Check for updates

The Journal of Industrial Economics

WILEY

ORIGINAL ARTICLE OPEN ACCESS

Optimal Simple Ratings

Hugo Hopenhayn[1] | Maryam Saeedi[2]

[1]Economics Department, UCLA, Los Angeles, California, USA | [2]Tepper Business School, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

Correspondence: Maryam Saeedi (msaeedi@andrew.cmu.edu)

Accepted: 24 February 2026

Keywords: certification | consumer surplus | producer surplus | simple ratings

ABSTRACT

We study optimal simple rating systems that partition sellers into a finite number of tiers. We show that optimal ratings must be threshold partitions, and that for linear supply and Cournot competition with constant marginal cost, optimal thresholds solve a k-means clustering problem requiring only the quality distribution. For convex (concave) supply functions, optimal thresholds are higher (lower) than the k-means solution. For log-concave distributions, two-tier certification captures at least 50% of maximum welfare gains from full disclosure, with five tiers typically achieving over 90%. Applications to eBay and Medicare Advantage data illustrate our method.

JEL Classification: D21, D47, D60, D82, L11

1 | Introduction

Online platforms and rating agencies overwhelmingly rely on simple, coarse rating systems rather than detailed quality disclosure. Michelin awards at most three stars, Airbnb designates "Superhosts," eBay distinguishes between badged and non-badged sellers, and California's Department of Public Health assigns restaurants letter grades A through C.[1] Yet despite having access to detailed seller quality data, intermediaries consistently compress this information into a handful of discrete signals. This article addresses two questions arising from this design choice. First, if a platform restricts itself to a finite partition of the quality space, what is the welfare-maximizing mechanism? Second, how large is the "surplus gap" between these simple mechanisms and the first-best outcome?

To address these questions, we investigate the optimal design and performance of rating systems when the market designer is constrained to use a small number of ratings. We primarily consider a competitive market with heterogeneous sellers whose qualities are drawn from a known distribution, with an extension to the canonical model of Cournot competition with constant marginal costs.[2] Buyers value quality but cannot observe it directly; they rely on ratings provided by a welfare-maximizing platform. 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. Because ratings pool sellers into groups, they create opposing distortions within each tier: high-quality sellers underproduce relative to full information (since they receive a price reflecting average tier quality), while low-quality sellers overproduce. The platform's problem is to set thresholds that balance these effects, which depend on the two sources of heterogeneity in our model. Our analysis makes three main contributions.

We first show that any optimal simple rating must take the form of a *threshold partition*: sellers are monotonically ordered by quality and divided by cutoff points into rating tiers. While intuitive, this result is nontrivial; it establishes that rating schemes which pool nonadjacent quality levels are strictly dominated by monotone threshold partitions. The design problem thus reduces

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *The Journal of Industrial Economics* published by The Editorial Board of The Journal of Industrial Economics and John Wiley & Sons Ltd.

from choosing an arbitrary partition to selecting a finite vector of thresholds.

Second, we derive a necessary condition for these thresholds with an intuitive economic interpretation: at any optimal cutoff, the planner is indifferent between assigning the marginal seller to the tier above or below. Moving her up raises her output (she is now pooled with higher-quality sellers and receives a higher price), and the welfare gain from this increased output must exactly offset the increased production cost. This intuitive criterion provides good guidance in the design of ratings, as pooling is more costly in areas where there is more quality dispersion or where supply is more responsive to prices. For the case of linear supply—and, as we show, for Cournot competition with constant marginal cost—this condition results in a closed-form solution: the optimal thresholds coincide with the solution to a standard k-means clustering problem as introduced by MacQueen et al. [5] and used extensively in machine learning and statistics. Platforms can therefore compute optimal ratings using only the distribution of seller qualities, without knowledge of demand elasticities or other market primitives. For more general supply functions, the curvature of supply matters. The k-means thresholds provide pointwise lower bounds when supply is convex and upper bounds when supply is concave.

Third, we quantify the efficiency of coarse ratings. For the broad class of log-concave quality distributions—which includes the uniform, normal, exponential, logistic, and Laplace—a simple two-tier certification captures at least 50% of the maximum welfare gain achievable under full information disclosure. This bound is tight: the Laplace distribution achieves exactly 50 percent. In practice, actual performance usually exceeds this lower bound, with two tiers capturing 55%–77% of potential gains for most log-concave distributions. Five categories typically recover over 90% of the full-information surplus. This efficiency arises because the first few thresholds successfully separate the most distinct segments of the quality distribution, maximizing the variance decomposition between groups.

