Optimal Simple Ratings*

Hugo Hopenhayn† **Maryam Saeedi**‡

UCLA Tepper, CMU

September 16, 2022

Abstract

We address an important practical problem related to rating design: If a market designer is restricted to a limited set of ratings, how should these be chosen? How does the choice depend on the objective of the market designer and on market characteristics? What is the value loss from using simple ratings? We provide a characterization of optimal ratings, which for our baseline setting is the solution to a standard clustering problem. Next we show that the value loss due to using simple ratings is small and drops sharply as the number of ratings grows. The methods are illustrated with an application to the Medicare Advantage insurance market and one for eBay.

Keywords: Adverse Selection, Certification, Simple Ratings, Consumer Surplus, Producer Surplus. **JEL codes:** D21, D47, D60, D82, L11

*We thank Marina Halac, Andy Skrzypacz, Ali Shourideh, and Ariel Zetlin-Jones for very valuable comments. We also thank Benjamin Vatter for generously sharing the data related to Medicare Advantage providers with us.

†hopen@econ.ucla.edu

‡msaeedi@andrews.cmu.edu

1 Introduction

Ratings and in particular simple rating mechanisms are used by many online trading platforms as well as many rating agencies. In this paper, we consider the question of optimal design and performance of rating systems when the market designer is constrained to use a small number of ratings. We also compute the welfare loss due to these simple ratings.

In practice, rating systems usually provide coarse signals of quality to buyers. For example, in California, restaurants are given grades A, B, C, or none based on a hygiene inspection score. Airbnb awards its top-quality hosts the Superhost badge, and eBay's high-quality sellers are classified as Top Rated Sellers. These quality ratings tend to use information that is not readily available to users. Many governmental and non-profit agencies certify firms that meet certain standards.[1] These examples raise two critical questions about simple rating design: First, given the number of ratings and our constraints, what are the criteria for the design of rating mechanisms? In particular, when there are only two tiers, how stringent should the standards for certification be?[2] Second, what is the welfare loss from using a coarse rating system instead of using the unconstrained optimal mechanism?

Our baseline model considers a competitive market with a large set of buyers and sellers, though our most important results apply equally to the canonical model of Cournot competition with constant marginal cost. Sellers are endowed with different levels of quality, which is the only source of product differentiation.[3] Our analysis focuses on two main sources of market heterogeneity: the distribution of seller qualities and the responsiveness of sellers' supply to prices. Intuitively, the heterogeneity and skewness of seller quality affect the spread of prices across ratings, while the responsiveness of supply determines the resulting reallocation of output across these categories.

[1]For example, the website ecolabelindex.com currently lists 456 certifiers for food and consumer products across 199 countries and 25 industry sectors. To the best of our knowledge, they all use these simple mechanisms, mostly certifying only a subset of the firms in the sector that meet some minimum requirements. Accessed August 31, 2022.

[2]Hui et al. (forthcoming) examine the effect of an increase in the requirements to become a badged seller on eBay. They find that this increase leads to a higher market share of high-quality sellers while decreasing the sales of sellers in the medium range of quality.

[3]While moral hazard might be a critical consideration in some markets, in others adverse selection might play a more critical role, as suggested by an empirical study using eBay data (see Hui et al. (2018)). Optimal rating design with moral hazard and adverse selection is considered in Saeedi and Shourideh (2022) in a simplified market environment. Shi et al. (2020) and Vatter (2021) also study optimal information disclosure in a model that includes moral hazard.

To the best of our knowledge, this is the first paper that systematically considers the interaction of these two factors and their impact on rating design.

We first derive a necessary condition defining the thresholds that correspond to an intuitive criterion. Consider a marginal seller with quality at the threshold between two adjacent intervals. For this threshold to be optimal, the planner should be indifferent between pooling this seller with those in the intervals above or below. This decision ultimately affects the demand faced by the seller, and thus its total output. The benefit of the increased output is the extra value generated by the additional sales, which at the optimum should equate the extra cost of production. Therefore, one of the key determinants of this trade-off is the supply behavior of sellers, in particular, the curvature of the supply function.

Next, we find a simple characterization for the optimal thresholds in the case of linear supply that also applies to the case of Cournot competition with constant marginal cost. These optimal thresholds are the solution to a standard clustering problem that involves only information regarding the distribution of qualities.[4] The thresholds also provide a useful benchmark for other cases, as discussed in the paper.

Regarding the performance of ratings, we show that a one-threshold partition closes at least half of the surplus gap between the no-information case and full-information case for quality distributions that follow some general conditions, such as log-concave density. In our numerical computations, we find that this partition closes from 46% to nearly 77% of the gap, depending on the underlying distribution of qualities. The loss due to coarse ratings diminishes rapidly as the number of thresholds increases, implying that a simple and cost-effective system with a few tiers can achieve a large part of the full-information disclosure value. To illustrate our methods, we calculate optimal certification thresholds based on data from two markets: eBay sellers and Medicare Advantage providers.

The design of a rating system faces the following challenges, which we address. How strict and how selective should the standards for certification be? And how does their choice depend on the distribution of quality and supply considerations? We find that an increase in right (resp.,

[4] This clustering problem can be solved by the $k-$means algorithm as introduced by MacQueen et al. (1967) and used extensively in machine learning and statistics.

left) skewness reduces (resp., increases) the share of producers with high ratings and increases (resp., decreases) the share of those with lower ratings. Similar considerations apply to the degree of convexity of the supply function. An intuition for this result is that optimal ratings trade off pooling in different regions. Pooling is more costly where there is more quality dispersion or where supply is more responsive to prices.

Related Literature Our paper is related to two strands of literature: first, the papers considering the impact of information disclosure on consumer and producer surplus; second, those concerning both the determinants of coarse rating systems and their performance.[5]

There is a large literature on certification and quality disclosure. Dranove and Jin (2010) provide an excellent survey of the earlier papers. Most of the literature focuses either on the incentives for firms to reveal their information or the incentives of certifiers to do so. The main question in this literature is how much information will be revealed in equilibrium and how this might depend on the nature of competition in the product or certification markets. This paper is a follow-up to Hopenhayn and Saeedi (Forthcoming), where we study the optimal information disclosure without restricting the planner to use a limited number of signals.

Coarse ratings have also been justified in the literature by their simplicity and overall performance. Wilson (1989) shows that losses relative to the full-information case are of order $1/n^2$ for a partition with n classes. This finding is consistent with our computed bounds in Section 4. Our theoretical bound on the gains from a two-tier certification is also related to the bounds found by the coarse matching literature, such as in McAfee (2002), Hoppe et al. (2011), and Shao (2016).

The most relevant empirical papers related to our theory are Vatter (2021), Saeedi (2019), Elfenbein et al. (2015), Fan et al. (2013), and Jin and Leslie (2003). Vatter (2021) studies the market for Medicare Advantage providers and finds optimal simple information disclosure mechanisms that maximize welfare. Saeedi (2019) studies the value of reputation mechanisms and establishes a positive signaling value for the certification done by eBay. Elfenbein et al. (2015) study the value of certification badges across different markets and find that certification provides more value when

[5]Our paper focuses on a setting where uncertainty is about seller quality and information is provided to consumers. There is a growing literature that focuses on the reverse channel, where an intermediary transmits information about buyers to sellers. For a survey see Bergemann and Bonatti (2019).

the number of certified sellers is low and when markets are more competitive. Fan et al. (2013) analyze the effect of badges on Taobao.com and find sellers offer price discounts to move up to the next reputation level. Jin and Leslie (2003) use data on restaurant hygiene ratings to examine the effect of an increase in product quality information to consumers on firms' choices of product quality. Our paper also relates to the literature that analyzes the effects of changes in marketplace feedback mechanisms on price and quality (e.g., Hui et al. (2016), Filippas et al. (2018), and Nosko and Tadelis (2015)). Hui et al. (2020) study the benefit of adding a second tier to the reputation mechanism to mitigate the cold-start problem and to promote entry.

Section 2 describes the model. In Section 3, we find the conditions for the design of optimal simple ratings, and in Section 4, we study the performance of these simple ratings relative to unconstrained optimal information disclosure. In Section 5, we consider other design issues, the role of asymmetries in the distribution of seller quality, consumer vs. producer surplus, and entry considerations. Section 6 concludes the paper. Proofs are relegated to the appendix unless otherwise specified.

2 The Model

This model is based on our previous work Hopenhayn and Saeedi (Forthcoming). There is a unit mass of sellers with qualities z distributed according to a continuous cumulative distribution function (cdf) $F(z)$ on the real numbers. Production technology is the same for all sellers and is given by a differentiable, strictly increasing, and strictly convex cost function of quantity $c(q)$ and, correspondingly, a strictly increasing twice continuously differentiable supply function $S(p)$. On the demand side, there is mass M of consumers who face a discrete choice problem, with preferences

$$U(z,\theta,p) = z + \theta - p,$$

where z is the quality of the good purchased, θ is a taste parameter measuring the preference for goods offered in this market vis-à-vis an outside option, and p is the price of the good. The taste parameter $\theta \geq 0$ is distributed according to a continuous and strictly increasing cdf $\Psi(\theta)$, while the outside good's utility (no purchase) is normalized to zero. Goods are differentiated only by their quality level, which is equally valued by all consumers. Given the linearity of the utility function in z, we can replace a good of quality z with a good of expected quality z and the utility of the consumer stays the same. Throughout the paper, we use z interchangeably as the quality or expected quality of a good.

