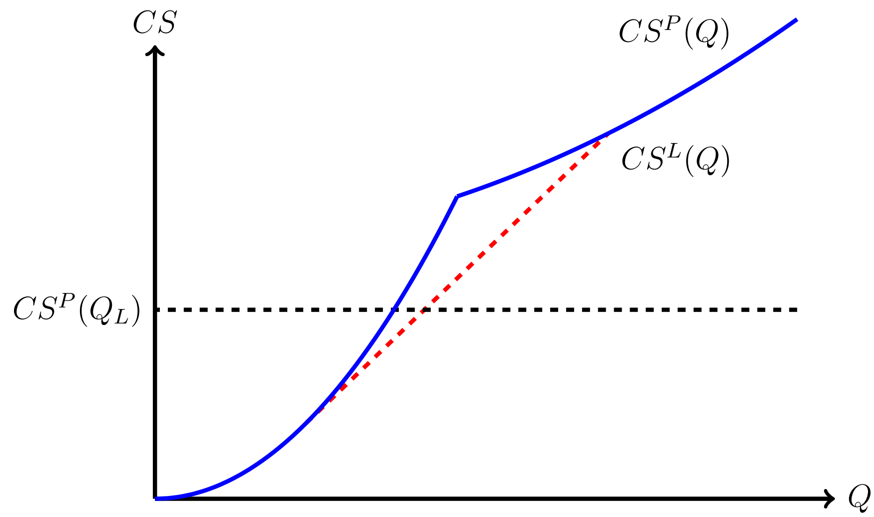

Figure 6: $CS^L(Q)$ (red) and $CS^P(Q)$ (blue) for $Q \in [Q_1^*, Q_2^*]$. The dashed line is $CS^P(Q_L)$.

consumers can benefit from a lottery only if (Q_L, p_H) is the global maximum under price posting. Notice also that because

$$\frac{\partial CS^L(Q)}{\partial Q} = \frac{1}{Q_2^* - Q_1^*}\left[R(Q_1^*) + \int_{Q_1^*}^{Q_2^*} P(x)dx - R(Q_2^*)\right] > 0, \tag{8}$$

there is a unique $\hat{Q} \in (Q_L, Q_2^*)$ such that

$$CS^L(\hat{Q}) = CS^P(Q_L).$$

That $\hat{Q} < Q_2^*$ follows from the facts that $CS^P(Q)$ is increasing in Q and that $CS^L(Q_2^*) = CS^P(Q_2^*)$. This implies the following:

Proposition 3. *Allowing the monopoly to use a lottery increases consumer surplus if and only if $Q^* > \hat{Q}$ and (Q_L, p_H) is the global maximum under price posting.*

Observe that $\hat{Q}$ can be larger than Q_H. In this case, a lottery always harms consumers because $Q^* < Q_H$.

Figure 6 illustrates (7) and Proposition 3. Notice that $CS^L(Q)$ in linear in general. This follows because the derivative in (8) is independent of Q. In other words, the linearity of $\overline{R}(Q)$ translates to $CS^L(Q)$ being linear.[6] For our leading example in (2) and assuming

[6] In contrast, $CS^P(Q)$ need not be convex outside the ironing range, where $R(Q)$ is concave, because $R'' = 2P' + P''Q < 0$ is compatible with $CS^{P''} = -P' - P''Q < 0$ since $P' < 0$.

$C(Q) = Q^2/15$, we have $Q_L = 78/16 \approx 4.69$ as the quantity associated with the global maximum under price posting and $\hat{Q} = 2451/512 \approx 4.79 < Q_H = 165/19 \approx 8.68$. Because $Q^* = 7.5$, $Q^* > \hat{Q}$ follows. Hence, consumer surplus with a lottery exceeds consumer surplus under price posting.

Taken at face value, Proposition 3 may seem to give some justification to the view that event organizers use rationing because they care for consumer surplus. After all, under the conditions stated in the proposition, consumer surplus is higher with a lottery than with a posted price mechanism. However, this alignment between what is good for the consumers and what the monopoly likes is a sheer coincidence. The monopoly does, by our assumptions, not care for consumer surplus. It uses a lottery mechanism because it maximizes profit.

4 Resale

Rationing, or "underpricing", goes hand in hand with resale because the inefficient allocation resulting from rationing opens scope for gains from trade. As mentioned, Bhave and Budish (2018) consider "the combination of low prices and rent seeking by speculators due to an active secondary market" to be the true puzzle in ticket pricing. Resale transaction prices that exceed the initial sale prices ("face values") are consistently observed in the real world and seem, at face value, difficult to reconcile with rational seller behaviour. As outlined in the introduction, while a variety of explanations have previously been put forward to justify systematic ticket "underpricing", it is difficult to explain why monopolists would pursue a pricing strategy that leads to profitable rent-seeking by speculators. Not surprisingly, sellers tend to dislike resale, with some going so far as seeing resale a threat to the existence of the event industry (Miranda, 2016).

There is thus ample motivation to analyze resale in the context of our theory of optimal rationing by a monopoly seller. We now provide such an analysis. We first show that resale market transaction prices that exceed the initial sale prices of the seller can be consistent with the seller exploiting an optimal pricing strategy involving rationing. We also show that, very generally, the seller is harmed by resale. Taken together, this shows that high resale prices can be consistent with the seller optimally inducing rationing (thus providing scope for resale) and with the seller disliking resale (and potentially taking steps to mitigate it).

In particular, provided the resale market is not too efficient, the seller does not dislike resale enough to prevent it from occurring by setting a market clearing price.

For the case in which a perfectly efficient resale market operates with some probability, we then provide a complete characterization of the optimal selling mechanism, showing, among other things, that the ironing parameters Q_1^* and Q_2^* do not vary with this probability, only the prices and hence the ironing marginal revenue. In turn, this implies that the optimal quantity produced varies with the probability with which the resale market operates. We conclude this section with an analysis of the welfare effects of resale prohibition. In particular, we provide conditions such that consumers surplus is larger with resale prohibition than when resale is permitted. Because the seller is better off without resale, this shows that resale prohibition can increase social surplus. That being said, social surplus increasing resale prohibition is a possibility, not a necessity.

4.1 Resale transaction prices

We begin our analysis by stipulating that the size of the resale market is negligible. In particular, we assume that the resale market is only active with probability ρ, and let ρ go towards 0. The benefit of studying this limiting case is that the seller's strategy will be as described in the previous section. Moreover, to make things interesting, we assume in what follows that this problem is such that $Q_1^* < Q^* < Q_2^*$, that is, absent resale rationing is strictly optimal. Recall that this implies

$$p_2^* < p_1^* < P(Q_1^*). \tag{9}$$

If one assumes that the resale market, if it operates, is characterized by random matching between buyers and sellers, with either side given the chance of making a take-it-or-leave-it offer with some probability, then the highest price offer made a by a seller in the resale market is $P(Q_1^*)$, which the buyer with willingness to pay $P(Q_1^*)$ is willing to pay. Thus, $P(Q_1^*)$ is also the highest transaction price in the resale market. Because of the inequalities in (9), this implies that the highest resale transaction price exceeds even the initial sale price in the premium market p_1^*. Note also that in equilibrium any resale transaction price has to exceed p_2^* because any successful buyer in the lottery market has a willingness to pay of

$v \geq P(Q_2^*) = p_2^*$. Thus, regardless of the specifics of the bargaining protocol and matching technology of the resale market, the resale market transaction prices will necessarily exceed the face value of the tickets sold at p_2^*. Moreover, because we let $\rho \to 0$, the seller's strategy of inducing rationing is optimal. Summarizing, we thus have the following proposition, which shows that resale with transaction prices that exceed the seller's initial prices is consistent with the use of a lottery mechanism in the primary market.

Proposition 4. *Resale transaction prices exceed the initial price of p_2^* and may even be larger than the initial price of p_1^*.*

Note that the previous proposition also applies if resale is *unanticipated* by the monopolist.

An interesting question concerns the empirical implications of our theory of optimal rationing by a monopoly seller. For $\rho \to 0$ (or unanticipated resale), a fundamental implication of our theory is that revenue under the optimal lottery mechanism with $Q \in (Q_1^*, Q_2^*)$ is a convex combination of the revenue associated with selling the quantities Q_1^* and Q_2^* at market clearing prices. This is equivalent to asking whether the incentive compatibility constraint

$$p_1^* = \alpha^* P(Q_1^*) + (1 - \alpha^*) p_2^*$$

is satisfied. Thus, if an analyst observes p_1^*, p_2^* and α^*, for example, by observing the premium and lottery prices a seller sets, the quantities sold in the premium and the lottery market and the numbers of users applying for a ticket in the lottery market, observations or estimations of $P(Q_1^*)$ will, via the incentive compatibility constraint, provide a test of the theory. In light of the preceding arguments about resale markets, if one is confident about having identified the upper bound on resale transaction prices, one can use this upper bound as an estimate of $P(Q_1^*)$.[7]

4.2 Seller is harmed by resale

We now turn to the analysis when resale is non-negligible and may occur on the equilibrium path. We begin with a very general result that states that the seller dislikes resale and

[7] Of course, there is some tension here as our theory is exactly correct only for $\rho \to 0$, in which case these observations will be noisy.

then make more specific assumptions to shed light on the seller's optimal strategy when it anticipates resale, on or off the equilibrium path.

To better appreciate both the generality of this result and the power (and, arguably, beauty) of the mechanism design approach used to prove it, denote by $U_B(v) \geq 0$ the expected payoff of a buyer —that is, of an agent who did not obtain an item in the primary market allocation— with value v from participating in the resale market and reconsider the incentive compatibility constraint for the marginal buyer whose value is $P(Q_1)$. Keeping the equilibrium structure and p_1, p_2, Q_1, Q_2, and Q (and hence α) fixed, this constraint becomes

$$P(Q_1) - p_1 = (1-\alpha)(P(Q_2) - p_2) + \alpha U_B(P(Q_1)), \tag{10}$$

where increases in $U_B(P(Q_1))$ can be interpreted as increases in the efficiency of the resale market. Notice that (10) is equivalent to

$$p_1 = \alpha(P(Q_1) - U_B(P(Q_1))) + (1-\alpha)p_2.$$

Thus, keeping everything else fixed, introducing or improving resale will harm the monopoly because it induces downwards pressure on p_1.

However, all else is not equal because resale also affects the participation constraint of the marginal agent with value $P(Q_2)$ who is indifferent between participating and being inactive. Without resale, this constraint binds by setting $p_2 = P(Q_2)$. Denote the expected payoff of a seller, that is, of an agent who obtained an item in the primary market, with value v by $U_S(v)$. Making the participation constraint bind means setting

$$p_2 = P(Q_2) + (1-\alpha)U_S(P(Q_2)).$$

Thus, the price that can be charged to the marginal agent who is indifferent between participating and not increases with the efficiency of resale. Moreover, the fraction $1-\alpha$ of this price increase can be passed on to agents who buy in the premium market because $p_1 = \alpha(P(Q_1) - U_B(P(Q_1))) + (1-\alpha)p_2$ by incentive compatibility. Thus, it seems that the answer as to whether resale benefits or harms the monopoly seller depends on the intricate details of the model, in particular, on the specifics of the resale market. If $(1-\alpha)^2 U_S(P(Q_2))$ is larger than $\alpha U_B(P(Q_1))$, then both p_1 and p_2 increase with resale, which would then imply that the seller must be better off with resale. Because $U_B(P(Q_1))$ and $U_S(P(Q_2))$ depend on the details of how the resale market is modelled as well as on the distribution from which values are drawn, an answer of even moderate generality seems difficult if not elusive.

We are now going to show that this is not the case by proving that the seller is harmed by effective resale without imposing any specific assumptions about how the resale market is modelled.