Beyond these main results, we characterize how optimal thresholds respond to the shape of the quality distribution. When the distribution is right-skewed, optimal certification becomes more stringent: a smaller fraction of sellers receives the top rating. This occurs because right skewness increases the quality of the certified pool, making the distortion from pooling a marginal seller upward relatively more costly. Both of our empirical applications—eBay seller quality and Medicare Advantage provider quality—exhibit left-skewed distributions, and in both cases, optimal certification includes approximately 65% of sellers, screening out the lower tail.

We also assess the robustness of the k-means thresholds when the true supply function is nonlinear. Numerical analysis shows that the linear thresholds perform remarkably well except when two conditions coincide: highly convex supply and the distribution of sellers' quality features long tails. In these cases, the optimal policy certifies a very small fraction of sellers who serve the majority of the market.

Our findings rationalize the widespread use of coarse ratings in practice. While simple mechanisms necessarily sacrifice some information, the welfare loss is low and diminishes rapidly as categories are added. Given the additional advantages of simplicity—transparency for consumers, lower cognitive costs, and reduced implementation burden—our analysis supports continued reliance on coarse ratings across many market settings. Concerning the practical aspect of rating design, the strong overall performance and modest data prerequisites of our linear thresholds offer a good starting point in design of rating systems.

1.1 | Related Literature

This article uses the canonical model we developed in Hopenhayn and Saeedi [6], where we study optimal information disclosure in a theoretical framework. In this article, we use this model to answer a very different question, motivated by the wide use of simple ratings. While our previous paper provides broad qualitative results (e.g., the optimality of interval ratings), this one goes much further, giving a detailed characterization for the design of optimal ratings with a finite partition. We provide an algorithm that solves for the optimal thresholds in the case of linear supply or Cournot and gives a benchmark more generally and illustrate our results in two applied settings. Here, we want to find the value loss that arises as a result of coarse rating instead of using the optimal unconstrained ratings, which is usually the aim of theoretical papers, such as our own previous work and many others. A lot of empirical work on ratings points to the use of coarse ratings in practice, Vatter [4], Saeedi [7], Elfenbein et al. [8], Fan et al. [9], and Jin and Leslie [10].[3] Other empirical papers analyze the effects of changes in marketplace feedback mechanisms on price and quality (e.g., Hui et al. [12], Filippas et al. [13], and Nosko and Tadelis [14]); Hui et al. [15] study the benefit of adding a second tier to the reputation mechanism to mitigate the cold-start problem and to promote entry. Even though the coarse ratings are very commonplace in practice and have been studied very extensively in empirical studies, they are not the outcome of most canonical theory papers.

In our study, we show that the losses due to coarse ratings can be low, and as a result, they might be justified because of the simplicity they provide over a more complex mechanism. Our results parallel the results of the coarse matching literature. Wilson [16] 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 [17], Hoppe et al. [18], and Shao [19].

Subsequent work has examined message-constrained information structures in persuasion environments. In particular, Lyu et al. [20] develop a general analysis of coarse information design with finitely many messages. Our results anticipate and complement this line: we establish implementable rating thresholds in a market-equilibrium setting with heterogeneous producers and endogenous supply, show that in the linear/Cournot benchmarks these thresholds coincide with k-means partitions, and quantify welfare performance. Their characterization extends the message-constraint perspective in a persuasion framework, while our focus is on market design and welfare in equilibrium. Complementary to our setting, Aybas and Turkel [21] analyze

14676451, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/joie.70025 by Maryam Saeedi - Carnegie Mellon University , Wiley Online Library on [12/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

persuasion with coarse communication, providing a geometric characterization under finite messages and tight bounds on the marginal value of added capacity.

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. Section 5 concludes the paper. Proofs are relegated to Appendix A unless otherwise specified.

2 | The Model and Equilibrium

In this section, we introduce the market environment, describe how information and ratings enter the model, characterize competitive equilibrium given a rating distribution, and derive the planner's objective.

2.1 | Environment

We consider a market with vertically differentiated products where quality is the only dimension of heterogeneity on the supply side. There is a unit mass of sellers, each endowed with a quality parameter $z \in \mathbb{R}$, distributed according to a continuous cumulative distribution function (cdf) $F(z)$ with density $f(z)$ on a support $[z_0, z_N] \subseteq \mathbb{R}$, where the bounds may be infinite.