We assume the following timing: (a) information about seller qualities is provided by the planner, (b) based on this information, consumers form posteriors about each seller's expected quality; (c) given these posteriors, perfectly competitive equilibrium prices are determined as a function of expected quality, considering the supply response of each seller to the corresponding price.[6] We interpret $F(z)$ as the distribution of the posterior means of quality of sellers given all the information the planner has. We assume all market participants have the same posterior information about the expected qualities of sellers after receiving any set of signals from the planner, represented by the distribution function $G(z)$.[7] In particular, in a finite rating system, we assume that G is a discrete distribution with point masses at the conditional mean qualities associated with each rating. We say that a seller has expected quality z if conditional on all signals received, that is the quality expected by all consumers.[8]

Given expected quality z, equilibrium prices arbitrage away the differences in expected quality, taking the form $p(z) = p(0) + z$, where $p(0)$ corresponds to the demand price of a hypothetical good of quality zero. This expression for prices guarantees that consumers are indifferent between goods with different signal realizations with any positive sales, which is a necessary condition for an equilibrium. The baseline price $p(0)$ determines the extent of the market, where the marginal consumer's θ is found by setting $U(0, \theta, p(0)) = 0$, or simply $\theta(p(0)) = p(0)$. All consumers with

[6]While throughout the paper we use the assumption of perfect competition among sellers, our main characterization of an optimal simple rating holds unchanged in the canonical case of Cournot equilibrium with constant marginal cost, as shown in Section 3.2.

[7]This representation of the information structure is consistent with the approach followed in Ganuza and Penalva (2010) and Gentzkow and Kamenica (2016). Given a common prior $F(z_0)$ over seller qualities and a signal structure π, we can let $G(z)$ be the distribution of the expected posterior of seller quality. Any information structure can thus be represented as a garbling of $F(z)$,which corresponds to the finest level of information available to the planner.

[8]If the planner gives all the information to buyers, then expected quality z will be equal to the actual quality of the seller, but when there is some pooling, this expected quality will be a function of other sellers that are pooled with the target seller. For example, a simple certification rating divides sellers into two groups, those certified and those not certified. As a result, there will be two different levels of expected quality and equilibrium prices, one for each group, with possibly many different levels of heterogeneity in quality within each group that are unobserved by buyers.

$\theta \geq p(0)$ will make their unit purchase, so aggregate demand is $Q = M(1 - \Psi(p(0)))$. Inverting this function, we can define an *inverse baseline demand* function,

$$P(Q) = \Psi^{-1}(1 - Q/M) = p(0). \tag{1}$$

On the supply side, each seller with expected quality z chooses its output, $q = S(p(z))$, so aggregate supply $Q = \int S(p(z))\, dG(z)$.

Definition. An (interior) *equilibrium,* given the distribution of expected qualities $G(z)$, is given by prices $p(z) = P(Q) + z$, where total quantity

$$Q = \int S(P(Q) + z)\, dG(z). \tag{2}$$

It is immediate to see that this last equation gives a necessary and sufficient condition for an equilibrium. Moreover, if $P(Q)$ is strictly decreasing and continuous, for any distribution G there will be a unique equilibrium value Q^* and under some regularity conditions, it will be interior.

Using the following assumption, Hopenhayn and Saeedi (Forthcoming) prove that a unique interior equilibrium exists.

Assumption 1. *There exists* $\widetilde{\theta}$ *in the support of* Ψ *such that*

$$M > \int S\left(\widetilde{\theta} + z\right) dG(z)$$

for all distributions G *such that* F *is a mean-preserving spread of* G*. In addition,* $\int S(p(0) + z)\, dG(z) > 0$ *for the same class of distributions.*

The first assumption rules out the possibility that all consumers make purchases in this market; in other words, we assume that the consumers are on the long side of the market.[9] The second assumption rules out no output as an equilibrium. While a corner equilibrium, if it exists, is also unique, we rule this out as a matter of convenience.

[9]As explained below, the assumption spans the set of all possible information structures.

2.1 Optimal Ratings

In order to define optimal ratings, we first derive expressions for consumer and producer surplus. Consider a consumer of type θ who buys a good of expected quality z and therefore receives utility $\theta + z - p(z)$. Given the equilibrium price $p(z) = P(Q) + z$, the consumer's net utility is $\theta - P(Q)$. It follows that total consumer surplus is

$$M \int_{P(Q)} (\theta - P(Q)) \, d\Psi(\theta) = \int_0^Q (P(x) - P(Q)) \, dx,$$

where the equality follows from the change of variables $x = M(1 - \Psi(\theta))$ and our definition of $P(Q)$: $P(Q) = \Psi^{-1}(1 - Q/M)$. This implies that consumer surplus will move in the same direction as market size, as given by total quantity Q. Producer surplus is defined more simply as total profits:

$$\Pi(Q) = \int \pi(P(Q) + z) \, dG(z).$$

A couple of observations are in order. First, notice that for fixed Q, producers benefit from better information. This follows from the convexity of the profit function and the fact that better information is defined by mean preserving spreads. But changes in the information structure might also affect the equilibrium value of total output Q, which is also market size as measured by the number of consumers served. This equilibrium effect has opposite impacts on consumer and producer surplus: while consumers prefer a larger market, producers prefer a smaller one. This countervailing effect of total quantity leads to a conflict of interest between consumers and producers in the design of an optimal rating system, as we analyze in Section 5.2.

With the exception of that section, we take the planner's objective to be the maximization of total surplus, the unweighted sum of consumer and producer surplus. For any information structure as given by the posterior distribution of mean quality G, total surplus is then given by

$$TS(G) = \int \pi(P(Q) + z) \, dG(z) + \int^Q (P(x) - P(Q)) \, dx,$$

where Q is the unique equilibrium output corresponding to G. This equation simplifies to

$$TS(G) = \int^{Q(G)} (P(x))\, dx + \int (z - c(q(z)))(q(z))\, dG(z)\,. \tag{3}$$

An *optimal simple rating* is given by the distribution G that maximizes this objective among a class of distributions defined in the following section.

3 Simple Ratings: Design

Most rating systems are coarse, ranking participants into a small number of categories. For example, in the case of Yelp, the partition involves five stars, including the possibility of half-stars. In the case of California restaurants, the partition involves three elements: A, B, and C. In addition, hundreds of governmental or non-profit certification agencies use a pass-fail or tiered signal for their certification method. In this section, we consider the question of optimal information design when the number of ratings the market designer can employ is limited. This restriction can be motivated not only by its wide use but also by its cost-effectiveness, as giving very precise information might be difficult or costly, and simple rankings might be easier to interpret. Moreover, as we find, most of the gains from optimal information provision can be achieved with a very limited number of ratings, which can make coarse ratings optimal if there is any cost associated with a more detailed rating mechanism.[10]

In this section, we focus on simple ratings that partition the set of sellers into N groups. We assume that consumers have no information other than that provided by the certifier.

Following our earlier discussion on information structures, the certifier's information can be summarized by a distribution of posterior mean qualities that, in order to avoid further notation, we denote by $F(z)$. This is the basis on which the certifier classifies sellers into rating bins. To simplify the exposition, we refer to the expected value z as the quality of the seller. We assume F is differentiable on its support with density $f(z)$.

[10]We do not model this cost component explicitly, but one can include it as part of the modeling assumptions. Including the cost of more precise information into the model requires many modeling assumptions, which we believe would interfere with the main message of the paper.

A *threshold partition* totally orders sellers into N quality intervals. The following lemma proves that threshold partitions are the best among the set of finite partitions.

Lemma 1. *An optimal simple rating is given by a threshold partition.*

Given this lemma, the design of an optimal simple rating system reduces to finding the vector of optimal thresholds, $\boldsymbol{z} = (z_1....z_{N-1})$, that divide sellers into N partitions, $\{[z_0, z_1], [z_1, z_2], ..., [z_{N-1}, z_N]\}$, where z_0 and z_N are the lower and upper supports of the distribution of expected qualities given the planner's information ($-\infty$ or $+\infty$ if unbounded), respectively.[11]

3.1 A Necessary Condition

In this section, we derive a simple and intuitive necessary condition to characterize these optimal thresholds. Let $M_k = m(z_{k-1}, z_k), k = 1, ..., N$ denote the conditional means of sellers' quality within the interval $[z_{k-1}, z_k]$. Let $Q(\boldsymbol{z})$ denote the unique equilibrium total quantity at the optimal threshold vector $\boldsymbol{z} = (z_1....z_{N-1})$. The prices for sellers in partition $[z_{k-1}, z_k]$ are denoted by $p_k = P(Q(\boldsymbol{z})) + M_k$, and quantities, by $q_k = S(p_k)$. From equation (3) it follows that total surplus for partition $\boldsymbol{z}$ is given by

$$W(z) = \int_0^{Q(\boldsymbol{z})} P(x)\,dx + \sum_{k=1}^{N} [F(z_k) - F(z_{k-1})] [M_k q_k - c(q_k)]. \quad (4)$$

Taking first-order conditions with respect to z_k proves the following necessary condition:

Lemma 2. *Let the thresholds $\boldsymbol{z} = (z_1, ..., z_{N-1})$ maximize (3). Then*

$$(P(Q(\boldsymbol{z})) + z_k)(q_{k+1} - q_k) = c(q_{k+1}) - c(q_k) \quad (5)$$

for all z_k.

Condition (5) has an intuitive interpretation. Consider a marginal seller with quality at the threshold between two adjacent intervals. For this threshold to be optimal, the planner should be

[11] Bergemann and Pesendorfer (2007) find a similar result in the context of optimal information design in auctions.

indifferent between pooling this marginal seller with those in the lower or upper interval. The left hand side shows the marginal value obtained by increasing the quantity of the marginal seller with quality z_k, from q_k to q_{k+1}. The right hand side shows the difference in cost. This condition highlights the relevance of the supply behavior of sellers, in particular, the curvature of the supply function, as it impacts both the changes in total output and the production cost.

Figure 1 provides a graphical representation of this necessary condition and its connection to the supply function. Three cases are considered: a linear supply function, given by the solid diagonal line, an upper convex supply (concave marginal cost function), and a lower concave supply (convex marginal cost function).[12] The area below the marginal cost function, i.e., supply function, between q_k and q_{k+1} equals the right hand side of (5), while the area under the line $P + z_k$ equals the left hand side of the equation. The difference between these two areas is

$$\int_{q_k}^{q_{k+1}} \left(P + z_k - C'\left(q\right)\right) dq,$$

which equals zero if and only if condition (5) holds, at the optimal threshold level for z_k. In the linear case, the integrand is positive up to point b (the triangle in blue) and negative thereafter. Point $P + z_k$ is such that the regions from a to b (the triangle in blue) and from b to c (the triangle in red) have the same areas. It is immediate that in the linear case, the corresponding value of $z_k = \left(M_{k+1} + M_k\right)/2$. It also follows easily that for the convex supply case, $P + z_k$ must be higher so the two corresponding areas will have the same area, while the converse holds for the concave supply case.