Our first set of assumptions merely stipulates that the resale market is anticipated by the seller and by the agents and that behavior in the resale market constitutes a (Bayes Nash) equilibrium. The importance of the second assumption is that it allows us to make use of incentive compatibility in the resale market. This implies that agents with higher values must obtain the good in every equilibrium of the resale market with a probability that is at least as high as the probability with which agents with lower values obtain it. In turn, this allows us to invoke the payoff equivalence theorem (see e.g. Myerson, 1981; Börgers, 2015). The payoff equivalence theorem implies that the expected payment the monopoly can extract from an agent with value v is, up to constant, pinned down by the probability with which the agent ultimately obtains the good, irrespective of whether the agent obtains it in the primary or in the secondary market. (Under profit maximization, the constant is pinned down by making the participation constraint bind.)

We say that the resale market is *effective* if the probability distribution of obtaining the good is not uniform across types. (Observe that with the lottery it is uniform; by incentive compatibility, the distribution can thus be only non-uniform if it assigns the good to agents with higher values with higher probability.)

We continue assuming that the problem is such that, without resale, the monopoly chooses rationing, i.e. $Q_1^* < Q^* < Q_2^*$.

Proposition 5. *The monopoly's profit with effective resale is smaller than without it.*

Intuitively, the reason why, absent resale, the monopoly chooses a uniform probability is that it would like to sell to the lower value agents (whose marginal revenue is higher) with higher probability than to the higher value agents (whose marginal revenue is lower) but is prevented from so doing by incentive compatibility: It cannot sell to lower value agents with higher probability than to higher value agents, so the best it can do is to sell to them with equal probability. Effective resale undoes this by shifting probability to higher value agents.

To illustrate both the logic behind Proposition 5 and to pave the way towards our analysis of optimal seller behaviour when resale is anticipated to occur on or off the equilibrium path, we now study the case where the resale market is perfectly competitive (or perfectly efficient) if it operates. Moreover, we assume, for now, that the resale market operates with probability 1, that is, $\rho = 1$. Finally, motivated by our analysis in the previous section, for now we restrict attention to the use of lottery mechanisms in the primary market.

Proposition 6. *Fix any lottery mechanism with Q_1, Q and Q_2 satisfying $Q_1 \leq Q \leq Q_2$. Then the equilibrium price and quantity traded in the secondary market, denoted p^* and q^*, are*

$$p^* = P(Q) \quad \textit{and} \quad q^* = \frac{(Q - Q_1)(Q_2 - Q)}{Q_2 - Q_1}.$$

Corollary 1. *Assume the monopolist faces a perfectly competitive secondary market with probability $\rho = 1$. Then the optimal lottery mechanism reduces to setting $Q_2 = Q$ and any $Q_1 \leq Q$ and $p_2 = P(Q)$ and $p_1 \geq P(Q)$.*

Proposition 6 and Corollary 1 imply that perfectly efficient resale is self-defeating in the sense that the monopoly seller will never choose a pricing strategy such that resale occurs on the equilibrium path.[8]

4.3 Efficient resale with some probability

To analyze resale that occurs on the equilibrium path in more detail, we now assume that the resale market is perfectly efficient when it operates and that it operates with probability $\rho \in [0, 1]$. With probability $1 - \rho$, there is no resale. This means that given $Q_1 \leq Q \leq Q_2$, the probability that an agent of type $v \in [P(Q), P(Q_1)]$ who participates in the lottery market

[8]This is reminiscent of the observation of Loertscher and Niedermayer (2019) that a monopoly platform has an incentive to drive out a competing exchange by using an inefficient mechanism if the competing exchange is "too" efficient. A subtle but important difference is that in our model the monopolist uses an inefficient pricing mechanism—rationing—if there is no competing exchange and an efficient mechanism—a market clearing price—if the secondary market is perfectly effective. In contrast, in Loertscher and Niedermayer (2019) entry by the sufficiently competing exchange is deterred by the use of an inefficient mechanism whereas without entry deterrence the pricing mechanism is efficient and consists of posted prices.

ends up with the good is

$$q^\rho(v) = \rho + (1-\rho)\frac{Q - Q_1}{Q_2 - Q_1} = \rho\alpha + (1-\alpha) \geq 1 - \alpha,$$

with strict inequality if $\rho > 0$. For an agent of type $v < P(Q)$, this probability is $q^\rho(v) = (1-\rho)(1-\alpha) \leq 1 - \alpha$ with strict inequality for $\rho > 0$.[9] Thus, resale shifts the probability distribution away from uniform probabilities for $\rho = 0$ to a distributions that gives higher (lower) weight to buyers with values above (below) $P(Q)$. Intuitively, because the buyers with the higher values have lower marginal revenue, resale will harm the seller.[10]

Making the individual rationality constraint bind for the type with value $P(Q_2)$ we have $(1-\rho)P(Q_2) + \rho P(Q) - p_2 = 0$, which implies

$$p_2 = (1-\rho)P(Q_2) + \rho P(Q). \tag{11}$$

The right-hand side of (11) captures winning the lottery and retaining the ticket with probability $1-\rho$ and selling the ticket in the resale market for a price of $P(Q)$ with probability ρ.

Making the incentive compatibility constraint for the type with value $v = P(Q_1)$ bind, and denoting all relevant variables by superscript ρ to indicate their dependence on the resale probability, we obtain

$$P(Q_1) - p_1^\rho = (1-\alpha)(P(Q_1) - p_2) + \alpha\rho(P(Q_1) - P(Q)). \tag{12}$$

[9] The latter is slightly less relevant as the incentive compatibility constraints will not bind for buyers with values below $P(Q)$.

[10] At this stage we are still implicitly assuming that even with resale the seller is restricted to using two prices only. Without resale, two prices are without loss of generality because the seller wants to keep the allocation probability for the higher types in the lottery market as small as possible, which by incentive compatibility then implies that this probability is uniform across agents in the lottery market. This logic in fact extends to the model with perfect resale occurring with probability ρ. To see the idea, consider a direct mechanism and assume that the designer (or seller) allocates the good to agents of type v with probability $g(v)$. Incentive compatibility in the designer's mechanism requires that $g(v)$ is non-decreasing in v. If the designer allocates the total quantity Q, the total probability that an agent of type v obtains the good is $q(v) = g(v) - \rho(1 - g(v))$ if $v \geq P(Q)$ and $q(v) = g(v) - \rho(1 - g(v))$ if $v < P(Q)$. These $q(v)$ are what matters for the payment it can extract from the agents. If the designer chose $g(v)$ to be non-uniform across agents for whom $g(v) \in (0, 1)$, this would imply $q(v) > \lambda + \rho(1-\lambda)$ for $v \geq P(Q)$ and $q(v) < \lambda - \rho(1-\lambda)$ for $v < P(Q)$, where λ is the probability resulting from the optimal lottery with uniform probability (and two prices). But this would mean that high types with low marginal revenue (low types with high marginal revenue) receive the good with higher (lower) probability than necessary. This result is formally stated in Proposition 7.

The first term on the right-hand side of (12) is associated with winning the lottery and paying p_2 and the second term comes from losing the lottery then participating in the resale market and paying $P(Q)$. Using (11) and rearranging gives

$$p_1^\rho = \alpha(1-\rho)P(Q_1) + \rho P(Q) + (1-\alpha)(1-\rho)P(Q_2).$$

Notice that since the price in the resale market is always $P(Q)$, p_1^ρ is now a convex combination of p_1^0 and $P(Q)$, with the weight on $P(Q)$ equal to the probability that the resale market operates.[11] Observe also that $p_1^1 = P(Q)$ and $p_2^1 = P(Q)$. Computing revenue for the monopolist, we then have

$$R^\rho(Q, Q_1, Q_2) = p_1^\rho Q_1 + p_2^\rho(Q - Q_1) = (1-\rho)R_\alpha(Q_1, Q_2) + \rho R(Q). \tag{13}$$

As we should, we obtain as special cases $R^0(Q, Q_1, Q_2) = R_\alpha(Q_1, Q_2)$ and $R^1(Q, Q_1, Q_2) = R(Q)$, matching the expressions computed in the proof of Corollary 1. Let $\overline{R}^\rho(Q)$ be the maximum revenue when the resale market operates with probability ρ. For arbitrary $\rho \in [0, 1]$, denote the maximizers of $R^\rho(Q, Q_1, Q_2)$ over (Q_1, Q_2) by $Q_i^*(\rho)$ for $i = 1, 2$. Then we have the following:

Proposition 7. *For any $\rho \in [0, 1]$ we have $\overline{R}^\rho(Q) = (1-\rho)\overline{R}(Q) + \rho R(Q)$. Thus, for any $\rho' > \rho \geq 0$ and any Q, $\overline{R}^{\rho'}(Q) \leq \overline{R}^\rho(Q)$. Moreover, for any $Q \in [Q_1^*(0), Q_2^*(0)]$ (i.e. any Q such that the monopolist optimally uses a lottery mechanism in the absence of resale) we have $Q_1^*(\rho) = Q_1^*(0)$ and $Q_2^*(\rho) = Q_2^*(0)$. Finally, for any $\rho \in [0, 1]$, the focus on lotteries that involve only two prices is without loss of generality.*

An illustration of the results of Proposition 7 for our leading example can be found in Figure 7. Here we see that as ρ increases from 0 to 1, the envelope of $\overline{R}^\rho$ of achievable revenue is continuously deformed from the convex hull $\overline{R}$ of the revenue function to the revenue function R itself. Similarly, the marginal revenue curve is continuously deformed from $\overline{R}'$ to R'. In this figure we also see that the lottery mechanism quantities $Q_1^*(\rho)$ and $Q_2^*(\rho)$ do not vary with Q. So as the resale market increases in efficiency, it is the lottery mechanism prices that adjust. Within the ironing range we also see that as the resale market becomes

[11]Recall that p_1^0 is itself a convex combination of $P(Q_1)$ and $P(Q_2)$, with the weight on $P(Q_2)$ equal to the probability of winning the lottery.

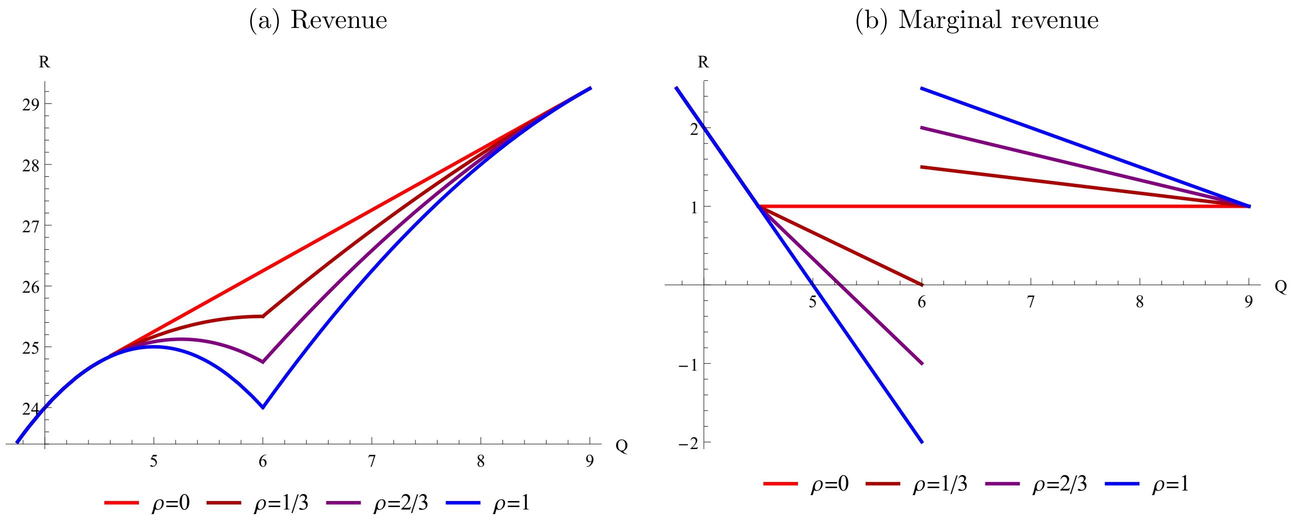

Figure 7: As ρ increases from 0 to 1, the envelope of the revenue function that specifies revenue achievable under a lottery mechanism is deformed from the convex full of the revenue function back to the revenue function itself. As a result, the effective marginal revenue function becomes upward sloping.

more efficient, the *effective* allocation probability for high value, low marginal revenue customers increases. Similarly, the allocation probability for low value, high marginal revenue customers decreases. Although the monopolist would optimally like to induce a uniform allocation probability within the ironing range (and indeed the monopolist still does this in the primary market), the effective allocations probabilities account for the fact that higher value customers are more likely to end up with a ticket after the resale market operates. This impacts the prices the monopolist can charge in the primary market, eroding the revenue of the monopolist.