All sellers share the same technology. Producing quantity $q \geq 0$ entails cost $c(q)$, where c is differentiable, strictly increasing, and strictly convex. The corresponding supply function $S(p)$ is the inverse of the marginal cost function, formally defined by the first-order condition $c'(S(p)) = p$. Given our assumptions on the cost function, $S(p)$ is strictly increasing and at least twice continuously differentiable.

On the demand side, there is a mass M of consumers who face a discrete choice between buying one unit of the good or taking an outside option. We index consumers by a taste parameter $\theta \geq 0$, distributed according to a continuous and strictly increasing cdf $\Psi(\theta)$ with support contained in $[0, \overline{\theta}]$ (possibly unbounded). The indirect utility of a consumer with type θ who buys from a seller of quality z at price p is

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

while the outside option is normalized to zero. The quasi-linear form implies that all consumers value quality increments equally: a one-unit increase in z increases utility by one unit for everyone. The linearity of U in z implies that consumers care only about the *expected* quality of the good they buy. Throughout the analysis we therefore use z to denote *expected* quality as perceived by consumers.[4]

The critical feature of our model is that quality z is not directly observable to consumers. Instead, a platform or certifying agency—henceforth *planner*—provides information about seller qualities through a rating system which they can commit to. Following the standard representation in information-design problems, any information structure can be represented by the induced distribution of posterior mean qualities, which we denote by G.[5]

The timing is as follows:

1. **Information provision stage**: The planner observes the true distribution of seller qualities $F(z)$ and implements a rating system that provides signals to consumers about seller quality.
2. **Belief formation stage**: Based on the rating system, consumers form posterior beliefs about each seller's quality. When ratings pool multiple sellers together, consumers know only the expected quality within each rating category.
3. **Market equilibrium stage**: Given consumer beliefs about quality, the market reaches a competitive equilibrium where prices adjust to clear markets and sellers choose their optimal production quantities.

Thus, the information structure implemented by the planner transforms the original distribution of true qualities $F(z)$ into a distribution of expected qualities $G(z)$. In particular, in a finite rating system, sellers are partitioned into finitely many groups, and buyers only observe the rating (group) to which each seller belongs. Within each group, buyers form a posterior mean about quality; hence G is discrete, with point masses at the conditional means of quality in each rating group.

2.2 | Equilibrium Analysis

Fix a distribution of perceived (expected) qualities G, in equilibrium, prices arbitrage away the differences in expected utility across sellers. Thus, all sellers with positive sales earn the same utility for consumers. Considering two sellers with expected qualities z and z' sold at prices p and p', we will have: $p' - p = z' - z$. Thus equilibrium prices must be of the form

$$p(z) = p(0) + z,$$

where $p(0)$ is the (endogenous) price of a hypothetical good of expected quality zero. We refer to $p(0)$ as the *baseline price*.

The baseline price $p(0)$ determines the market size, where the marginal consumer's θ is found by setting $U(0, \theta, p(0)) = 0$, or simply $\theta = p(0)$. All consumers with $\theta \geq p(0)$ will make their unit purchase. Hence, the aggregate demand is

$$Q^d = M(1 - \Psi(p(0))).$$

Inverting this relationship gives us the *inverse demand* as a function of the total quantity Q:

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

We interpret $P(Q)$ as the inverse demand function evaluated at the hypothetical quality-zero good. Given Q, the equilibrium price for a seller with expected quality z is therefore

$$p(z) = P(Q) + z.$$

14676451, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/joie.70025 by Maryam Saeedi - Carnegie Mellon University , Wiley Online Library on [12/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

On the supply side, each seller with expected quality z chooses its output, $q = S(p(z))$, thus aggregate supply equals $Q^s = \int S(p(z))dG(z)$. Market equilibrium requires supply to equal demand, $Q^s = Q^d = Q$, leading to the equilibrium condition:

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 the 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 [6] 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.[6] The second assumption rules out no output as an equilibrium. Although when it exists, the corner equilibrium is also unique, we rule this out as a matter of convenience.

2.3 | Welfare Analysis

To evaluate different rating systems, we need to specify the welfare criterion. We adopt a total surplus standard, measuring the sum of consumer and producer surplus without distributional weights. This welfare standard is appropriate when the planner seeks efficiency. We now derive consumer surplus, producer surplus, and total surplus, making explicit how each depends on the rating distribution G.