Proposition 1. *Let the thresholds* $\boldsymbol{z} = \left(z_1, ..., z_{N-1}\right)$ *maximize (3), and denote by* $M_k = E\left(z | z_{k-1} \leq z \leq z_k\right)$ *the corresponding conditional means. Then*

1. $z_k = \left(M_k + M_{k+1}\right)/2$ *if the supply function* $S\left(p\right)$ *is linear;*

2. $z_k > \left(M_k + M_{k+1}\right)/2$ *if the supply function is strictly convex in the interval* $\left[M_k, M_{k+1}\right]$*; and*

3. $z_k < \left(M_k + M_{k+1}\right)/2$ *if the supply function is strictly concave in the interval* $\left[M_k, M_{k+1}\right]$.

[12] We follow the practice of putting price on the y-axis and quantity on the x-axis.

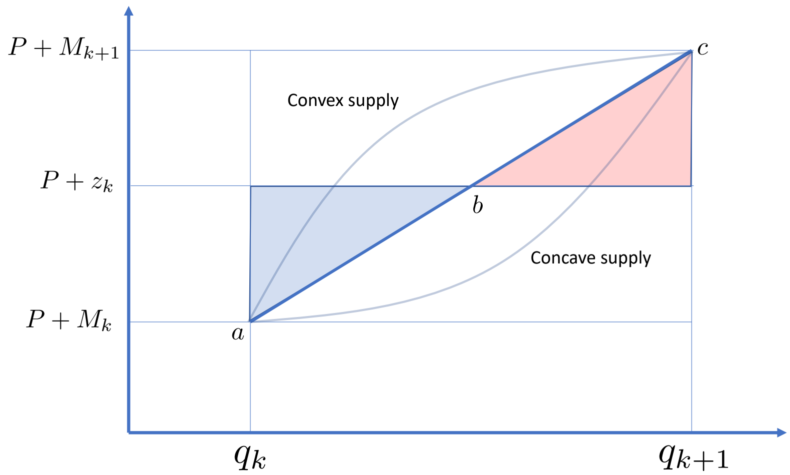

Figure 1: Optimal Pattern

A simple characterization for the solution in the linear supply case and conditions for uniqueness are provided in the following proposition.

Proposition 2. *If the supply function is linear, the optimal thresholds* $\boldsymbol{z} = (z_1, ..., z_N)$ *are the ones that minimize*

$$\sum_{k=1}^{N} \int_{z_{k-1}}^{z_k} (z - M_k)^2 \, dF(z). \tag{6}$$

If in addition F *has log-concave density, the solution to this minimization problem is unique.*

According to this proposition, the optimal thresholds for the linear supply case are the ones that minimize the sum of the variance of qualities within partitions. This objective coincides with the popular $k-$means criteria for clustering as introduced by MacQueen et al. (1967), commonly used in the machine learning and statistics literature. Therefore, estimating the optimal thresholds is a trivial task, because many software programs incorporate algorithms to solve this problem. Additionally, the linear supply function helps us simplify the profit function for the sellers. For a seller with expected quality z, profits are equal to $p(z)^2/2 = (P(Q)+z)^2/2$. Therefore, for any distribution G of observed qualities, total profits are

$$\begin{aligned}\Pi &= \frac{1}{2}\int\left(P\left(Q\right)+z\right)^{2}dG\left(z\right)\\ &= \frac{1}{2}P\left(Q\right)^{2}+P\left(Q\right)\bar{z}+\frac{1}{2}\int z^{2}dG\left(z\right).\end{aligned} \tag{7}$$

In addition, when the supply function is linear, the equilibrium condition 2 reduces to

$$Q=\alpha\int P\left(\left(Q\right)+z\right)dG\left(z\right)=\alpha P\left(Q\right)+\bar{z},$$

so the equilibrium quantity Q does not depend on G and thus on the partition. Consequently, maximizing total profits as given in equation (7) is equivalent to maximizing $\int z^{2}dG\left(z\right)$.

Proposition 1 suggests that the thresholds for the linear case can be lower (resp., upper) bounds for the case of convex (resp., concave) supply. While this proposition gives the criteria for local deviations for a single threshold, convex to the right and concave to the left, starting with those obtained for the linear case, it does not imply an ordering of the whole vector of thresholds. The following proposition gives the conditions for total ordering.

Proposition 3. *Suppose the quality distribution $F\left(z\right)$ has a log-concave density. Let $\left(z_{1}^{L},...,z_{N-1}^{L}\right)$ be the optimal thresholds for the linear supply case. The optimal vector of thresholds $\left(z_{1},...,z_{N-1}\right)$ for a convex (resp., concave) supply function is pointwise higher (resp., lower) than $\left(z_{1}^{L},...,z_{N-1}^{L}\right)$.*

The formula in equation (6) gives a simple characterization for the optimal thresholds in the linear supply case that depends only on the distribution of qualities, and, by the previous proposition, provides a lower (resp., upper) bound when the supply function is convex (resp., concave).

This proposition also suggests that the thresholds for the linear case can provide a good reference point in solving for the optimal thresholds for non-linear supply. In practice, online platforms, such as eBay and Airbnb, often experiment with different certification thresholds. The thresholds for the linear case are very easy to compute and provide a good starting point, while Proposition 3 indicates the direction of improvement.

There is an additional reason why the solution to the linear case is of interest. As we show in the following section, this solution coincides with the optimal thresholds under Cournot competition with constant marginal costs, a workhorse model of imperfect competition. In consequence, the numerical results and applications we provide in Section 4 for the linear supply case also apply to Cournot competition.

3.2 Cournot Competition

Up to here we have assumed that sellers are price takers and are in perfect competition. In this section, we extend our analysis to the case of Cournot competition among sellers with constant marginal cost, c. Our main result is that the optimal thresholds in this setting coincide with those derived above for the perfect competition case with linear supply.

There are in total n sellers in the market. Assuming an interior solution, the first-order condition for a seller of quality level z is given by

$$MR(z) = P'(Q)\,q(z) + P(Q) + z = c. \tag{8}$$

Summing over all sellers gives

$$P'(Q)\,Q + nP(Q) + n\bar{z} = nc, \tag{9}$$

where $\bar{z}$ is the average quality of sellers in the market, which is assumed to be exogenous.[13] Equation (9) determines Q independently of the distribution of z up to its mean, which is thus independent of the information structure.[14] Furthermore, from equation (8) it follows that

$$q(z) = \frac{P(Q) + z - c}{-P'(Q)},$$

[13]The value Q that solves this equation is unique provided marginal revenue is decreasing for the average-sized seller.

[14]This is true for an interior equilibrium where no sellers are excluded from production. A sufficient condition is that $P(Q) + \underline{z} > c$, where Q is the solution to equation (9). .

so it is linear in z, given Q. Seller profits are given by

$$\begin{aligned}\pi(z) &= \left[(P(Q)+z)-c\right]q(z) \\ &= -P'(Q)\,q(z)^2 \\ &= \frac{\left[(P(Q)+z)-c\right]^2}{-P'(Q)} \\ &= \frac{(P(Q)-c)^2+2(P(Q)-c)z+z^2}{-P'(Q)}.\end{aligned}$$

Consumer surplus is independent of the information structure, because it depends only on Q. Thus, maximizing total surplus is equivalent to maximizing total profits. For any information structure $G(z)$, total profits are given by

$$\Pi = \frac{1}{-P'(Q)}\left[(P(Q)-c)^2+2(P(Q)-c)\bar{z}+\int z^2 dG(z)\right].$$

All the terms involving Q and $\bar{z}$ are independent of the information structure G. Thus, total surplus can be written as

$$S_0 + a\int z^2 dG(z) \tag{10}$$

for constants S_0 and a. Hence surplus maximization is equivalent to maximizing $\int z^2 dG(z)$, as in the case of perfect competition with linear supply (see equation 7).[15]

4 Simple Ratings: Performance

In this section, we consider the performance of a simple rating system in the case of perfect competition with linear supply, or, equivalently, Cournot competition with constant marginal cost. We focus on the simplest case of a two-tier rating system that certifies sellers whose expected qualities are above a predetermined threshold. This type of rating is simple to interpret ("certified or not") and widely used in practice; for example, many governmental or non-profit agencies certify a subgroup of sellers while not giving any information about the others. Section 4.1 gives a theoretical

[15]Here we have considered quantity competition. For a model of price competition with partially informed consumers, see Moscarini and Ottaviani (2001).

bound on performance. Section 4.2 provides numerical calculations to gauge the performance for a wide range of distribution functions, including those more widely used in applications. In Section 4.3 we derive the optimal thresholds for two particular scenarios, Medicare Advantage insurers and eBay sellers, based on distributions of qualities identified in other papers.