Finally, assuming that the monopolist's cost function $C(Q)$ is such that $C'(Q)$ is increasing, the optimal quantity Q^* is characterized by

$$(\overline{R}^{\rho})'(Q^*) = C'(Q^*).$$

Refer to Figure 8 for an illustration. For any $\rho > 0$ this equation may have multiple solutions, in which case we need to check which of these corresponds to the global maximum of profit $\overline{R}^{\rho}(Q) - C(Q)$.

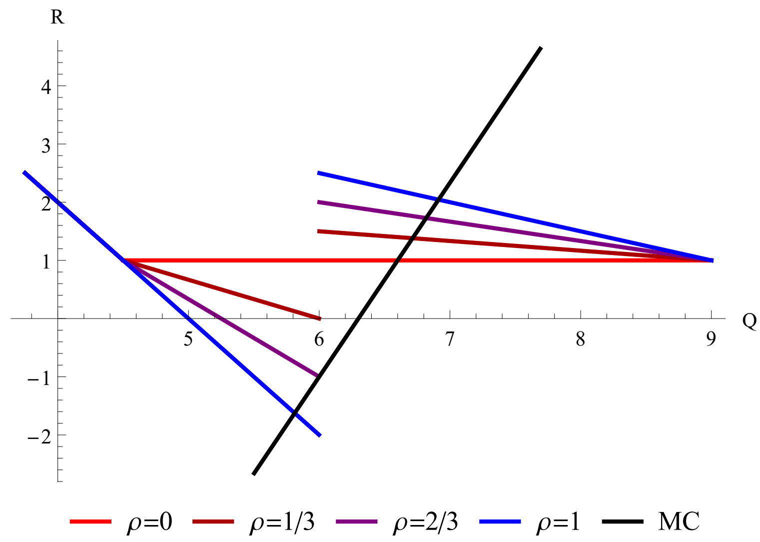

Figure 8: The optimal quantity Q^* can still be determined, for general ρ, as the intersection of marginal revenue and marginal cost.

4.4 Consumer surplus enhancing resale prohibition

We next discuss distributional and welfare effects of resale prohibition under the assumption that without prohibition the secondary market would be perfectly efficient with probability $\rho = 1$ if the seller induces rationing. These assumptions imply that one will never observe a secondary market in operation, with or without resale prohibition. This is obvious when resale is prohibited. Without prohibition, it follows from our observations above that, anticipating a perfectly efficient secondary market if it induces rationing, the monopolist optimally posts a single price and thereby prevents resale from occurring (see Corollary 1). For the purpose of this analysis, we impose the same assumptions as in Subsection 3.3. That is, we assume that the profit-maximization problem under price posting has the two local maxima (Q_L, p_H) and (Q_H, p_L) with $Q_L < Q_H$ and $p_H = P(Q_L) > p_L = P(Q_H)$ as illustrated in Figure 5.

Because resale always harms the monopoly, it is no surprise that the monopoly always benefits from resale prohibition. Interestingly, however, in our model it may well be the case the consumers also benefit from resale prohibition. Our preceding analysis then implies that the monopoly will choose price posting when resale is not prohibited.

Proposition 3 then sheds light on the question when resale prohibition increases consumer surplus as it implies the following corollary:

Corollary 2. *Assume that resale, if not prohibited, is perfectly efficient. Then, consumer surplus is higher when resale is prohibited if and only if $Q^* > \hat{Q}$ and (Q_L, p_H) is the global maximum under price posting.*

Although this may sound counterintuitive at first, the channel through which consumer surplus increasing resale prohibition becomes possible is simple. When resale is efficient, the monopoly will stay clear of rationing (and of opening the scope for resale) and instead choose the profit maximizing posted price-quantity pair. When the quantity under price posting is smaller than under the lottery, the reduction in consumer surplus from the inefficiency of the lottery allocation may be more than offset by the increase in consumer surplus resulting from the fact that a larger quantity is being allocated.[12] For example, for the piecewise linear demand function in (2), consumer surplus is higher under resale prohibition if the monopoly's cost function is $C(Q) = Q^2/15$.

5 Extension: Heterogeneous goods

Up to now, we have assumed a homogenous good. This assumption is useful as it highlights the role of and rationale for rationing when revenue is not concave, but it is, obviously, restrictive. For example, different categories of seats at an event venue may differ in quality such as first row seats that are more prestigious and higher quality than other seats. As mentioned in the introduction, seats of different qualities are often bunched together and sold at a uniform price, for example, at the Australian Open. As we show next, our model sheds new light on this phenomenon as well.

To account for quality differences, we now extend our baseline model by letting, for $i = 1, \ldots, n$, θ_i be the quality level of the good in category i with the θ_i's satisfying $\theta_n > 0$ and, for all $i < n$, $\theta_i > \theta_{i+1}$. The utility of a consumer with value v of obtaining a good in category i is $\theta_i v$. In this extension section, we only consider the problem of optimally

[12] Of course, the revenue they pay is also higher under the lottery, both because the quantity is larger and because the lottery generates more revenue than price posting.

selling, abstracting from the problem of producing the good and the different categories.[13] Let $k_i \geq 0$ be the mass of units available in category i and let $K = \sum_{i=1}^{n} k_i$ be aggregate capacity. As before, we assume that consumers have single-unit demands independently drawn from a continuous distribution F that gives rise to an inverse demand function $P(Q)$ for goods of quality 1, and we denote the revenue of selling Q at the price $P(Q)$ by $R(Q)$. We assume $K < \overline{Q}$, where $P(\overline{Q}) = 0$. Notice that if we normalize $\theta_1 = 1$ and assume $k_i = 0$ for all $i > 1$, this model specialises to the one analyzed in Subsection 3.1.

For $i < n$, letting $\Delta_i := \theta_i - \theta_{i+1}$, the market clearing prices $\boldsymbol{p} = (p_1, \ldots, p_n)$ for selling the total capacity K satisfy $p_n = \theta_n P(K)$, and, for $i < n$,

$$p_i = p_{i+1} + \Delta_i P(K_{(i)}), \tag{14}$$

where $K_{(i)} = \sum_{j=1}^{i} k_j$. Iterative substitution then yields

$$p_i = \theta_n P(Q) + \sum_{j=i}^{n-1} \Delta_j P(K_{(j)}).$$

More generally, the market clearing prices for selling the quantity $Q \leq K$ are

$$p_{m(Q)} = \theta_{m(Q)} P(Q) \quad \text{and, for } i < m(Q), \quad p_i = p_{i+1} + \Delta_i P(K_{(i)}),$$

where $m(Q) \in \{1, \ldots, n\}$ is the index such that $K_{(m(Q)-1)} < Q \leq K_{(m(Q))}$. Iterative substitution then yields

$$p_i = \theta_{m(Q)} P(Q) + \sum_{j=i}^{m(Q)-1} \Delta_j P(K_{(j)}). \tag{15}$$

Putting all of these calculations together, we have the following lemma.

Lemma 1. *Revenue $R^\theta(Q)$ when selling $Q \leq K$ at market clearing prices is given by*

$$R^\theta(Q) = R(Q)\theta_{m(Q)} + \sum_{j=1}^{m(Q)-1} R(K_{(j)})\Delta_j. \tag{16}$$

[13] In some applications, this is a reasonable approximation to the problem sellers; for example, event venues will often have fixed number of prestigious front row seats. At any rate, the assumption highlights the key to the optimality of rationing.

In light of Lemma 1 and our baseline analysis, one might intuitively expect that revenue under the optimal mechanism is given by

$$\overline{R}^{\theta}(Q) = \overline{R}(Q)\theta_{m(Q)} + \sum_{j=1}^{m(Q)-1} \overline{R}(K_{(j)})\Delta_j,$$

that is, the convex hull of $R^{\theta}(Q)$. We will shortly show that this intuition is correct.

Under the class of lottery mechanisms described in Section 3, all lotteries had binary outcomes, with winners receiving a ticket and losers missing out. The natural implementation was to ration losing agents so that they did not make a payment. When tickets are heterogeneous there is scope for the monopolist to construct lotteries with multiple outcomes differentiated by ticket quality. The natural implementation in this case is to think of each lottery as a "category" of uniformly priced tickets that are available for purchase. For example, a monopolist may price tickets by venue section but the quality of a given ticket might actually depend on the row number of the corresponding seat. In principle any category of tickets can also be rationed and for convenience we accommodate this by allowing lotteries to include tickets of quality $\theta_{n+1} = 0$, where $k_{n+1} = \infty$.[14]

Motivated by the previous observations, we now introduce *generalized lottery mechanisms*. Under a generalized lottery mechanism that sells Q tickets, the monopolist offers a collection of ticket categories $\mathcal{I} \subset \mathcal{P}(\{1, \ldots, m(Q), n+1\})$, where $\mathcal{I}$ is subject to three restrictions.[15] First, only tickets of consecutive qualities can be used to create a ticket category.[16] Second, for any ticket category that includes tickets of at least three qualities, tickets that are of one of the interior quality levels cannot be included in any other ticket category.[17] Third, the entire mass of Q tickets must be included in some category. It follows that random allocation (ironing) in the *interior* involves bunching and uniform pricing of different ticket categories, while random rationing only occurs for the lowest quality category (which necessarily includes tickets of quality $m(Q)$). The precise mass of tickets included in each category lottery together with the appropriate incentive constraints then pin down the price

[14]The natural implementation for these lotteries is to first ration an appropriate mass of consumers so that all remaining consumers pay to enter a lottery in which they are guaranteed a ticket.

[15]Here, $\mathcal{P}(X)$ denotes the power set of the set X.

[16]We consider $m(Q)$ and $n+1$ to be consecutive qualities.

[17]For example, if we have a category $I = \{i, i+1, i+2, i+3\}$ then tickets of quality θ_{i+1} and θ_{i+2} cannot be included in another category.

of each ticket category.

It turns out that the optimal selling mechanism is in fact a generalized lottery mechanism and ticket categories that include tickets of more than a single quality correspond to a generalized ironing procedure that is applied to regions where the revenue function is convex. This is stated formally in the following proposition:

Proposition 8. *Revenue under the optimal selling mechanism is given by*

$$\overline{R}^{\theta}(Q) = \overline{R}(Q)\theta_{m(Q)} + \sum_{j=1}^{m(Q)-1} \overline{R}(K_{(j)})\Delta_j.$$

Furthermore, this revenue is achieved by the generalized lottery mechanism.

In principle, the monopolist could decide not to sell the Q highest quality tickets from the mass of K tickets available, However, the previous proposition shows that this is not optimal and from this point forward we can assume, without loss of generality, that $Q = K$ (which in turn implies that $m(Q) = n$).