Total surplus is maximized with full information, as shown in Hopenhayn and Saeedi (Forthcoming). Our measure of performance of a simple rating system is the fraction of the gap between the full-information and no-information cases that is closed with this rating system. Given that in the case of linear supply, or Cournot equilibrium, total quantity and consumer surplus are invariant to the information structure, our performance measure coincides with the fraction of the gap in producer surplus that is covered by the simple rating, using equation 7. Under full information, total profits are $\frac{1}{2}P\left(Q\right)^2 + P\left(Q\right)\bar{z} + \frac{1}{2}\int z^2 dF\left(z\right)$. Therefore, the surplus gap with respect to the full-information case is

$$\Delta\Pi = \frac{1}{2}\left(\int z^2 dF\left(z\right) - \int z^2 dG\left(z\right)\right)$$

for any distribution of a rating system, G, that is a garbling of F.[16] In particular, the maximum surplus gap 11between the full-information and no-information cases is

$$\overline{\Delta\Pi} = \frac{1}{2}\left(\int z^2 dF\left(z\right) - \bar{z}^2\right). \tag{11}$$

For a threshold partition $\left(z_1, ..., z_{N-1}\right)$, G has N mass points at the conditional means $M_1, ..., M_N$; therefore, we can write the surplus gap as

$$\begin{aligned} \Delta\Pi &= \frac{1}{2}\sum_{k=1}^{N}\int_{z_{k-1}}^{z_k}\left(z^2 - M_k^2\right)dF\left(z\right) \\ &= \frac{1}{2}\sum_{k=1}^{N}\int_{z_{k-1}}^{z_k}\left[\left(z - M_k\right)^2\right]dF\left(z\right). \end{aligned} \tag{12}$$

This equation corresponds to the loss function used in $k - means$ clustering, given that at the optimal thresholds, as defined earlier, the expected values M_k are precisely the centroids of the corresponding intervals $\left[z_{k-1}, z_k\right]$. We are interested in seeing how much of the possible total sur-

[16]The total quantity stays the same given the assumption of linear supply.

plus gain from information, as expressed in equation 11, is captured by 12, or, simply, the following ratio:

$$\gamma \equiv \frac{\overline{\Delta\Pi} - \Delta\Pi}{\overline{\Delta\Pi}} = 1 - \frac{\sum_{k=1}^{N} \int_{z_{k-1}}^{z_k} (z - M_k)^2 \, dF(z)}{\int z^2 dF(z) - \bar{z}^2}$$
$$= \frac{\sum_{k=1}^{N} (F(z_k) - F(z_{k-1}))(M_k - \bar{z})^2}{\int (z - \bar{z})^2 \, dF(z)},$$

which is the ratio of the variance between the conditional mean qualities and total variance. Intuitively, this bound is a measure of the relative importance of the variance between the means of the partitions, separated by their ratings, and the variance that remains in each pool. This connection to variance decomposition is used below to derive a theoretical bound on ratings' performance.

4.1 Theoretical Bounds

The simplest coarse rating scheme is a two-tier certification, widely used in many settings. The next proposition provides a useful bound for the gains from certification that builds on the variance decomposition described above. The corollary that follows gives sufficient conditions so that a two-tier rating achieves at least half of the surplus of the full-information case.

Proposition 4. *The relative performance of a two-tier setting satisfies*

$$\gamma \geq \frac{1}{1 + \max\{cv_1^2, cv_2^2\}},$$

where cv_1 is the coefficient of variation of $z - \bar{z}$ conditional on $z < \bar{z}$, and cv_2 is the coefficient of variation of $z - \bar{z}$ conditional on $z \geq \bar{z}$, where $\bar{z}$ is the mean of sellers' qualities.

Corollary 1. *Suppose that the distribution F has an increasing hazard rate and a decreasing reverse hazard rate. Then a two-tier rating achieves at least half of the surplus of the full-information case.*

Proof. From a well-known result from Stoyan and Daley (1983) (pp. 16–19), the conditions of this corollary imply that $cv_1 < 1$ and $cv_2 < 1$. Using the bound in Proposition 4 completes the proof. □

The conditions given in the corollary are satisfied by a large class of distributions that include all those with log-concave densities, such as uniform, normal, exponential and double exponential, logistic, extreme value, and many others with some restriction on parameters (e.g., power function $F(z) = z^c$ for $c \geq 1$.) Related bounds for two-sided matching problems can be found in McAfee (2002); Hoppe et al. (2011); Shao (2016). The results of Wilson (1989) imply that the losses from $N-$ratings are of order $1/N^2$.

4.2 Numerical Results

We now examine the numerical results for a variety of distribution functions often used in the economics literature. Table 1 reports the share of the total surplus gap that is closed with partitions of different sizes n. As can be seen from the calculations, a one-threshold (certification) partition closes from near 50% to almost 80% of the total surplus gap, depending on the underlying distribution of qualities. The only case where one threshold cannot reach 50% of the benefits is the Pareto distribution, which does not satisfy the conditions stated in 1. The gains are diminishing as the number of thresholds increases. Even though total surplus increases with the number of tiers, our numerical results suggest that most gains are attained with a small number of ratings.

The numerical results and the previous theoretical bounds suggest that in practice we may not need very complicated mechanisms to attain most of the benefits from information provision. In this paper, we do not explicitly model any cost related to providing information; but in practice, providing information may involve various costs for both the market designer and consumers. First, we have assumed that the market designer has costless access to information on the quality of sellers; however, in practice, getting precise information may increase the market designer's cost of designing the rating system. Secondly, it might be costly for the market designer to convey finer information to consumers and for them to process this information. Therefore, these costs and the small benefits from more complex information mechanisms might justify the coarse information mechanisms seen in practice.

Table 1: Optimal Thresholds

Distribution	Case	Mean/Median	z^*	$1-F(z^*)$	Share of Surplus Gap Closed			
					$n=2$	$n=3$	$n=5$	$n=10$
Pareto	$\alpha=3$	1.19	2.73	0.05	0.46	0.68	0.84	0.94
	$\alpha=4$	1.12	1.84	0.09	0.54	0.74	0.89	0.97
Exponential	all	1.45		0.20	0.65	0.82	0.93	0.98
$F(z)=z^{\alpha}$	$\alpha=0.5$	1.32	0.41	0.36	0.77	0.90	0.97	0.99
$z\in[0,1]$	$\alpha=2$	0.94	0.62	0.62	0.72	0.87	0.95	0.99
Log-normal	$\sigma=0.25$	1.03	1.09	0.36	0.63	0.81	0.92	0.98
$(\mu=0)$	$\sigma=1$	1.64	4.25	0.07	0.55	0.75	0.89	0.97

Note: The above calculations correspond to the linear supply case.

4.3 Empirical Application

Medicare Advantage Providers Here we apply our method of finding optimal thresholds to Medical Advantage providers. The data, which generously shared with us by Benjamin Vatter, consists of estimates of the quality of products offered by insurance firms across the United States (details on the source of this data are given in Vatter (2021)). Our application is mainly meant to illustrate the workings of our method using an empirically relevant distribution, and is an overly-simplified version of how this market works. In particular, we pool all the national data and abstract from the local nature of many of these insurance markets, as well as abstracting from moral hazard considerations.[17]

Figure 2 gives the baseline discrete distribution of qualities along with a density kernel estimation, which is the one used in our calculations. The optimal two-tier rating is defined by a threshold quality equal to 14.9, which is slightly below the mean quality, implying that 66.5% of the firms are certified. This threshold achieves 64% of the gap between the full-information and no-information cases. Notice that by certifying a high fraction of firms, this rating serves mostly the purpose of screening out the lower tail of quality, which as seen in Section 5.1 is optimal when the distribution of qualities is left skewed, as in this case.

eBay Sellers' Ratings Nosko and Tadelis (2015) provides a quality measure given by the per-

[17] All these features and other important considerations relevant to the design of a rating system for this industry are carefully considered in Vatter's excellent paper.

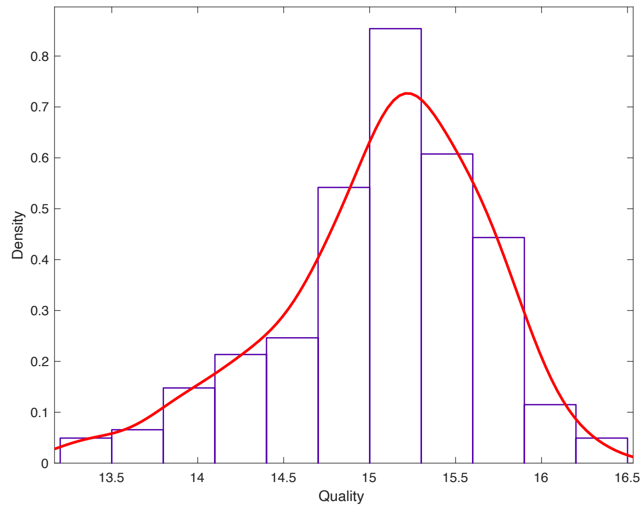

Figure 2: Medicare Advantage Quality Distribution

centage time a seller got positive feedback (as opposed to negative or none). The distribution of this statistic across sellers is given in Figure 3 together with a density kernel estimator.[18] If we interpret this statistic as an ex ante probability of a good (vs. a bad) experience and the expected utility from this purchase as $(1 - P(good))\, u(bad) + P(good)u(good)$, then expected utility is an affine transformation of the probability of a good experience. Based on this interpretation, we can use this distribution to calculate the optimal certification threshold, as we did for other distributions above. Table 1 reports the results for the kernel estimate of this distribution. According to our calculations, more than 65% of sellers should be certified, closing about 63% of the surplus gap. As in the Medicare application, the distribution of qualities is also left-skewed and certification serves the purpose of mainly screening out the lower tail.

5 Other Design Considerations

In this section we explore other features that are relevant to the design of optimal ratings. In Section 5.1, we consider the role of the distribution of seller qualities F, and in particular, its skewness, as already emphasized in our applications. In Section 5.2, we discuss the conflict of interest between

[18] The data for the histogram comes directly from Table 4 in Nosko and Tadelis (2015).

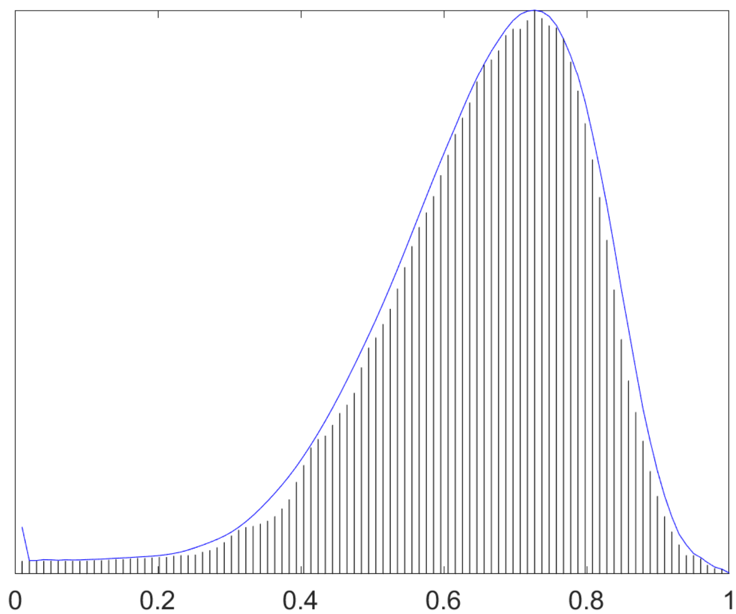

Figure 3: Distribution of Percent Positive Responses for eBay Sellers

consumers and producers in the design of simple ratings. Section 5.3 considers the case of vertical differentiation, where buyers differ in their preference for quality, and the matching between goods' quality and consumer's type becomes important. Section 5.4 considers the role of entry.

5.1 Skewness and Optimal Thresholds

The optimal thresholds, as depicted in equation 6, will depend on the distribution of sellers' quality (i.e., F distribution). In this section, we study how skewness in the distribution of qualities impacts this optimal choice. In particular, we show that in the simple case of a two-tier certification, the optimal threshold is skewed in the same direction as the distribution of qualities. Then, we extend this result, providing general comparative statics for the vector of thresholds with respect to an appropriately defined skewness ordering.

Before proceeding to the analysis, we provide some intuition behind our results. Consider the case of one certification threshold, z^*. The following trade-off appears when deciding how strictly to draw the line separating the upper and lower segments. When putting z^* in the upper group, there is an upward distortion of the supply of the seller at z^*, which is a function of the distance $M_H - z^*$. This distance also measures the extent to which the seller at z^* gains from being pooled with higher-quality sellers. When putting z^* in the lower group, there is a downward distortion of the supply of the seller at z^*, which is a function of $M_L - z^*$. This distance also measures the extent to which the seller at z^* loses from being pooled with lower-quality sellers. Right skewness (resp., left skewness) of the distribution $F(z)$ will increase (resp., decrease) the upward distortion and decrease (resp., increase) the downward distortion, making it optimal to have more restrictive (resp., less restrictive) certification standards.

The condition given in Proposition 2 for linear supply functions implies

$$z^* = \frac{1}{2}\left(M_L(z^*) + M_H(z^*)\right), \tag{13}$$

which can be used to relate this threshold to properties of the distribution. Consider first the case of a symmetric distribution (i.e., where the median, z_{median}, equals the mean, $\bar{z}$). Since for any z^*, $F(z^*) M_L + (1 - F(z^*)) M_H = \bar{z}$, setting the threshold $z^* = \bar{z} = z_{median}$ would satisfy the above condition.

The same reasoning suggests that when F is skewed, the optimal threshold will also be skewed relative to the mean in the same direction. This can be easily proved, as follows. Consider the case of a right skewed distribution where $\bar{z} > z_{median}$. Let $M_L(.)$ and $M_H(.)$ denote functions equal to the conditional average of the quality of sellers below and above any value within the range of qualities, respectively. Furthermore, denote $g(z) = \frac{1}{2}(M_L(z) + M_H(z))$. Following Proposition 1, the optimal threshold is a fixed point of this function. When $z \to z_{max}$ (or as $z \to \infty$ in the case of unbounded support), $g(z) \to \frac{1}{2}\bar{z} + \frac{1}{2}z < z$, and when $z \to z_{min}$(or as $z \to -\infty$ in the case of unbounded support), $g(z) \to \frac{1}{2}z_{min} + \frac{1}{2}\bar{z} > z$. For $z = z_{median}$, $g(z) = \bar{z} > z$. Since the function $g(z)$ is increasing and continuous, the unique fixed point z^* must be to the right of z_{median} and, as a consequence, $z^* > \bar{z}$, as illustrated in Figure 4. The result for the case of left skewness can be shown similarly. Consistently with these results, the empirical examples in Section 4.3 have both left-skewed distributions and the optimal thresholds we found are below the respective means.

Table 1 shows the optimal threshold for a series of distributions, as well as the corresponding

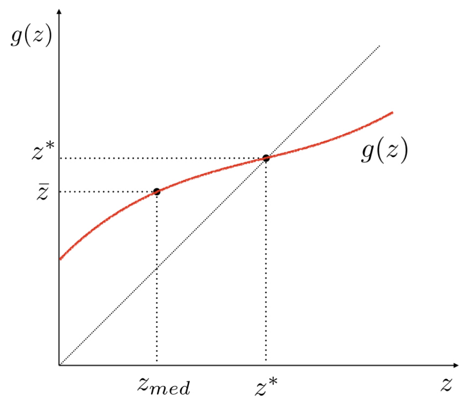

Figure 4: z^* When the Mean Is Greater than the Median

fraction of certified sellers. All distributions in our example are skewed to the right except for one, so according to our argument, $z^* > \bar{z} > z_{median}$ and it is optimal to have a smaller share of sellers certified. This is shown in the fifth column of Table 1. As an example, for the Pareto distributions only a small fraction should get certified, 5% when the power parameter is 3, and 9% when the power parameter is 4.[19] For the exponential distribution, only 20% of sellers should be certified regardless of the hazard rate.

Now, we extend our findings to the case of multiple signals under a stronger skewness order. This skewness order, the convex (resp., concave) order, was originally proposed by Van Zwet (1964).

Definition. Distribution $\tilde{F}$ is more skewed to the right than F if $\tilde{F}^{-1}(F(x))$ is convex; equivalently, there exists an increasing convex function $g(x)$ such that $\tilde{F}(g(x)) = F(x)$.[20]

We can think of this ordering as stretching to the right the quality scale with the transformation $g(x)$. As an example, if F is a uniform distribution in $[0,1]$ and $g(x) = x^2$, then $\tilde{F}(x^2) = x$ or, equivalently, $\tilde{F}(x) = x^{1/2}$.

Proposition 5. *Suppose the supply function is linear. Let F be a distribution with log-concave density and $\tilde{F}$ a distribution such that $\tilde{F}(g(z)) = F(z)$, where g is a strictly convex increasing function. Let $\{l_k\}$ be the optimal thresholds for F, and $\{g(z_k)\}$, the optimal thresholds for $\tilde{F}$. Then $z_k > l_k$ for all k.*

[19]When $\alpha \leq 2$, the value of z^* is undefined, as total surplus is strictly increasing in z^* in all the support.
[20]Note that this definition implies that $F^{-1}(F(x)) = g^{-1}(x)$ is concave.

This proposition implies that for all k, $\tilde{F}(g(z_k)) = F(z_k) > F(l_k)$, so the percentiles defined by the two optimal thresholds are ordered. In particular, for a two-tier certification rating, the share of certified sellers should be lower for distribution $\tilde{F}$. An example is given in Table 1 for the case of power distributions $F(z) = z^{\alpha}$. It is easily shown that the distribution with $\alpha = 0.5$ is more skewed to the right than the one with $\alpha = 2$.[21] Consistently with the previous proposition, the share of certified sellers is lower when $\alpha = 0.5$.

5.2 Consumer and Producer Surplus

We have focused on total surplus as the objective function. However, it is easy to show that at the optimal thresholds, there is generically a conflict of interest between consumers and producers, and that the optimal choice balances off these conflicting interests. The difference lies in the equilibrium effects of ratings: consumers' surplus increases with total output, while profits decrease. These two opposing effects balance each other exactly at the optimal thresholds. In which direction would consumers like the threshold to move? In particular, would consumers prefer stricter or less strict criteria for certification? Again, the answer depends on the properties of the supply function.

In the case of concave supply, total output decreases with increasing information, so total output and consumer surplus are maximized with thresholds at the extremes of the distribution, i.e., a trivial partition with no information. When supply is linear, total quantity is independent of the amount of information, so consumer surplus is the same for any threshold. This implies that the optimal threshold is also the one that maximizes profits. The following proposition provides sufficient conditions that determine the direction of change of output (and consumer surplus) at the optimal thresholds. The direction of change of producer surplus has the opposite sign.

Proposition 6. *Let $\boldsymbol{z} = (z_1, ..., z_{N-1})$ be the thresholds that maximize total surplus. If $S''(p)/S'(p)$ is decreasing (resp., increasing) in p, then $dQ(\boldsymbol{z})/dz_k$ and $dCS(\boldsymbol{z})/dz_k$ are negative (resp., positive) at $\boldsymbol{z}$.*

To illustrate the above results, consider a simple example. Suppose the supply function $s(p) = p^{\theta}$ (cost function $c(q) = q^{1+\theta}/(1+\theta)$). Therefore, $S''(p)/S'(p) = (\theta - 1)/p$. For $\theta > 1$, this

[21] Take $g(x) = x^4$.

expression is decreasing in p. Therefore, starting at the surplus maximizing thresholds, consumers would prefer lower thresholds, while producers would prefer higher ones. Therefore, if the planner puts more weight on consumers, it should lower the thresholds, while if it puts more weight on sellers, it should increase them. The reverse occurs when $\theta < 1$.

5.3 Heterogeneous Preference for Quality

In this extension, we consider a demand system where agents have heterogeneous preference for quality, and sellers have inelastic supply. While by construction, improvements in information do not increase total quantity, they contribute to welfare by increasing the correlation between average seller quality and consumer preference for quality. The optimal threshold is defined by a slightly modified formula that weighs differences in the sellers' quality gap in each interval by the respective gap in consumers' preferences. As a result, skewness in consumers' preferences for quality has similar implications to the ones observed for skewness in producers' quality.