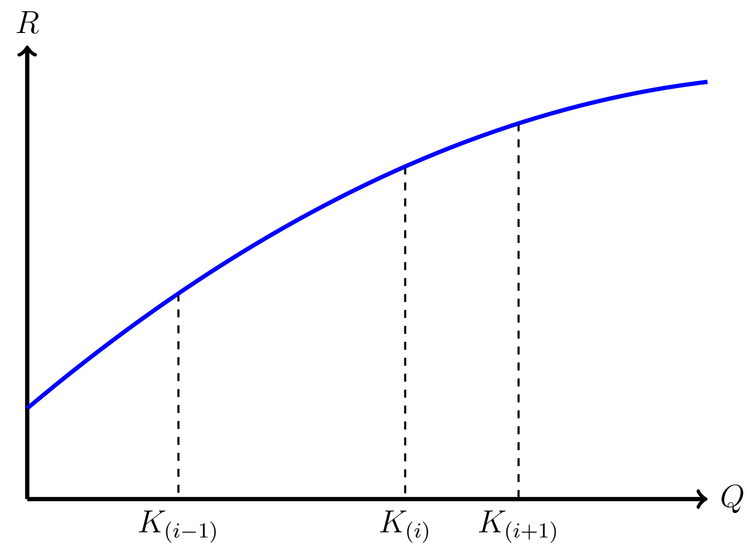

Figure 9: For the $K_{(i)}$ where the revenue function is concave, we use posted prices for the associated categories under the generalized lottery mechanism.

We now provide a description of the optimal generalized lottery mechanism constructed in the proof of Proposition 8. There are three cases to consider. The first case, illustrated in

Figure 9, applies to regions where the revenue function is concave and ticket categories correspond to ticket quality. In particular, for any $i \in \{1, \ldots, n\}$ such that $R(K_{(i)}) = \overline{R}(K_{(i)})$, there exists a stand-alone ticket category $\{i\}$ and customers with $v \in [P(K_{(i)}), P(K_{(i-1)})]$ are allocated a ticket of quality θ_i at the market clearing price.

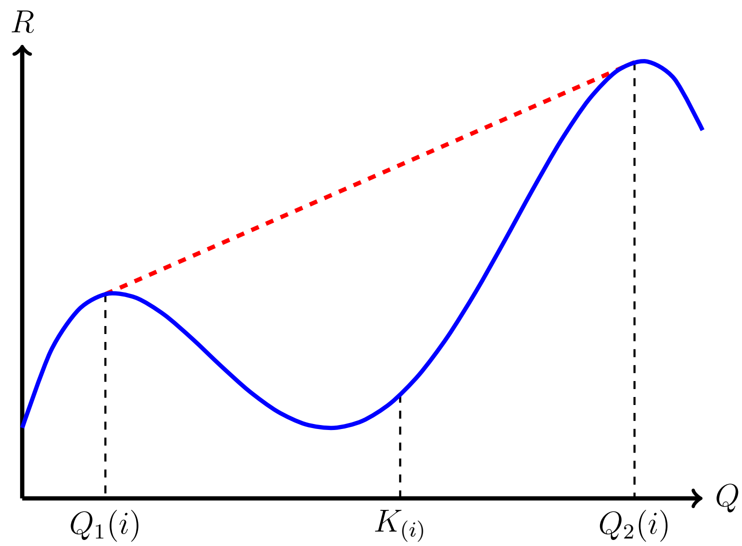

Figure 10: When a single $K_{(i)}$ falls within a convex region, a ticket category $\{i, i+1\}$ is created. The number of ticket categories expands by one relative to the number of quality levels.

In regions where the revenue function is convex, lotteries are required under the optimal mechanism. In particular, for any $i \in \{1, \ldots, n\}$ such that $R(K_{(i)}) < \overline{R}(K_{(i)})$ there exists, by assumption, $Q_1(i)$ and $Q_2(i)$ with $K_{(i)} \in [Q_1(i), Q_2(i)]$ such that

$$\overline{R}(K_{(i)}) = \alpha(i)R(Q_1(i)) + (1 - \alpha(i))R(Q_2(i)),$$

where

$$\alpha(i) = \frac{Q_2(i) - K_{(i)}}{Q_2(i) - Q_1(i)}.$$

The interval $[Q_1(i), Q_2(i)]$ corresponds to the mass of tickets included in a single ticket category. The second case, illustrated in Figure 10, applies to regions where the number of ticket categories expands by one relative to the number of quality levels. Specifically, if $K_{(i-1)} \leq Q_1(i) < Q_2(i) \leq K_{(i+1)}$, the ticket category $\{i, i+1\}$ is created and agents

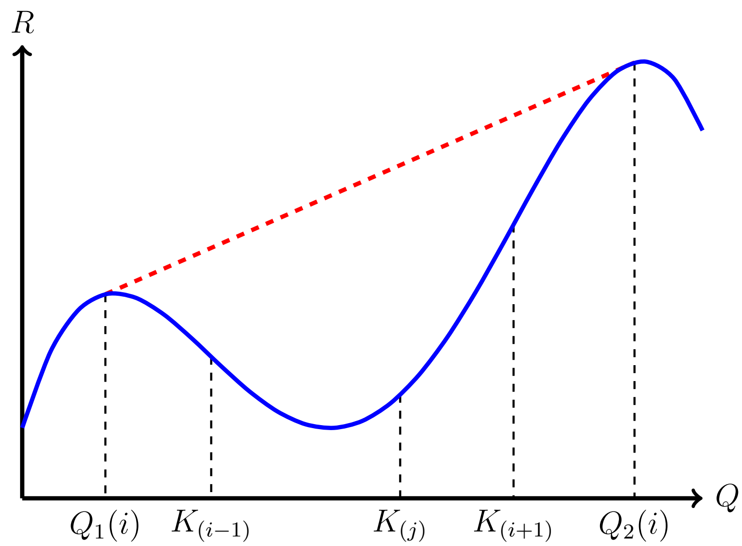

Figure 11: When multiple quantity cutoffs, say $K_{(i-1)}$, $K_{(i)}$ and $K_{(i+1)}$, fall within a single convex region, a ticket category $\{i-1, i, i+1, i+2\}$ is created. The number of ticket categories decreases by one relative to the number of quality levels.

with values $v \in [P(Q_2(i)), P(Q_1(i))]$ enter a lottery and receive a ticket of quality θ_i with probability $1-\alpha(i)$ and θ_{i+1} with probability $\alpha(i)$. The third case, illustrated in Figure 11, applies to regions in which the number of ticket categories weakly contracts relative to the number of quality levels. In particular, if $K_{(i-1)} \leq Q_1(i) < Q_2(i) \leq K_{(i+1)}$ fails to hold then we create a new ticket category $I_j = I \cup \{\max\{I\}+1\}$, where $I = \{\ell \in \{1, \dots, n\} : K_{(\ell)} \in [Q_1(i), Q_2(i)]\}$. If $\ell \in \{\min\{I\}+1, \dots, \max\{I\}-1\}$ then the entire mass k_ℓ of tickets of quality θ_ℓ are included in the associated lottery, along with a mass $K_{(\min\{I_j\})} - Q_1(i)$ of tickets of quality $\theta_{\min\{I\}}$ and a mass $Q_2(i) - K_{(\max\{I_j\})}$ of tickets of quality $\theta_{\max\{I_j\}}$. This completes our description of the allocation rule associated with the optimal selling mechanism. The ticket category prices are then straightforward to compute given the incentive constraints and for the sake of brevity, we defer the interested reader to the proof of Proposition 8.

Interestingly, allowing for heterogenous goods, non-concave revenue and the seller to use an optimal mechanism also provides a solution to the long-standing problem of which goods are to be treated as identical, which is sometimes referred to as *conflation* (see, e.g., Levin and Milgrom, 2010). In our model, conflation is a function of the quality differentials of the various goods available, the quantities in which these are available, and the curvature of the revenue function $R(Q)$.

6 Related literature

There is a large literature on ticket pricing and ticket resale. For an excellent overview, see, for example, Courty (2003a) and the references in Bhave and Budish (2018). Rosen and Rosenfield (1997) analyze ticket pricing from the perspective of second-degree price discrimination while Courty (2003b) introduces uncertainty about demand. Becker (1991) considered the prevalence of non-market clearing pricing in the events industry a major conundrum and provided a theory based on social interactions to explain the phenomenon. As far as we know, the connection to non-monotone marginal revenue that gives rise to optimal rationing (and, from the seller's perspective, optimal prohibition of resale) and that is at the heart of our paper, has not been made in the literature.

Mussa and Rosen (1978) first applied ironing techniques to a non-linear pricing problem and Myerson (1981) introduced the concept of ironing in a mechanism design setup. While the difference between different qualities of goods that is central in Mussa and Rosen (1978) and the probability of being served that is a the center of attention in Myerson (1981) may largely be a matter of interpretation, the quality interpretation may have clouded the view that ironing implies rationing and random allocations. To the best of our knowledge, ours is the first paper that shows how a seller, who is endowed with quantities (or capacities) of vertically differentiated goods, can combine these goods into new quality categories to obtain the convex hull of the revenue function. This problem is absent in the model Mussa and Rosen (1978) analyze because there the seller can produce arbitrary quality levels without any restrictions other than those imposed by the cost function. Put differently, in the heterogenous goods extension of our model the key choice problem of the seller is how to combine and price given sets of goods of given quality. In Mussa and Rosen, this problem is moot because the seller can just choose the desired quality.

As we show, increasing marginal costs are necessary for rationing to be strictly optimal with homogeneous goods. In Myerson (1981), marginal costs are strictly increasing because the seller has an endowment of one unit, which, with more than one buyer, becomes a binding constraint that can be interpreted as marginal costs of infinity at the second unit. In his setting, rationing is strictly optimal when multiple buyers have the same ironed virtual type and this ironed virtual type is the highest among all virtual types. Of course, rationing induces an inefficient allocation, which opens the scope for resale. Interestingly, while resale that arises from the inefficiency in an optimal auction due to discrimination based on virtual types when the buyers draw their values from hetereogenous distributions has been analyzed (see, e.g. Zheng, 2002), ours is, as far as we are aware, the first paper to analyze resale that arises from the inefficiency due to strictly optimal rationing.

Bulow and Roberts (1989) analyze ironing in a monopoly setting but assume constant marginal costs, so that rationing is not required for profit maximization.[18] As just discussed, if the quantity sold is allocated efficiently, there is no scope for resale. Without invoking mechanism design argument, using linear programming with discrete types, Wilson (1988) analyzes monopoly pricing with non-monotone marginal revenue and increasing marginal costs but does not allow for resale. As noted by Bulow and Roberts (1989), the first occurrence of ironing in the context of monopoly pricing is due to Hotelling (1931).

Dworczak et al. (2019) also consider mechanisms that involve two posted prices, where one price involves trade with certainty and the other involves rationing. However, this occurs in a fundamentally different setting and for different reasons. Dworczak et al. (2019) only consider regular environments and focus on efficiency but assume that one unit of the numeraire is not necessarily worth the same to all agents to capture inequality. Rationing arises under efficiency as it provides a means for the designer to redistribute units of the numeraire from "rich" to "poor" agents. Chan and Eyster (2003) is a precursor to that in the following sense. For a college admissions problem, they show that when colleges have, because of affirmative action, a preferences for students with lower scores students over students with intermediate scores, the colleges will use uniform randomization over a certain set of test scores if lower scores are correlated with the group that is favoured by the affirmative action and when the college cannot, because of fairness constraints, admit students with lower scores with higher probability than students with higher scores.

[18] A number of other papers, including Harris and Raviv (1981), Riley and Zeckhauser (1983) and Stokey (1979), that also show the optimality of posted price selling mechanisms, which is sometimes referred to as the "no-haggling" result, assume constant marginal costs (up to maximum demand). Samuelson (1984) is an early example of a paper that finds the optimality of a two-price selling mechanism, which arises in a setting where a buyer and seller have interdependent values.