We examine briefly the determination of optimal thresholds when consumers differ in their preferences for quality for the case of certification, i.e., $N = 2$. Suppose consumers' preferences are given by the utility function $u = \theta z + \theta_0 - p$ for a good of quality z, à la Mussa and Rosen (1978). Consumers differ in their preference for quality θ and for the value they assign the inside vs. outside good θ_0, which is distributed in the population according to some joint distribution $\Psi(\theta, \theta_0)$. As earlier, seller qualities z are distributed according to the cdf $F(z)$. For simplicity, we restrict our analysis to a partition of sellers into two groups defined by threshold z^* with qualities z_L and z_H, respectively. Given prices p_L and p_H, consumers will be split into three groups: those that do not consume and those that consume either the H or L product, with demands $D_H(p_L, p_H)$ and $D_L(p_L, p_H)$, respectively. Prices p_L and p_H will be equilibrium prices provided that $D_H(p_L, p_H) = (1 - F(z^*))\, q(p_H)$ and $D_L(p_L, p_H) = F(z^*)\, q(p_L)$. As in our previous case, there is a unique equilibrium under fairly general conditions.

Lemma 3. *The optimal choice of threshold z^* satisfies the following first-order necessary condition:*

$$\Pi(p_H) - \Pi(p_L) = (z^* - z_L)\, \theta_L q_L + (z_H - z^*)\, \theta_H q_H, \tag{14}$$

where θ_L is the average preference for quality of consumers who purchase the L product, and θ_H, of those who purchase the H product.

This formula has an intuitive explanation. The first term is the loss of profits of those sellers that transition from the H to the L group, when z^* is marginally increasing. The second term measures the effect of the increase in the averages z_L and z_H as z^* is increased, valued at the quality preference of the average consumer in each group and weighted by their respective market sizes.

Vertical Differentiation with Inelastic Supply

To establish further results, we consider the canonical model of vertical differentiation where consumers differ only in their preference for quality θ and where sellers supply inelastically one unit of output.[22] Given equilibrium prices p_L and p_H, all consumers above a threshold θ^* buy an H product, while all those between $\underline{\theta}$ and θ^* buy an L product, where $\underline{\theta} z_L = p_L$ and $\theta^* (z_H - z_L) = p_H - p_L$. Substituting in equation (14) gives the condition

$$(z^* - z_L)(\theta^* - \theta_L) \quad = \quad (z_H - z^*)(\theta_H - \theta^*) .$$

Notice that this equation is a modified version of equation (16), where the gaps between z^* and the respective means are weighted by the corresponding preference gaps. This equation highlights the role of the complementarities between average quality and preference for quality in the determination of the optimal threshold. In particular, when both distributions are symmetric, this also implies that the optimal threshold z^* (and also θ^*) will equal the corresponding mean. Moreover, when z and θ have the same distribution, the optimal threshold is also given by our baseline condition, as given in equation (16).[23] As an example, if both have uniform distributions, then when $\theta^* = z^* = 1/2$, this condition will hold.

[22] This case can be reinterpreted as a one-to-one matching environment with surplus function θz.

[23] It is interesting to note that when all consumers have the same preference for quality and supply is inelastic, welfare is independent of z^*, as the average product quality is not affected by its choice.

5.4 Entry

In our previous analysis we did not consider explicitly the effect of changes in z^* on entry. Many of our results extend to settings where the distribution of qualities of sellers is not affected by entry. We discuss here two scenarios: one where entrants are ex ante differentiated and one where they are ex ante homogeneous.

Consider first the case of differentiated entrants. Our analysis extends without modification to the following scenario. Suppose there is a mass n of entrants that are differentiated in qualities z and fixed (or entry) costs f. Assume qualities are independent from fixed costs and are given by distribution F and Φ, respectively. For a given threshold partition z^*, we can define the aggregate supply functions S_L and S_H as follows. Let $S_H(p) = S(p) N_H(p)$, where $N_H(p) = n(1 - F(z^*)) \Phi(\pi(p))$. This supply function combines the effect of prices on the intensive and extensive margin. We can define similarly $S_L(p)$. Our analysis remains unchanged if we substitute $S(p)$ by $\hat{s}(p) = S(p) n\Phi(f(p))$, so total supplies are $S_L(p) = F(z^*)\hat{s}(p)$ and $S_H(p) = \hat{s}(p)(1 - F(z^*))$.[24]

For the homogeneous case, assume there is a set N of potential entrants that draw their qualities independently from distribution F upon entry, after paying an entry cost f, which is distributed according to cdf $\Phi(f)$. For fixed output, improved information results in a mean preserving spread of expected qualities and thus prices. Given that profit functions are convex in prices, this results in an increase in expected profits and a consequent increase in entry. In the case of linear supply, where in the absence of entry, total output does not change, additional entry results in an increase in total output and thus consumer surplus. In the case of concave supply, we have seen that total output decreases. This increases profits over what is produced by the mean preserving spread of average qualities, thus inducing entry and mitigating, if not totally undoing, the drop in total output that would result in the absence of entry. Finally, note that if all potential entrants were to have the same entry cost, all surplus gains from improved information would accrue to consumers, as expected, and average profits would remain unchanged. The above results apply in particular

[24]The properties of these modified supply functions will now depend both on the individual supply functions and the distribution of fixed costs. There exist assumptions on the latter that will guarantee that the modified supply functions are linear, convex, or concave when each of these properties holds for the original supply functions.

to the effect of introducing a certification mechanism in a market where there is none.

6 Final Remarks

In this paper we considered the optimal design of quality ratings in markets with adverse selection and limited signals. Ratings reallocate demand across producers, impacting not only the average quality of goods consumed but also average cost. The optimal thresholds in a discrete rating system optimize this trade-off. Optimal ratings thus depend on the characteristics of the market, given by the distribution of producers' quality, the elasticity of supply, and consumers' preferences. We find that the optimal thresholds in the case of a convex (resp., concave) supply function are pointwise higher (resp., lower) than those in the linear case. Intuitively, in the case of a simple certification rating with two groups, more elastic supply leads to a higher threshold and lower share of certified sellers. We also find that skewness in the distribution of seller qualities matters for optimal ratings, which move in the direction of the skew.

We have given a simple characterization for the optimal thresholds in the case of linear supply as the solution to standard clustering problems. Our results thus provide a straightforward and easy-to-compute method for the design of rating systems. This method is used to derive bounds on the performance of the rating system as a function of the number of categories. We first theoretically showed that a simple certification mechanism, or a two-tier rating, is enough to reach half of the benefits of the best rating mechanism in a large family of sellers' quality distributions such as log-concave densities. As an example, we found that for the exponential family of distributions, 65% of the total surplus gains from the full-information case can be achieved with only two categories. We also applied our method to two empirical problems in order to find the optimal thresholds and the gain from them. The large gains in total surplus with a very simple threshold mechanism suggest that the added cost of a more complex one might not be compensated by the gains from it. This finding could explain the popularity of these simple schemes among market designers.

References

Bergemann, D. and A. Bonatti (2019): "Markets for information: An introduction," *Annual Review of Economics*, 11. 4

Bergemann, D. and M. Pesendorfer (2007): "Information structures in optimal auctions," *Journal of economic theory*, 137, 580–609. 10

Dranove, D. and G. Z. Jin (2010): "Quality disclosure and certification: Theory and practice," *Journal of Economic Literature*, 48, 935–963. 4

Elfenbein, D. W., R. Fisman, and B. McManus (2015): "Market structure, reputation, and the value of quality certification," *American Economic Journal: Microeconomics*, 7, 83–108. 4

Fan, Y., J. Ju, and M. Xiao (2013): "Losing to Win: Reputation Management of Online Sellers," . 4, 5

Filippas, A., J. J. Horton, and J. Golden (2018): "Reputation inflation," in *Proceedings of the 2018 ACM Conference on Economics and Computation*, ACM, 483–484. 5

Ganuza, J.-J. and J. S. Penalva (2010): "Signal Orderings Based on Dispersion and the Supply of Private Information in Auctions," *Econometrica*, 78, 1007–1030. 6

Gentzkow, M. and E. Kamenica (2016): "A Rothschild-Stiglitz Approach to Bayesian Persuasion," *American Economic Review*, 106. 6

Hopenhayn, H. and M. Saeedi (Forthcoming): "Optimal Information Disclosure and Market Outcomes," *Theoretical Economics.* 4, 5, 7, 16, 32

Hoppe, H. C., B. Moldovanu, and E. Ozdenoren (2011): "Coarse matching with incomplete information," *Economic Theory*, 47, 75–104. 4, 18

Hui, X., Z. Liu, and W. Zhang (2020): "Mitigating the Cold-start Problem in Reputation Systems: Evidence from a Field Experiment," *Available at SSRN 3731169.* 5

Hui, X., M. Saeedi, Z. Shen, and N. Sundaresan (2016): "Reputation and regulations: evidence from eBay," *Management Science*, 62, 3604–3616. 5

Hui, X., M. Saeedi, G. Spagnolo, and S. Tadelis (forthcoming): "Raising the Bar: Certification Thresholds and Market Outcomes," *American Economic Journal: Microeconomics*. 2

Hui, X., M. Saeedi, and N. Sundaresan (2018): "Adverse selection or moral hazard, an empirical study," *The Journal of Industrial Economics*, 66, 610–649. 2

Jin, G. Z. and P. Leslie (2003): "The effect of information on product quality: Evidence from restaurant hygiene grade cards," *The Quarterly Journal of Economics*, 118, 409–451. 4, 5

MacQueen, J. et al. (1967): "Some methods for classification and analysis of multivariate observations," in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, Oakland, CA, USA, vol. 1, 281–297. 3, 12

McAfee, R. P. (2002): "Coarse matching," *Econometrica*, 70, 2025–2034. 4, 18

Mease, D. and V. N. Nair (2006): "Unique optimal partitions of distributions and connections to hazard rates and stochastic ordering," *Statistica Sinica*, 1299–1312. 34

Moscarini, G. and M. Ottaviani (2001): "Price competition for an informed buyer," *Journal of Economic Theory*, 101, 457–493. 15

Mussa, M. and S. Rosen (1978): "Monopoly and product quality," *Journal of Economic theory*, 18, 301–317. 25

Nosko, C. and S. Tadelis (2015): "The limits of reputation in platform markets: An empirical analysis and field experiment," Tech. rep., National Bureau of Economic Research. 5, 19, 20