Che et al. (2013) derive the optimal assignment when agents are budget constrained. They show that under certain conditions, lotteries are optimal, and analyze resale, assuming an otherwise competitive resale market in which the initial seller can levy a tax on transactions. While the empirical implications of their model and ours are similar, the driving forces are different. In our setting, there are no budget constraints, and rationing occurs because the seller maximizes profits. As noted, the fact that, and the reason why, in our model the monopoly prevents resale from occurring on the equilibrium path if the resale market technology is sufficiently effective resembles the observations made by Loertscher and Niedermayer (2019), who show that a monopoly intermediary prevents entry by a competing exchange if and only if the technology of the competing exchange is sufficiently effective.

As far as we are aware, Meng and Tian (2019) provide the first instance of a model in which ironing is, in a sense, non-horizontal. The same phenomenon occurs in our model with imperfect resale, and the reasons are related. Without additional constraints, the designer would like to keep the probabilities uniform across agents. For some reason—resale in our setting, second period allocation and information elicitation in Meng and Tian (2019)— the designer cannot do that and is restricted to award the higher types with higher than uniform probability, which makes the ironing increasing rather than horizontal. Of course, it will be optimal to choose these probabilities as small as possible.

7 Conclusions

Non-market clearing prices that induce excess demand, rationing, and thereby open scope for resale, are a persistent feature of reality but have been deemed puzzling for theory. By charging a higher, market clearing price, it would seem that the seller could kill two birds with one stone—prevent resale *and* generate more revenue. Analyzing an otherwise standard monopoly pricing problem without restricting the seller to set market clearing prices and revenue to be concave, we show that "underpricing" that induces random rationing and opens scope for resale is part of the optimal selling strategy for a monopoly. Rationing is strictly profit maximizing only if marginal costs are strictly increasing. We also show that resale always harms the seller, and that a necessary condition for consumers to be better off with random rationing than with market clearing prices is that, with market clearing pricing, the local maximum characterized by a small quantity and high price is the global maximum. In an extension to heterogenous goods, we show that, in general, non-market clearing prices are still an essential part of the optimal selling mechanism. However, non-market clearing pricing may now take the form of selling goods of different qualities at a uniform price, thereby randomly allocating the goods of heterogenous qualities to the consumers with heterogenous valuations who purchase at the same price.

It does not take much imagination or experience to gather that the mechanism design methodology developed by Roger Myerson was initially met with skepticism and criticism based on the grounds that it was abstract and technical, maybe naturally begging the question of where one observes the designs laid out there. While in the nearly four decades since, partly driven by market design on the Internet, there has been a wide arrange of applications of his methodology, a central piece of this methodology— ironing—has remained somewhat under the radar, still begging the question as to where, if at all, one ever observes this concept in the real world. One message emerging from our paper is that it may have been hidden in plain sight. It explains both underpricing and rationing in ticket pricing and the bunching of tickets of different quality into a single price category. This gives some reason for skepticism towards often heeded calls for "realism". Just because a concept is formulated in the abstract does not mean that it does not have clear-cut counterpart in the real world once one looks closely enough.

There are many avenues for future empirical and theoretical research. For one thing, whether revenue under market clearing pricing is concave is an empirical question. From a theoretical perspective, it would be interesting to extend the monopoly model we have analyzed here to a model of quantity competition in which each firm can decide whether it wants to iron its own residual demand function.

References

Alaei, S., H. Fu, N. Haghpanah, and J. Hartline (2013): "The Simple Economics of Approximately Optimal Auctions," in *54th Annual IEEE Symposium on Foundations of Computer Science*, 628–637.

Becker, G. S. (1991): "A Note on Restaurant Pricing and Other Examples of Social Influences on Price." *Journal of Political Economy*, 99, 1109 – 1116.

Bhave, A. and E. Budish (2018): "Primary-Market Auctions for Event Tickets: Eliminating the Rents of "Bob the Broker"?" Working paper.

Börgers, T. (2015): *An Introduction to the Theory of Mechanism Design*, New York: Oxford University Press.

Bulow, J. and J. Roberts (1989): "The Simple Economics of Optimal Auctions," *Journal of Political Economy*, 97, 1060–1090.

Chan, J. and E. Eyster (2003): "Does Banning Affirmative Action Lower Student Quality?" *American Economic Review*, 93, 858 – 872.

Che, Y.-K., I. Gale, and J. Kim (2013): "Assigning Resources to Budget-Constrained Agents," *Review of Economic Studies*, 80, 73–107.

Courty, P. (2003a): "Some Economics of Ticket Resales," *Journal of Economic Perspectives*, 17, 85–97.

——— (2003b): "Ticket Pricing under Demand Uncertainty," *Journal of Law & Economics*, 46, 627–652.

Dworczak, P., S. D. Kominers, and M. Akbarpour (2019): "Redistribution through Markets," Working paper.

Gershkov, A., J. K. Goeree, A. Kushnir, B. Moldovanu, and X. Shi (2013): "On the Equivalence of Bayesian and Dominant Strategy Implementation," *Econometrica*, 81, 197–220.

Harris, M. and A. Raviv (1981): "A Theory of Monopoly Pricing Schemes with Demand Uncertainty," *American Economic Review*, 71, 347–365.

Hartline, J. (2017): "Mechanism Design and Approximation," Book draft.

Hotelling, H. (1931): "The Economics of Exhaustible Resources," *Journal of Political Economy*, 39, 137–175.

Levin, J. and P. Milgrom (2010): "Online Advertising: Heterogeneity and Conflation in Market Design," *American Economic Review, Papers and Proceedings*, 100, 97–112.

Loertscher, S. and A. Niedermayer (2019): "Entry-Deterring Agency," Working paper.

Meng, D. and G. Tian (2019): "Sequential Nonlinear Pricing of Experience Goods With Network Effects," Working paper.

Miranda, L.-M. (2016): "Stop the Bots from Killing Broadway," *New York Times*.

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

Myerson, R. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58–78.

Riley, J. and R. Zeckhauser (1983): "Optimal Selling Strategies: When to Haggle, When to Hold Firm," *Quarterly Journal of Economics*, 98, 267–289.

Rosen, S. and A. M. Rosenfield (1997): "Ticket Pricing," *The Journal of Law and Economics*, 40, 351 – 376.

Samuelson, W. (1984): "Bargaining under Asymmetric Information," *Econometrica*, 52, 995–1005.

Stokey, N. (1979): "Intertemporal Price Discrimination," *Quarterly Journal of Economics*, 93, 355–371.

Wilson, C. A. (1988): "On the Optimal Pricing Policy of a Monopolist," *Journal of Political Economy*, 96, 1645–176.

Zheng, C. Z. (2002): "Optimal Auction with Resale," *Econometrica*, 70, 2197–2224.

A Proofs

A.1 Non-concave revenue that is a sum of concave revenues

We are now going to show that when market revenue $R(Q)$ arises as the sum of n revenue functions $R_i(Q)$, R is not necessarily globally concave, even if each of the R_i are well-behaved in the sense that they are twice continuously differentiable and concave. Here we focus on the case where the largest willingness to pay $\overline{p}_i := P_i(0)$ differs across the markets, where $P_i(Q)$ is the willingness to pay in market i. We will assume that $\overline{p}_i > \overline{p}_{i+1}$ for all $i \in \{1, \ldots, n-1\}$ and denote by $D_i(p)$ the demand function and by $\tilde{R}_i(p)$ the revenue function, as a function of price, in market i. Let $D(p) = \sum_i D_i(p)$ be the aggregate demand function. Assuming all D_i are decreasing, $D(p)$ is decreasing and hence invertible. Denoting by $P(Q)$ this inverse, $R(Q) = P(Q)Q$ as usual. However, it turns out to be easier work with the functions $\tilde{R}_i(p)$. Total revenue given p is

$$\tilde{R}(p) = \sum_i \tilde{R}_i(p).$$

Wherever $\tilde{R}(p)$ is twice continuously differentiable, which occurs at all p such that all $\tilde{R}_i(p)$ are twice continuously differentiable, we have

$$\tilde{R}''(p) = \sum_i \tilde{R}_i''(p).$$

However, at the $n-1$ points $\overline{p}_2, \ldots, \overline{p}_n$ the revenue function is not differentiable. At every point of non-differentiability $\overline{p}_i$, it satisfies

$$\tilde{R}'(\overline{p}_i + \varepsilon) = \sum_{j=1}^{i-1} \tilde{R}_j'(\overline{p}_i + \varepsilon) > \sum_{j=1}^{i} \tilde{R}_j'(\overline{p}_i + \varepsilon) = \tilde{R}'(\overline{p}_i - \varepsilon)$$

because $\tilde{R}_i'(p)|_{p=\overline{p}_i} = \overline{p}_i D'(\overline{p}_i) < 0$. That is, at every point of non-differentiability, the derivative $\tilde{R}'$ is increasing in p. Thus, $\tilde{R}(p)$ is not globally concave. Because $R(Q) = \tilde{R}(P(Q))$, it follows that $R(Q)$ also fails to be globally concave.

Because $R(Q)$, respectively $\tilde{R}(p)$, only fail to be concave in the neighborhood of points that are not differentiable, and because there are such points if and only if $\overline{p}_i \neq \overline{p}_j$ for some i and j (and analogously, $\underline{p}_i \neq \underline{p}_j$ where $\underline{p}_i$ is such that $D_i(p) = D_i(\underline{p}_i)$ for all $p \leq \underline{p}_i$), it also follows that $R(Q)$ is globally concave if and only if $\overline{p}_i = \overline{p}_j$ and $\underline{p}_i = \underline{p}_j$ for all i and j.

A.2 Proof of Proposition 1 and Theorem 1

To prove Proposition 1 and Theorem 1 we utilize the equivalence of monopoly pricing problems and optimal auction design. While this connection was first observed by Bulow and Roberts (1989), we follow the proof methodology of Alaei et al. (2013).

Proof. For ease of exposition, in this proof we introduce the normalize the mass of consumers to 1 (i.e. set $\mu = 1$), which implies that $Q \in [0, 1]$. As noted by Bulow and Roberts (1989), the monopolist's revenue maximization problem is equivalent to designing an optimal auction when the auctioneer (seller) faces a single buyer with a private value drawn from the distribution F. In what follows, we refer to the problem with a continuum of buyers as the *monopolist's problem* and to the problem in which the designer faces a single buyer as the *auctioneer's problem.*

We first express the monopolist's problem using concepts and results from mechanism design. Specifically, fix Q and let $\langle \boldsymbol{x}, \boldsymbol{t} \rangle$ denote the selling mechanism chosen by the monopolist, where $x(\hat{v})$ and $t(\hat{v})$ respectively denote the probability that a buyer is allocated a unit of the good and the price that buyer pays when the buyer reports to be of type $\hat{v}$.[19] Bayesian incentive compatibility then requires that, for all $v, \hat{v} \in [0, P(0)]$, we have

$$vx(v) - t(v) \geq vx(\hat{v}) - t(\hat{v}).$$

Similarly, individual rationality requires

$$vx(v) - t(v) \geq 0.$$

Finally, feasibility requires

$$\int_0^{P(0)} x(v) f(v) \, dv \leq Q.$$

[19] Here, we are considering a standard mechanism format, where we think of buyers as observing the mechanism $\langle \boldsymbol{x}, \boldsymbol{t} \rangle$ before reporting a type $\hat{v}$ to the monopolist and then receiving a unit of the good with probability $x(\hat{v})$ and paying a price of $t(\hat{v})$. Of course, given such a mechanism there is an equivalent implementation where buyers pay a transfer only upon receiving a unit of the good. Essentially, the monopolist offers a menu of pairs $(x(\hat{v}), t(\hat{v}))$ and since an arriving buyer is free to report any type $\hat{v}$, they can choose any item from this menu.