Saeedi, M. (2019): "Reputation and adverse selection: theory and evidence from eBay," *The RAND Journal of Economics*, 50, 822–853. 4

Saeedi, M. and A. Shourideh (2022): "Optimal Rating Design," . 2

Shao, R. (2016): "Generalized coarse matching," *Games and Economic Behavior*, 100, 142–148. 4, 18

Shi, Z. J., K. Srinivasan, and K. Zhang (2020): "Design of Platform Reputation Systems: Optimal Information Disclosure," *Kannan and Zhang, Kaifu, Design of Platform Reputation Systems: Optimal Information Disclosure (March 14, 2020).* 2

Stoyan, D. and D. J. Daley (1983): "Comparison Methods for Queues and Other Stochastic Models." *JOHN WILEY & SONS, INC., 605 THIRD AVE., NEW YORK, NY 10158, USA, 1983.* 17

Van Zwet, W. (1964): "Convex transformations of random variables, Math," *Centrum, Amsterdam*, 1964. 23

Vatter, B. (2021): "Quality disclosure and regulation: Scoring design in medicare advantage," Tech. rep., Working Paper. 2, 4, 19

Wilson, R. (1989): "Efficient and competitive rationing," *Econometrica: Journal of the Econometric Society*, 1–40. 4, 18

A Proofs

Proof of Lemma 1

Proof. Consider a partition of the set of sellers into sets $S_1, ..., S_N$. Suppose there are two sets S_k, S_{k+1} that are not totally ordered in quality with means $M_k \leq M_{k+1}$ and mass G_k and G_{k+1}. By reordering elements of these two sets, one can substitute S_k and S_{k+1} with two new disjoint sets S'_k and S'_{k+1} of equal measures to the original ones, where $S_k \cup S_{k+1} = S'_k \cup S'_{k+1}$ and $S'_k < S'_{k+1}$, element-wise. By construction, $M'_k \leq M_k \leq M_{k+1} \leq M'_{k+1}$ and $G'_k M'_k + G'_{k+1} M'_{k+1} = G_k M_k + G_{k+1} M_{k+1}$. This corresponds to a mean preserving spread of the original distribution of means. Using Proposition 1 in Hopenhayn and Saeedi (Forthcoming), this results in higher total surplus. □

Proof of Lemma 2

To totally differentiate equation (3) with respect to z_k, first note that by the envelope condition, we can ignore the effect on the output choices $q_1, ..., q_N$. In particular, this implies that $\partial Q / \partial z_k = f(z_k)(q_k - q_{k+1})$. Since $M_k = \int_{z_{k-1}}^{z_k} z dF(z) / (F(z_k) - F(z_{k-1}))$, it follows that

$$\frac{\partial (F(z_k) - F(z_{k-1})) M_k}{\partial z_k} = f(z_k) z_k, \frac{\partial (F(z_{k+1}) - F(z_k)) M_{k+1}}{\partial z_k} = -f(z_k) z_k.$$

The result now follows by totally differentiating (3) and setting it equal to zero.

Proof of Proposition 1

To prove this proposition, we need an intermediate step, which is proven using the following lemma.

Lemma 4. *The optimal thresholds satisfy the following condition:*

$$\frac{z_k - M_k}{M_{k+1} - M_k} S(p_k) + \frac{M_{k+1} - z_k}{M_{k+1} - M_k} S(p_{k+1}) = \frac{\int_{p_k}^{p_{k+1}} S(p) dp}{p_{k+1} - p_k}, \tag{15}$$

where M_k and M_{k+1} are the conditional mean qualities for the two groups, and p_k and p_{k+1} are the equilibrium prices.

Proof. First note that

$$\begin{aligned}\left(P\left(Q\right)+z_k\right)\left(q_{k+1}-q_k\right) &= \left(P\left(Q\right)+M_{k+1}-M_{k+1}+z_k\right)q_{k+1} \\ &- \left(P\left(Q\right)+M_k-M_k+z_k\right)q_k \\ &= p_{k+1}q_{k+1}-p_kq_k-\left(M_{k+1}-z_k\right)q_{k+1}-\left(z_k-M_k\right)q_k.\end{aligned}$$

Substituting in (5) and rearranging gives

$$\left(M_{k+1}-z_k\right)q_{k+1}+\left(z_k-M_k\right)q_k=\pi_{k+1}-\pi_k.$$

Equation (15) follows by substituting $\pi_{k+1}-\pi_k=\int_{p_k}^{p_{k+1}}S\left(p\right)dp$, using $q_{k+1}=S\left(p_{k+1}\right)$ and $q_k=S\left(p_k\right)$, and dividing the left hand side by $\left(M_{k+1}-M_k\right)$ and the right hand side by the equivalent value $p_{k+1}-p_k$. □

We use the expression found in Lemma 4. Equation (15) equates the expected value of $S\left(p\right)$ under two lotteries. The left hand side lottery has weights $\alpha=\left(z_k-M_k\right)/\left(M_{k+1}-M_k\right)$ on price p_k and $\left(1-\alpha\right)$ on price p_{k+1}. The second lottery is uniform between these two extreme prices. When S is linear, it must be the case that $\alpha=1/2$, and this implies that

$$z_k-M_k=M_{k+1}-z_k. \tag{16}$$

When S is convex, $\alpha>1/2$ so $z_k-M_k>M_{k+1}-z_k$, so the optimal threshold is above the one defined by equation (16), while the reverse occurs when s is concave. This concludes the proof.

Proof of Proposition 2

Without loss of generality, let $S(p) = p$, so the cost function $c(q) = \frac{1}{2}q^2$. Consider now the objective function (3) for this case:

$$W(\boldsymbol{z}) = \int_0^Q P(x)\,dx + \sum_{k=1}^{N}\left[F(z_k) - F(z_{k-1})\right]\left[M_k(P + M_k) - \frac{1}{2}\left((P + M_k)^2\right)\right] \quad (17)$$

$$= \int_0^Q P(x)\,dx + \sum_{k=1}^{N}\left[F(z_k) - F(z_{k-1})\right]\left[\frac{1}{2}M_k^2 - \frac{1}{2}P^2\right]. \quad (18)$$

After suppressing the terms that are unaffected by the partition, maximizing this expression is equivalent to maximizing

$$\sum_{k=1}^{N}\left[F(z_k) - F(z_{k-1})\right](M_k - \bar{z})^2, \quad (19)$$

where $\bar{z} = \sum_{k=1}^{N}\left[F(z_k) - F(z_{k-1})\right]M_k$ is the mean seller quality, which is independent of the partition. The above expression is the variance between partitions. Since total variance is fixed, maximizing (19) is equivalent to minimizing (6). Uniqueness of the thresholds is guaranteed when the distribution has log-concave density, as shown in Mease and Nair (2006).

Proof of Proposition 3

We use the following properties of distributions with log-concave densities (see Lemma 1 in Mease and Nair (2006)):

$$\mathbb{E}(z|s \le z \le s + d) - s \text{ is decreasing in } s \text{ for } d > 0 \text{ and} \quad (20)$$

$$s - \mathbb{E}(z|s - d \le z \le s) \text{ is increasing in } s \text{ for } d > 0, \quad (21)$$

and these properties are preserved when conditioning on intervals.

Lemma 5. *Suppose F is a distribution with log-concave density, and let $m(a, b) = E_F(z|a \le z \le b)$. Suppose the vector of thresholds $\{l_k\}_{k=1}^{N-1}$ satisfies*

$$l_k - m\left(l_{k-1}, l_k\right) = m\left(l_k, l_{k+1}\right) - l_k, \tag{22}$$

and let $z_1, \ldots, z_{N-1}$ *be a vector such that*

$$z_k - m\left(z_{k-1}, z_k\right) > m\left(z_k, z_{k+1}\right) - z_k. \tag{23}$$

Then $z_k > l_k$ *for all* k.

To prove Lemma 5, we first use the following:

Claim. Under the assumptions of Lemma 5, suppose that for some k, $z_k < l_k$ and $z_{k+1} - z_k \geq l_{k+1} - l_k$. Then $z_{k-1} < l_{k-1}$ and $z_k - z_{k-1} \geq l_k - l_{k-1}$.

Proof. Note that

$$\begin{aligned} z_k - m\left(z_{k-1}, z_k\right) &> m\left(z_k, z_{k+1}\right) - z_k \\ &\geq m\left(z_k, z_k + l_{k+1} - l_k\right) - z_k \\ &\geq m\left(l_k, l_{k+1}\right) - l_k \\ &= l_k - m\left(l_{k-1}, l_k\right). \end{aligned} \tag{24}$$

The first inequality follows from (23), the second one, from the monotonicity of m, the third, from (20), and the last, from (22). Now consider $k-1$. We will show that $z_k - z_{k-1} \geq l_k - l_{k-1}$.

Suppose, by way of contradiction, that $z_k - z_{k-1} < l_k - l_{k-1}$. Then

$$\begin{aligned} z_k - m\left(z_{k-1}, z_k\right) &\leq l_k - m\left(l_k - \left(z_k - z_{k-1}\right), l_k\right) \\ &\leq l_k - m\left(l_{k-1}, l_k\right) \end{aligned}$$

where the first inequality follows from condition (21), and the second one, from the monotonicity of m. This inequality contradicts (24), proving that $z_k - z_{k-1} \geq l_k - l_{k-1}$. Given that $z_k < l_k$, this also guarantees that $z_{k-1} < l_{k-1}$. □

We now prove Lemma 5. Let h denote the highest k for which $z_k < l_k$. By the definition of h, $z_{h+1} - z_h > l_{h+1} - l_h$. Using inductively the previous claim, it follows that the same is true for all $k = 1, ..., h$. For $k = 1$, the claim would imply that $z_0 < l_0$, which cannot be true if the distribution had a lower bound, since in that case both z_0 and l_0 should equal this lower bound. For unbounded support, an argument similar to the one used in the claim can be used to generate a contradiction. This completes the proof.

Proof of Proposition 3.