Following standard mechanism design arguments of Myerson (1981), under any optimal incentive compatible and individual rational mechanism we must have

$$t(v) = vx(v) - \int_0^v x(u)\,du,$$

where $x(v)$ is non-decreasing in v. The revenue of the monopolist under any optimal incentive compatible and individually rational mechanism is then given by

$$\int_0^{P(0)} t(v)\,dv = \int_0^{P(0)} \left(vx(v) - \int_0^v x(u)\,du \right) f(v)\,dv = \int_0^{P(0)} \left(v - \frac{1-F(v)}{f(v)} \right) x(v)f(v)\,dv.$$

Letting $\Phi(v) = v - \frac{1-F(v)}{f(v)}$ denote the virtual value function of Myerson (1981), the problem faced by the monopolist is to maximize

$$\int_0^{P(0)} \Phi(v)x(v)f(v)\,dv \tag{17}$$

subject to the constraint that $x(v) \in [0,1]$ is increasing in v, as well as the feasibility constraint

$$\int_0^{P(0)} x(v)f(v)\,dv \leq Q.$$

The objective (17) is of course the same objective function faced by an auctioneer who sells an object to a buyer with private type v drawn from the distribution F. The monopolist faces an additional feasibility constraint, namely that the object is allocated to the buyer with an ex ante probability of at most Q.[20]

We now solve the monopolists' optimization problem. Since the feasibility constraint restricts the mass of units sold, we will ultimately rewrite the objective function so that the variable of integration is the mass of units sold. First, we proceed by rewriting the objective function in quantile space. In particular, let $\psi(v) = 1 - F(v)$ denote the quantile of the value v (i.e. the mass of consumers with a value of at least v) and let $y(z) = x \circ \psi^{-1}(z)$ denote the quantile allocation rule. Our objective function can then be rewritten

$$\int_0^1 \left(\frac{z}{f(F^{-1}(1-z))} - F^{-1}(1-z) \right) y(z)\,dz = \int_0^1 R'(z)y(z)\,dz,$$

[20] The optimal mechanism for selling Q units also corresponds to the optimal selling mechanism for an unconstrained auctioneer with an appropriately chosen reservation value c.

where $R(z)$ is the revenue generated by selling to all types that fall within the quantile z at the market clearing posted price of $P(z) = F^{-1}(1-z)$. Integration by parts then yields

$$\int_0^1 zF^{-1}(1-z)(-y'(z))\,dz = \int_0^1 R(z)(-y'(z))\,dz.$$

Following the analysis of Alaei et al. (2013) (see also Hartline (2017)), any incentive compatible allocation rule $y(z)$ is non-increasing and can therefore be expressed as a convex combination of reverse Heaviside step functions $H(q-z)$ (where the reverse Heaviside step function $H(q-z)$ corresponds to the allocation induced by a posted price mechanism with price $F^{-1}(1-q)$ and quantity sold q). Therefore, if we fix an allocation rule $y(z)$ and represent it as a convex combination of reverse Heaviside step functions, we can compute revenue by taking the corresponding convex combination of revenues for each associated posted price mechanism. This is precisely how revenue is computed in the last expression for the objective function. It immediately follows that the maximum achievable revenue that can be generated by selling the quantity q is $\overline{R}(q)$, where $\overline{R}$ is the convex hull of R. Changing the variable of integration from quantiles z to quantities q and incorporating the feasibility constraint, we then have that revenue under the optimal mechanism is given by

$$\int_0^1 \overline{R}'(q)H(Q-q)\,dq = \int_0^1 \overline{R}(q)\delta(Q-q)\,dq = \overline{R}(Q),$$

where $\delta(x)$ denotes the Dirac delta function which has a point mass at $x = 0$.[21] The statements of Proposition 1 and Theorem 1 then follow immediately from the fact that whenever Q is such that $\overline{R}(Q) > R(Q)$, $\overline{R}(Q)$ can always be expressed as a convex combination two values (and this convex combination is unique when R has two local maxima). □

A.3 Proof of Proposition 2

Proof. By the proof of Theorem 1, when the monopolist sells the quantity Q using the optimal mechanism, revenue is given by $\overline{R}(Q)$. The monopolist thus seeks to chose the quantity Q in order to maximize profits which are given by $\overline{R}(Q) - C(Q)$. By Alexandrov's

[21] Recall that $H'(x) = \delta(x)$ and that for a sufficiently well-behaved function f we have $\int_{-\infty}^{\infty} f(x)\delta(x)\,dx = f(0)$. Thus, we also see that our last expression for the objective function (which involves the derivative of the allocation rule $y(z)$) is well-defined even if $y(z)$ includes discrete jumps.

theorem $\overline{R}$ is twice differentiable almost everywhere with $\overline{R}'' \leq 0$. The corresponding first-order condition is simply $\overline{R}'(Q^*) = C'(Q^*)$ and $C'' > 0$ is then a sufficient condition for a maximum. □

A.4 Proof of Proposition 4

Proof. Suppose that the lottery market operates (note that this is an off-path event since we are considering the limit as $\rho \to 0$). Then in equilibrium any successful buyer in the lottery market has a willingness to pay of $v \geq P(Q_2^*) = p_2^*$. It immediately follows that all transaction prices in the secondary market will exceed p_2^*. If we assume that the secondary market is a market of perfect information characterized by random matching between buyers and sellers, with either side given the change of making a take-it-or-leave-it offer with some probability, then in equilibrium the highest transaction price in the secondary market is $P(Q_2^*) > p_1^*$. □

A.5 Proof of Proposition 5

Proof. For $\hat{v} \in [P(Q_2^*), P(Q_1^*)]$, let $\rho(\hat{v})$ be the ultimate probability (consisting of the probability of winning in the lottery plus the probability of obtaining the good in the resale market respectively minus the probability of selling it in the resale market) of being allocated a unit of the good when there is resale. Let $\lambda^* = (Q^* - Q_1^*)/(Q_2^* - Q_1^*)$ be the (uniform) probability that an agent of type v obtains the good in the lottery the monopoly induces when there is no resale. We have $\rho(\hat{v}) > \lambda^* > \rho(v)$ for $\hat{v}$ sufficiently high and v sufficiently low, and $\rho(.)$ increasing by incentive compatibility.

The rest of the proof follows from the optimality of the lottery when there is no resale and a revealed preference argument. Specifically, the lottery with allocation probability λ^* and parameters Q_1^* and Q_2^* implements the optimal mechanism derived by Myerson (1981) when N agents draw their values independently from the distribution F that gives rise to $P(Q)$ as N goes to infinity. Because allocating the good to agents with values $v \in [P(Q_2^*), P(Q_1^*)]$ with probability λ^* is strictly optimal in this mechanism and because the allocation rule that allocates the good to agents of these types with probability $\rho(v)$ is admissible in Myerson's problem but not chosen by the designer, it follows that λ^* is strictly revealed preferred to

$\rho(v)$. This implies that the seller is strictly worse off with effective resale. $\square$

A.6 Proof of Proposition 6

Proof. By assumption, the consumers that participate in the lottery are those with values that lie between $P(Q_2)$ and $P(Q_1)$. Since a mass of $Q_2 - Q_1$ consumers participate in the lottery and only $Q - Q_1$ units are allocated under the lottery, the total mass of units that can be supplied in the secondary market is given by $Q - Q_1$ and the maximum quantity demanded in the secondary market is $Q_2 - Q$. It follows that for $q_S \in [0, Q - Q_1]$ and $q_D \in [0, Q_2 - Q]$ the supply and demand schedules are given by

$$P^S(q_S) = P\left(Q_2 - \frac{Q_2 - Q_1}{Q - Q_1} q_S\right) \quad \text{and} \quad P^D(q_D) = P\left(Q_1 + \frac{Q_2 - Q_1}{Q_2 - Q} q_D\right).$$

In a competitive equilibrium in the resale market, we have $q_D = q_S \equiv q^*$ and $P^S(q^*) = P^D(q^*) \equiv p^*$. Because $P^S(q^*) = P^D(q^*)$ is equivalent to

$$Q_2 - \frac{Q_2 - Q_1}{Q - Q_1} q^* = Q_1 + \frac{Q_2 - Q_1}{Q_2 - Q} q^*,$$

we obtain

$$q^* = \frac{(Q - Q_1)(Q_2 - Q)}{Q_2 - Q_1}.$$

Plugging q^* back into $P^S(q^*)$ yields $p^* = P(Q)$. $\square$

A.7 Proof of Corollary 1

Proof. When the resale market is perfectly efficient, the binding incentive compatibility constraint for the consumer with value $v = P(Q_1)$ becomes

$$P(Q_1) - p_1 = (1 - \alpha)(P(Q_1) - P(Q_2)) + \alpha(P(Q_1) - P(Q))$$

which gives us

$$p_1 = (1 - \alpha)P(Q_2) + \alpha P(Q). \tag{18}$$

Revenue for the monopolist is the given by

$$\begin{aligned} R(Q, Q_1, Q_2) &= Q_1[(1 - \alpha)P(Q_2) + \alpha P(Q)] + Q_2 P(Q_2) \\ &= QP(Q_2) - \alpha Q_1(P(Q) - P(Q_2)). \end{aligned}$$

Observe that for any $Q_2 > Q$ and any $Q_1 \in [0, Q]$, we have

$$R(Q, Q_1, Q_2) \leq R(Q, Q_1, Q) = QP(Q) = R(Q).$$

Thus, with perfect resale the optimal "lottery" for the monopoly is degenerate and consists of setting the market clearing price $P(Q)$. (Any $Q_1 \in [0, Q]$ and any $p_1 \in (P(Q), P(Q_1)]$ will be optimal as no one will buy at $p_1 > P(Q)$.) Effectively, the monopoly's profit-maximization problem reduces to the standard case in which a single market clearing price that satisfies $R'(Q) = C'(Q)$ is chosen. (Note that by assumption there will be at least two local maxima that satisfy $R'(Q) = C'(Q)$ and the monopolist will optimally select whichever of these corresponds to the global maximum.) □

A.8 Proof of Proposition 7

Proof. Given any $\rho \in [0, 1]$ we have

$$\begin{aligned}
\overline{R}^{\rho}(Q) &= \max_{Q_1, Q_2} R^{\rho}(Q, Q_1, Q_2) \\
&= \max_{Q_1, Q_2} ((1-\rho) R_{\alpha}(Q_1, Q_2) + \rho R(Q)) \\
&= (1-\rho) \max_{Q_1, Q_2} R_{\alpha}(Q_1, Q_2) + \rho R(Q) \\
&= (1-\rho)\overline{R}(Q) + \rho R(Q),
\end{aligned}$$

where this last line follow immediately from the proof of Proposition 1 and Theorem 1. Since $\overline{R}(Q) \geq R(Q)$ for all Q, the second statement of the proposition follows immediately from the previous expression for $\overline{R}^{\rho}(Q)$. We also immediately have that $Q_1^(\rho) = Q_1^*(0)$ and $Q_2^*(\rho) = Q_2^*(0)$, since the above maximization problem shows that $Q_1^*(\rho)$ and $Q_2^*(\rho)$ are independent of ρ.