Let $\{l_k\}$ denote the optimal thresholds for the linear supply function, and $\{z_k\}$, those for the convex supply function. Lemma 1 and equations (22) and (24) hold, so Lemma 5 proves the proposition.

Proof of Proposition 4.

Proof. Let $\bar{M}_1$ and $\bar{M}_2$ be the conditional mean of z below and above the mean $\bar{z}$, respectively. By the variance decomposition,

$$\begin{aligned}
\int (z-\bar{z})^2 dF(z) &= \int^{\bar{z}} \left(z-\bar{M}_1\right)^2 dF(z) + \int_{\bar{z}} \left(z-\bar{M}_2\right)^2 dF(z) \\
&\quad + F(\bar{z})\left(\bar{M}_1-\bar{z}\right)^2 + (1-F(\bar{z}))\left(\bar{M}_2-\bar{z}\right)^2 \\
&= F(\bar{z})\left(cv_1^2+1\right)\left(\bar{M}_1-\bar{z}\right)^2 + (1-F(\bar{z}))\left(cv_2^2+1\right)\left(\bar{M}_2-\bar{z}\right)^2 \\
&\leq \left(\max\left\{cv_1^2, cv_2^2\right\}+1\right)\left(F(\bar{z})\left(\bar{M}_1-\bar{z}\right)^2 + (1-F(\bar{z}))\left(\bar{M}_2-\bar{z}\right)^2\right),
\end{aligned}$$

where the second equality follows from

$$cv_1 = \frac{\int^{\bar{z}} \left((z-\bar{z}) - \left(\bar{M}_1-\bar{z}\right)\right)^2 dF(z)}{F(\bar{z})\left(\bar{M}_1-\bar{z}\right)^2}$$

,and similarly for cv_2. From the above inequality,

$$\frac{F(\bar{z})\left(\bar{M}_1-\bar{z}\right)^2 + (1-F(\bar{z}))\left(\bar{M}_2-\bar{z}\right)}{\int (z-\bar{z})^2 dF(z)} \geq \frac{1}{1+\max\left\{cv_1^2, cv_2^2\right\}}.$$

This gain corresponds to setting $z^* = \bar{z}$, so it is a lower bound to the gains under the optimal threshold. $\square$

Proof of Proposition 5

To prove this proposition, we first need to show the following lemma.

Lemma 6. *Let $g(z_1), ..., g(z_{N-1})$ be the optimal thresholds for $\tilde{F}$. Let $M_k = m(z_{k-1}, z_k) = E_F(z_{k-1} \leq z \leq z_k)$. Then $z_k - M_k > M_{k+1} - z_k$.*

Proof. Let $\tilde{M}_k = E_{\tilde{F}}(g(\tilde{z}_{k-1}) \leq z \leq g(\tilde{z}_k))$. Note that by strict convexity of g, $\tilde{M}_k > g(M_k)$. It follows that

$$\begin{aligned} z - M_k &> z_k - g^{-1}\left(\tilde{M}_k\right) \\ &= g^{-1}(g(z_k)) - g^{-1}\left(\tilde{M}_k\right) \\ &= g^{-1}\left(\tilde{M}_{k+1}\right) - g^{-1}(g(z_k)) \\ &> M_{k+1} - z_k. \end{aligned}$$

$\square$

To prove the proposition, let the vector $\{l_k\}$ be the optimal thresholds for F, and $\{z_k\}$, the optimal thresholds for $\tilde{F}$. Equation (22) follows from the necessary condition for optimal thresholds, and (24) follows from the previous lemma.

Proof of Proposition 6

To prove this proposition, we need to show the following lemma first:

Lemma 7. *The term $dQ(\boldsymbol{z})/dz_k$ has the same sign as*

$$\frac{z_k - M_k}{M_{k+1} - M_k} S'(p_k) + \frac{M_{k+1} - z_k}{M_{k+1} - M_k} S'(p_{k+1}) - \frac{\int_{p_k}^{p_{k+1}} S'(p)\, dp}{p_{k+1} - p_k}. \tag{25}$$

Proof. Total output is

$$Q = \sum_{k=1}^{N} \left(F\left(z_k\right) - F\left(z_{k-1}\right)\right) S\left(p_k\right),$$

where $p_k = P\left(Q\right) + M_k$. Differentiating with respect to z_k and using

$$\begin{aligned}
\left(F\left(z_k\right) - F\left(z_{k-1}\right)\right) \frac{\partial M_k}{\partial z_k} &= f\left(z_k\right)\left(z_k - M_k\right) \\
\left(F\left(z_{k+1}\right) - F\left(z_k\right)\right) \frac{\partial M_{k+1}}{\partial z_k} &= f\left(z_k\right)\left(M_{k+1} - z_k\right)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial Q}{\partial z_k} &= f\left(z_k\right)\left(S\left(p_k\right) - S\left(p_{k+1}\right)\right) \\
&+ f\left(z_k\right)\left[S'\left(p_k\right)\left(M_k - z_k\right) + S'\left(p_{k+1}\right)\left(M_{k+1} - z_k\right)\right] \\
&+ \sum_{k=1}^{N} \left(F\left(z_k\right) - F\left(z_{k-1}\right)\right) S'\left(p_k\right) P'\left(Q\right) \frac{\partial Q}{\partial z_k},
\end{aligned}$$

we get

$$\frac{\partial Q}{\partial z_k} = \frac{f\left(z_k\right)\left[S\left(p_k\right) - S\left(p_{k+1}\right) + S'\left(p_k\right)\left(M_k - z_k\right) + S'\left(p_{k+1}\right)\left(M_{k+1} - z_k\right)\right]}{1 - \sum_{k=1}^{N} \left(F\left(z_k\right) - F\left(z_{k-1}\right)\right) S'\left(p_k\right) P'\left(Q\right)}.$$

The denominator is positive, since $S'\left(p_k\right) > 0$ and $P'\left(Q\right) < 0$, so $\partial Q/\partial z_k$ has the same sign as

$$S\left(p_k\right) - S\left(p_{k+1}\right) + S'\left(p_k\right)\left(M_k - z_k\right) + S'\left(p_{k+1}\right)\left(M_{k+1} - z_k\right),$$

and since $S\left(p_k\right) - S\left(p_{k+1}\right) = -\int_{p_k}^{p_{k+1}} S'\left(p\right) dp$ and $p_{k+1} - p_k = M_{k+1} - M_k$, equation (25) follows. □

Proof of Proposition 6

Letting $\alpha\left(z_k\right) = \frac{M_{k+1} - z_k}{M_{k+1} - M_k}$, we can rewrite equation (15) as

$$S\left(p_k\right) + \alpha\left(z_k\right)\left(S\left(p_{k+1}\right) - S\left(p_k\right)\right) = S\left(p_k\right) + \frac{\int_{p_k}^{p_{k+1}} S\left(p\right) - S\left(p_k\right) dp}{p_{k+1} - p_k},$$

so

$$\alpha\left(z_k\right)=\frac{\int_{p_k}^{p_{k+1}}\frac{S(p)-S(p_k)}{S(p_{k+1})-S(p_k)}dp}{p_{k+1}-p_k}. \tag{26}$$

To evaluate dQ/dz_k at the optimal thresholds $z_1, ..., z_k$, we rewrite equation (25) in a similar fashion using the expression for $\alpha\left(z_k\right)$ given by equation (26).

$$\begin{aligned}\frac{dQ}{dz_k} &= \alpha\left(z_k\right)\left(S'\left(p_{k+1}\right)-S'\left(p_k\right)\right)-\frac{\int_{p_k}^{p_{k+1}}\left(S'\left(p\right)-S'\left(p_k\right)\right)dp}{p_{k+1}-p_k}\\ &= \frac{\int_{p_k}^{p_{k+1}}\frac{S(p)-S(p_k)(S'(p_{k+1})-S'(p_k))}{S(p_{k+1})-S(p_k)}dp-\int_{p_k}^{p_{k+1}}\left(S'\left(p\right)-S'\left(p_k\right)\right)dp}{p_{k+1}-p_k},\end{aligned}$$

so a sufficient condition for dQ/dz_k to be positive (resp., negative) is that

$$\frac{\left(S\left(p\right)-S\left(p_k\right)\right)\left(S'\left(p_{k+1}\right)-S'\left(p_k\right)\right)}{S\left(p_{k+1}\right)-S\left(p_k\right)}-\left(S'\left(p\right)-S'\left(p_k\right)\right)>0\left(resp.,<0\right),$$

or, equivalently,

$$\frac{S'\left(p_{k+1}\right)-S'\left(p_k\right)}{S\left(p_{k+1}\right)-S\left(p_k\right)}-\frac{S'\left(p\right)-S'\left(p_k\right)}{S\left(p\right)-S\left(p_k\right)}>0\left(resp.,<0\right). \tag{27}$$

A sufficient condition for this equation to hold is that

$$\frac{S'\left(p\right)-S'\left(p_k\right)}{S\left(p\right)-S\left(p_k\right)} \tag{28}$$

increasing (resp., decreasing) in p (for all $p>p_k$). The derivative of (28) with respect to p has the sign of

$$\begin{aligned}&S''\left(p\right)\left(S\left(p\right)-S\left(p_k\right)\right)-S'\left(p\right)\left(S'\left(p\right)-S'\left(p_k\right)\right)\\ =&S''\left(p\right)\int_{p_k}^{p}S'\left(x\right)dx-S'\left(p\right)\int_{p_k}^{p}\left(S''\left(x\right)\right)dx,\end{aligned}$$

which in turn has the sign of

$$\frac{S''\left(p\right)}{S'\left(p\right)}-\frac{\int_{p_k}^{p}\frac{\left(S''\left(x\right)\right)}{S'\left(x\right)}S'\left(x\right)dx}{\int_{p_k}^{p}S'\left(x\right)dx}.$$

The second term is a weighted average of the coefficient of absolute risk aversion of S for values between p_k and p. So, if $S''\left(x\right)/S'\left(x\right)$ is increasing (resp., decreasing) in x, then this difference will be positive (resp., negative).