It only remains to show that the restriction to two-price lottery mechanism is without loss of generality. The main difficulty here, is that the *effective* values of the customers in the primary market are endogenous to the induced resale market outcome. However, letting ρ and Q be given, then we know that the *effective* distribution of types faced by the monopolist in the primary market is given by $\hat{F}(x) = F(x)$ for $x \geq (1-Q)/\overline{Q}$ and $\hat{F}(x) = (1-\rho)F(x)$ for $x < (1-Q)/\overline{Q}$ (i.e. the effective distribution function $\hat{F}$ has a jump of mass ρ at the value $P(Q)$, reflecting the expected return for these types given that they retain an allocated ticket with probability $1-\rho$ and sell it in the secondary market with probability ρ. We then just apply standard mechanism design arguments (see the proof of Proposition 1 and Theorem 1) to the selling problem with the effective inverse demand function to show that it suffices to restrict attention to two-price lottery mechanisms. effective inverse demand curve $\hat{P}$ to see that it suffices to restrict attention to lottery mechanisms. □

A.9 Proof of Lemma 1

Proof. Starting from

$$R^\theta(Q) = (Q - K_{m(Q)-1})p_{m(Q)} + \sum_{i=1}^{m(Q)-1} k_i p_i$$

and using (15) we have

$$\begin{aligned} R^\theta(Q) &= (Q - K_{m(Q)-1})\theta_{m(Q)}P(Q) + \sum_{i=1}^{m(Q)-1} k_i \left(\theta_{m(Q)}P(Q) + \sum_{j=i}^{m(Q)-1} \Delta_j P(K_{(j)})\right) \\ &= Q\theta_{m(Q)}P(Q) + \sum_{i=1}^{m(Q)-1} k_i \sum_{j=i}^{m(Q)-1} \Delta_j P(K_{(j)}). \end{aligned}$$

Interchanging the order of summation and simplifying then yields

$$\begin{aligned} R^\theta(Q) &= Q\theta_{m(Q)}P(Q) + \sum_{j=1}^{m(Q)-1} \sum_{i=1}^{j} k_i \Delta_j P(K_{(j)}) \\ &= Q\theta_{m(Q)}P(Q) + \sum_{j=1}^{m(Q)-1} K_{(j)}\Delta_j P(K_{(j)}) \\ &= R(Q)\theta_{m(Q)} + \sum_{j=1}^{m(Q)-1} R(K_{(j)})\Delta_j. \end{aligned}$$

□

A.10 Proof of Proposition 8

To prove this proposition, we follow the same methodology used to prove Proposition 1 and Theorem 1.

Proof. For ease of exposition, in this proof we introduce the normalization $\mu = 1$ (i.e. set the mass of consumers to 1), which implies that $Q \in [0, 1]$. In this case, we will see that the monopolist's revenue maximization problem is equivalent to designing an optimal multi-unit auction where the auctioneer (seller) faces a single buyer with a one-dimensional private value drawn from the distribution F (rather than an allocation problem involving heterogeneous goods). In particular, in the multi-unit allocation problem, each additional "unit" allocated to a given agent corresponds to purchasing an additional "unit" of quality. So if an agent purchases i units in the multi-unit allocation problem, this corresponds to purchasing a good of quality θ_{n-i+1} in the original problem.

We first express the monopolist's problem using concepts and results from mechanism design. Specifically, let $\langle \boldsymbol{x}, \boldsymbol{t} \rangle$ denote the selling mechanism chosen by the monopolist. For each possible buyer report $\hat{v} \in [0, P(0)]$, the allocation rule $\boldsymbol{x}(\hat{v}) = (x_1(\hat{v}), \ldots, x_n(\hat{v}))$ encodes a probability distribution over the outcomes $\{1, \ldots, n+1\}$, where outcome $i \in \{1, \ldots, n+1\}$ corresponds to the buyer receiving a good of quality θ_i.[22] For $i \in \{1, \ldots, n+1\}$, $x_i(\hat{v})$ denotes the probability that a buyer that reports to be of type $\hat{v}$ is allocated a good of quality θ_i. Similarly, $t(\hat{v})$ denotes the transfer paid by a buyer that reports to be of type $\hat{v}$. Letting $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$, Bayesian incentive compatibility then requires that, for all $v, \hat{v} \in [0, P(0)]$, we have

$$v(\boldsymbol{\theta} \cdot \boldsymbol{x}(v)) - t(v) \geq v(\boldsymbol{\theta} \cdot \boldsymbol{x}(\hat{v})) - t(\hat{v}).$$

Similarly, individual rationality requires

$$v(\boldsymbol{\theta} \cdot \boldsymbol{x}(v)) - t(v) \geq 0.$$

Finally, feasibility requires that, for all $i \in \{1, \ldots, n\}$,

$$\int_0^{P(0)} x_i(v) f(v)\, dv \leq k_i \quad \text{and} \quad \sum_{i=1}^{n} \int_0^{P(0)} x_i(v) f(v)\, dv \leq Q.$$

Equivalently, letting $X_{(i)}(v) = \sum_{j=1}^{i} x_j(v)$, feasibility requires that, for all $i \in \{1, \ldots, n\}$,

$$\int_0^{P(0)} X_{(i)}(v) f(v)\, dv \leq K_{(i)} \quad \text{and} \quad \int_0^{P(0)} X_{(n)}(v) f(v)\, dv \leq Q.$$

Following standard mechanism design arguments (see, e.g., Myerson (1981)), under any optimal incentive compatible and individual rational mechanism we must have

$$t(v) = v(\boldsymbol{\theta} \cdot \boldsymbol{x}(v)) - \int_0^{v} (\boldsymbol{\theta} \cdot \boldsymbol{x}(u))\, du,$$

[22] Recall that we introduced the convention $\theta_{n+1} = 0$ and $K_{n+1} = \infty$ for convenience.

where $\boldsymbol{\theta} \cdot \boldsymbol{x}(v)$ is non-decreasing in v. The revenue of the monopolist under any optimal incentive compatible and individually rational mechanism is then given by

$$\begin{aligned}\int_0^{P(0)} t(v)\,dv &= \int_0^{P(0)} \left(v(\boldsymbol{\theta} \cdot \boldsymbol{x}(v)) - \int_0^v (\boldsymbol{\theta} \cdot \boldsymbol{x}(u))\,du\right) f(v)\,dv \\ &= \int_0^{P(0)} \left(v - \frac{1-F(v)}{f(v)}\right)(\boldsymbol{\theta} \cdot \boldsymbol{x}(v))f(v)\,dv.\end{aligned}$$

Denoting by $\Phi(v) = v - \frac{1-F(v)}{f(v)}$ the virtual value function, the problem faced by the monopolist is thus to maximize

$$\int_0^{P(0)} \Phi(v)(\boldsymbol{\theta} \cdot \boldsymbol{x}(v))f(v)\,dv, \tag{19}$$

subject to the constraint that $\boldsymbol{\theta} \cdot \boldsymbol{x}(v) \in [0,1]$ is increasing in v, as well as the feasibility requirements that, for all $i \in \{1,\ldots,n\}$,

$$\int_0^{P(0)} X_{(i)}(v)f(v)\,dv \le K_{(i)} \quad \text{and} \quad \int_0^{P(0)} X_{(n)}(v)f(v)\,dv \le Q.$$

Since the feasibility constraints restrict the mass of goods sold for each quality level, as well as the total quantity of goods sold, we will ultimately rewrite the objective function so that the variables of integration are the cumulative mass of goods sold. We proceed by first rewriting the objective function in terms of the cumulative allocation rules $X_{(i)}(v)$. In particular, if we adopt the convenient notation $\Delta_n = \theta_n$, which is natural given the convention $\theta_{n+1} = 0$, then we can rewrite the objective function as follows:

$$\begin{aligned}\int_0^{P(0)} \Phi(v)(\boldsymbol{\theta} \cdot \boldsymbol{x}(v))f(v)\,dv &= \sum_{i=1}^n \int_0^{P(0)} \Phi(v)\theta_i x_i(v)f(v)\,dv \\ &= \sum_{i=1}^n \int_0^{P(0)} \Phi(v)\Delta_i X_{(i)}(v)f(v)\,dv.\end{aligned}$$

This objective function is the same as the objective function faced by an auctioneer designing a multi-unit auction involving a single buyer with private type v drawn from the distribution F.

Next, we rewrite the objective function in quantile space. In particular, let $\psi(v) = 1 - F(v)$ denote the quantile of the value v (i.e. the mass of consumers with a value of at least v) and let $Y_{(i)}(z) = X_{(i)} \circ \psi^{-1}(z)$ denote the ith cumulative quantile allocation rule.

Our objective function can be rewritten

$$\sum_{i=1}^{n}\int_0^1\left(\frac{z}{f(F^{-1}(1-z))}-F^{-1}(1-z)\right)\Delta_i Y_{(i)}(z)\,dz=\sum_{i=1}^{n}\int_0^1 R'(z)\Delta_i Y_{(i)}(z)\,dz,$$

where $\Delta_i R(z)$ is the revenue associated with selling an $(n-i+1)$th unit to all types within the quantile z at the market clearing posted price $\Delta_i P(z)$. Integration by parts yields

$$\sum_{i=1}^{n}\int_0^1 zF^{-1}(1-z)\Delta_i(-Y'_{(i)}(z))\,dz=\sum_{i=1}^{n}\int_0^1 R(z)\Delta_i(-Y'_{(i)}(z))\,dz.$$

Next, by Gershkov et al. (2013), we can restrict attention to allocation rules implementable in dominant strategies without loss of generality. This restriction is useful because any allocation rule implementable in dominant strategies is monotone in the sense that $X_{(i)}(v)$ is increasing in v for all $i \in \{1,\ldots,n\}$.[23] Thus, we can restrict attention to allocation rules such that $Y_{(i)}(z)$ is non-increasing in z for all $i \in \{1,\ldots,n\}$ which, following the analysis of Alaei et al. (2013) (see also Hartline (2017)), can be expressed as a convex combination of reverse Heaviside step functions $H_i(q-z)$.[24] If we fix an allocation rule $Y_{(i)}(z)$ and represent it as a convex combination of reverse Heaviside step functions, we can compute the revenue contribution from allocating an jth unit to some agents by taking the corresponding convex combination of revenue contributions for each associated posted price mechanism. This is precisely how revenue is computed in the last expression for the objective function. It immediately follows that an upper bound on the revenue that can be generated by selling an $(n-i+1)$th unit to a mass of q agents is $\Delta_i\overline{R}(q)$, where $\overline{R}$ is the convex hull of R.[25] Changing the variable of integration from quantiles z to quantities q and incorporating the feasibility constraints for each quality i, an upper bound on the level of revenue that can be

[23] Note that the Bayesian incentive compatibility requirement that $\boldsymbol{\theta}\cdot\boldsymbol{x}(v)$ is increasing in v does not immediately imply that the $X_{(i)}(v)$ are all increasing in v because of the $\boldsymbol{\theta}$ vector weighting.

[24] In this problem the reverse Heaviside step function $H_i(q-z)$ corresponds to the allocation where an $(n-i+1)$th unit is sold to a mass q of agents under the market clearing posted price of $\Delta_i F^{-1}(1-q)$).

[25] At this stage in the proof of Proposition 1 and Theorem 1, we immediately had that this upper bound was achievable (and in particular, achievable using a lottery mechanism). Here, however, we face additional constraints that have not yet been addressed: A jth unit can only be allocated to agents that have already been allocated $j-1$ units. Therefore, if lotteries are involved in the allocation at multiple quality levels, these lotteries may need to be "coordinated" so that we never attempt to randomly allocate a jth unit to an agent that was not randomly allocated a $(j-1)$th unit in a previous lottery. However, we will shortly see that this upper bound is in fact achievable because whenever lotteries are used for adjacent quality levels, the interval of types involved in each lottery is the same. This property allows these lotteries to be coordinated and the aforementioned constraints are satisfied without losing any revenue.

achieved under the optimal mechanism is

$$\sum_{i=1}^{n} \int_0^1 \overline{R}'(q)\Delta_i H_i(K_{(i)} - q)\, dq.$$

Finally, we need to incorporate the constraint that a mass of at most Q units is sold. From the previous expression, we see that it is optimal to sell as many higher quality goods as is feasible, since higher quality goods make a greater revenue contribution. Adopting the notation from Section 5, this means the lowest quality good allocated is $m(Q)$. Therefore, incorporating this last feasibility constraint, we have

$$\begin{aligned}
&\sum_{i=1}^{m(Q)-1} \int_0^1 \overline{R}'(q)\Delta_i H(K_{(i)} - q)\, dq + \int_0^1 \overline{R}'(q)\theta_{m(Q)} H(Q - q)\, dq \\
&= \sum_{i=1}^{m(Q)-1} \int_0^1 \overline{R}(q)\Delta_i \delta(K_{(i)} - q)\, dq + \int_0^1 \overline{R}(q)\theta_{m(Q)} \delta(Q - q)\, dq \\
&= \sum_{i=1}^{m(Q)-1} \overline{R}(K_{(i)})\Delta_i + \overline{R}(Q)\theta_{m(Q)}, \qquad (20)
\end{aligned}$$

where $\delta(x)$ denotes the Dirac delta function which has a point mass at $x = 0$. This last equation is precisely the convex hull of revenue under market clearing posted prices (see (16)).

To complete the argument we describe an allocation that achieves the upper bound in terms of the multi-unit allocation setting. Since the Q highest quality tickets are allocated under this upper bound, without loss of generality we can simplify the exposition by setting $Q = K$ (which implies that $m(Q) = n$). We begin by considering how to allocate all agents their first units. If $R(Q) = \overline{R}(Q)$, these units are simply allocated to all agents with $v \geq P(Q)$. If $R(Q) < \overline{R}(Q)$, then there exists $Q_1(n)$ and $Q_2(n)$ with $Q \in [Q_1(n), Q_2(n)]$ such that

$$\overline{R}(Q) = \alpha(n)R(Q_1(n)) + (1 - \alpha(n))R(Q_2(n)),$$

where

$$\alpha(n) = \frac{Q_2(n) - Q}{Q_2(n) - Q_1(n)}.$$

Under the optimal allocation all agents with values such that $v \geq P(Q_1(n))$ are then allocated a first unit with certainty, while agents such that $v \in [P(Q_2(n)), P(Q_1(n))]$ are allocated a first unit with probability $1 - \alpha(n)$.

Now consider allocating some agents their second unit. If $R(K_{(n-1)}) < Q_1(n)$ (which holds if $R(K_{(n-1)}) = \overline{R}(K_{(n-1)})$ and may also hold otherwise) then we allocate the second units in the same manner as the first unit. In particular, even if a lottery is involved in the allocation of both first and second units we must have $R(K_{(n-1)}) < \overline{R}(K_{(n-1)})$. Therefore, we do not need to worry about "coordinating" these lotteries (see footnote 25) since any agent that participates in a lottery for the second unit is necessarily allocated a first unit. If $R(K_{(n-1)}) > Q_1(n)$ then we have $R(K_{(n-1)}) < \overline{R}(K_{(n-1)})$, as well as

$$\overline{R}(K_{(n-1)}) = \alpha(n-1)R(Q_1(n)) + (1 - \alpha(n-1))R(Q_1(n)),$$

where

$$\alpha(n-1) = \frac{Q_2(n) - K_{(n-1)}}{Q_2(n) - Q_1(n)}.$$

So under the optimal allocation, agents with $v \geq P(Q_1(n))$ are allocated a second unit with certainty, while agents with $v \in [P(Q_2(n)), P(Q_1(n))]$ must participate in a lottery in which they are allocated two units with probability $1 - \alpha(n-1)$. So under the optimal allocation, agents with values in the interval $[P(Q_2(n)), P(Q_1(n))]$ first participate in a lottery for a first unit, and the successful agents then participate in a lottery for a second unit. From an ex ante perspective, the agents with values within the interval $[P(Q_2(n)), P(Q_1(n))]$ that are allocated two units are selected uniformly at random, which is how we can achieve the upper bound given in (20). Iterating, we proceed in this manner, constructing the optimal allocation unit by unit until the appropriate allocation of the nth units is determined.

This multi-unit allocation rule is isomorphic to the allocation rule of a generalized lottery mechanism, which we now describe together with the ticket category prices. Before proceeding we first make note of a useful property of generalized lottery mechanisms. Specifically, the three restrictions on ticket categories under a generalized lottery mechanism imply that ticket categories have a natural quality ordering: For any categories $I, I' \in \mathcal{I}$ we have $\min\{I'\} \geq \max\{I\}$ or $\min\{I\} \geq \max\{I'\}$. We can thus index ticket category quality by j

in a well-defined manner, where I_j denotes the jth highest quality ticket category. We also introduce the vector $\boldsymbol{y}(I_j) = (y_1(I_j), \ldots, y_n(I_j))$, where $y_\ell(I_j)$ is the probability of receiving a ticket of quality θ_ℓ in the category j lottery and let $p(I_j)$ denote the price of ticket category j.

To describe the generalized lottery mechanism and we proceed by considering three cases. First, for any $i \in \{1, \ldots, n\}$ such that $R(K_{(i)}) = \overline{R}(K_{(i)})$ we introduce a ticket category $\{i\}$. Letting j denote the quality index of category $\{i\}$, agents with values $v \in [P(K_{(i)}), P(K_{(i-1)})]$ are allocated a ticket of quality θ_i and, for sufficiently small $\varepsilon > 0$, agents with values $v \in [P(K_{(i)}) - \varepsilon, P(K_{(i)}))$ are allocated a ticket from category $j+1$. Suppose that $i < n$. Then aside from the knife-edge case $K_{(i)} = Q_1(i+1)$, we have $I_{j+1} = \{i+1\}$ and the price of tickets in category $I_j = \{i\}$ is given by

$$p(I_j) = p(I_{j+1}) + \Delta_i P(K_{(i)}).$$

When $K_{(i)} = Q_1(i+1)$ we have

$$p(I_j) = p(I_{j+1}) + (\theta_i - \boldsymbol{\theta} \cdot \boldsymbol{y}(I_{j+1}))P(K_{(i)}).$$

And for the $i = n$ case, $p(I_j) = \theta_n P(Q)$.

Next, for any $i \in \{1, \ldots, n\}$ such that $R(K_{(i)}) < \overline{R}(K_{(i)})$, a non-trivial ticket category that corresponds to a lottery needs to be created. By assumption, there exists $Q_1(i)$ and $Q_2(i)$ with $K_{(i)} \in [Q_1(i), Q_2(i)]$ such that

$$\overline{R}(K_{(i)}) = \alpha(i)R(Q_1(i)) + (1 - \alpha(i))R(Q_2(i)),$$

where

$$\alpha(i) = \frac{Q_2(i) - K_{(i)}}{Q_2(i) - Q_1(i)}.$$

The interval $[Q_1(i), Q_2(i)]$ then maps to the mass of tickets included in the lottery, as we explicitly describe in the next two cases.

For the second case, we suppose that $R(K_{(i)}) < \overline{R}(K_{(i)})$ and $K_{(i-1)} \leq Q_1(i) < Q_2(i) \leq K_{(i+1)}$ hold. Here, a ticket category $I_j = \{i, i+1\}$ needs to be created, where the corresponding lottery includes a mass of $K_{(i)} - Q_1(i)$ tickets of quality θ_i and $Q_2(i) - K_{(i)}$ tickets of quality θ_{i+1}. Agents with values $v \in [P(Q_2(i)), P(Q_1(i))]$ are allocated a ticket from category I_j. For sufficiently small $\varepsilon > 0$, agents with values $v \in [P(Q_2(i)) - \varepsilon, P(Q_2(i)))$ are allocated a ticket from category I_{j+1} and agents with values $v \in (P(Q_1(i)), P(Q_1(i)) + \varepsilon]$ are allocated a ticket from category I_{j-1}. Suppose that $i < n$. Then aside from the knife-edge cases where $Q_2(i) = K_{(i+1)}$ or $Q_2(i) = Q_1(i+1)$, we have $I_{j+1} = \{i+1\}$ and the price of the category $I_j = \{i, i+1\}$ tickets is

$$p(I_j) = p(I_{j+1}) + (1 - \alpha(i))\Delta_i P(Q_2(i)).$$

If $Q_2(i) = K_{(i+1)}$ or $Q_2(i) = Q_1(i+1)$ we have

$$p(I_j) = p(I_{j+1}) + ((1 - \alpha(i))\theta_i + \alpha(i)\theta_{i-1} - \boldsymbol{\theta} \cdot \boldsymbol{y}(I_{j+1}))P(Q_2(i)).$$

When $i = n$ the natural implementation is for agents to pay only if they aren't rationed and

$$p(I_j) = \theta_n P(Q_2(i)).$$

Finally, to ensure that we have a complete specification of all ticket category prices, we need to price the ticket category containing only tickets of quality θ_i, if it exists. Specifically, aside from the knife-edge case where $Q_1(i) = K_{(i-1)}$ (or $Q_1(i) = 0$ when $i = 1$), we have have $I(j-1) = \{i\}$ with

$$p(I_{j-1}) = p(I_j) + \alpha(i)\Delta_i P(Q_1(i)).$$

The third and final case that needs to be considered is when we still have $R(K_{(i)}) < \overline{R}(K_{(i)})$ but $K_{(i-1)} \leq Q_1(i) < Q_2(i) \leq K_{(i+1)}$ fails to hold. Letting $I = \{\ell \in \{1, \ldots, n\} : K_{(\ell)} \in [Q_1(i), Q_2(i)]\}$, a ticket category $I_j = I \cup \{\max\{I\} + 1\}$ needs to be created. For $\ell \in \{\min\{I_j\} + 1, \ldots, \max\{I_j\} - 1\}$ the entire mass k_ℓ of tickets of quality θ_ℓ are included in the ticket category, along with a mass $K_{(\min\{I_j\})} - Q_1(i)$ of tickets of quality $\theta_{\min\{I\}}$ and a mass $Q_2(i) - K_{(\max\{I_j\})}$ of tickets of quality $\theta_{\max\{I_j\}}$. Suppose that $\max\{I_j\} < n$. Then aside from the knife-edge cases where $Q_2(i) = K_{(\max\{I_j\})}$ or $Q_2(i) = Q_1(\max\{I_j\})$, we have $I_{j+1} = \{\max\{I_j\}\}$ and the price of tickets in category I_j is given by

$$p(I_j) = p(I_{j+1}) + (\boldsymbol{\theta} \cdot \boldsymbol{y}(I_j) - \theta_{\max\{I_j\}})P(Q_2(i)).$$

If $Q_2(i) = K_{(\max\{I_j\})}$ or $Q_2(i) = Q_1(\max\{I_j\})$, this becomes

$$p(I_j) = p(I_{j+1}) + (\boldsymbol{\theta} \cdot (\boldsymbol{y}(I_j) - \boldsymbol{y}(I_{j+1})))P(Q_2(i)).$$

When $i = n$ the natural implementation is for agents to pay only if they aren't rationed. Letting $y_{n+1}(I_j)$ denote the probability of rationing we have

$$p(I_j) = (\boldsymbol{\theta} \cdot \boldsymbol{y}(I_j))P(Q_2(i))/y_{n+1}(I_j).$$

Finally, for completeness, we need to price the ticket category containing only tickets of quality $\theta_{\min\{I_j\}}$, if it exists. Specifically, aside from the knife-edge case where $Q_1(i) = K_{\min\{I_j\}-1}$ (or $Q_1(i) = 0$ when $i = 1$), we have $I_{j-1} = \{\min\{I_j\}\}$ with

$$p(I_{j-1}) = p(I_j) + (\theta_{\min\{I_j\}} - \boldsymbol{\theta} \cdot \boldsymbol{y}(I_j))P(Q_1(i)).$$

□