2.3.1 | Consumer Surplus

Consider a consumer of type θ who buys a good of expected quality z and receives utility $\theta + z - p(z)$. Given the equilibrium price $p(z) = P(Q) + z$, the consumer's net utility is $\theta - P(Q)$. Notice that the consumer's surplus is independent of which product they purchase—the quality premium in price exactly offsets the quality benefit. The consumer purchases if and only if $\theta - P(Q) \geq 0$, or equivalently, $\theta \geq P(Q)$. Hence total consumer surplus is

$$M \int_{P(Q)} (\theta - P(Q))d\Psi(\theta).$$

To evaluate this integral, we use the change of variables $x = M[1 - \Psi(\theta)]$. When $\theta = P(Q)$, we have $x = Q$ (the total mass of purchasing consumers), and when $\theta \to \infty$, we have $x \to 0$. Substituting:

$$CS = \int_0^Q (P(x) - P(Q))dx,$$

where $P(x) = \Psi^{-1}(1 - x/M)$. Thus, consumer surplus is monotone in market size Q. Importantly, consumer surplus depends on the information structure only through its effect on total quantity Q. Better information does not directly benefit consumers through better matching (since prices adjust to reflect quality differences) but only indirectly if it expands market participation.

2.3.2 | Producer Surplus

Given G and the induced equilibrium $Q(G)$, a seller with expected quality z chooses $q(z) = S(P(Q(G)) + z)$ and earns profits

$$\pi(z) = (P(Q(G)) + z)\ q(z) - c(q(z)).$$

Producer surplus is simply total profits:

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

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. This equilibrium effect has opposite impacts on consumer and producer surplus: whereas 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.

2.3.3 | Total Surplus

We take the planner's objective to be the maximization of total surplus, the unweighted sum of consumer and producer surplus. Thus, total surplus is then given by

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

This simplification is by canceling the term involving $P(Q)$ from consumer and producer surplus which is equal to $P(Q)Q$. 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 | Design of Optimal Simple Ratings

In this section, we focus on the design of simple ratings where the planner can use limited number of signals to maximize total

14676451, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/joie.70025 by Maryam Saeedi - Carnegie Mellon University , Wiley Online Library on [12/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

surplus. As shown in Hopenhayn and Saeedi [6], Proposition 3i, the unconstrained optimal mechanism in this setting is full revelation of quality.

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.

3.1 | The Structure of Optimal Ratings

Before analyzing specific rating designs, we establish a fundamental structural property that dramatically simplifies the optimization problem. Among all possible ways to partition sellers into N groups, which structure maximizes welfare?

Proposition 1. *An optimal simple rating system with N categories must take the form of a threshold partition that sorts sellers monotonically by their quality.*

This structural result is powerful: it reduces the infinite-dimensional problem of choosing an optimal partition to the finite-dimensional problem of choosing $N-1$ threshold values $z_1, z_2, \ldots, z_{N-1}$ that divide sellers into quality intervals $[z_0, z_1], [z_1, z_2], \ldots, [z_{N-1}, z_N]$.[7] Henceforth we restrict attention to such threshold partitions.

3.2 | Necessary Conditions for Optimal Thresholds

In this section, we derive a simple and intuitive necessary condition to characterize these optimal thresholds. Let $\mathbf{z} = (z_1, z_2, \ldots, z_{N-1})$ denote the vector of thresholds; and $M_k = m(z_{k-1}, z_k), k = 1, \ldots, N$ denote the conditional means of sellers' quality within the interval $[z_{k-1}, z_k]$; and $Q(\mathbf{z})$ denote the unique equilibrium total quantity. Under this rating system, all sellers in group k are perceived to have expected quality M_k, face the same price p_k, and choose the same quantity q_k, where

$$p_k = P(Q(z)) + M_k, \qquad q_k = S(p_k).$$

From Equation (3), it follows that the total surplus for partition $\mathbf{z}$ is given by

$$TS(\mathbf{z}) = \int_0^{Q(z)} P(x)dx + \sum_{k=1}^{N} \left[F(z_k) - F(z_{k-1})\right]\left[M_k q_k - c(q_k)\right]. \tag{4}$$

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

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

$$\left(P(Q(\mathbf{z})) + z_k\right)\left(q_{k+1} - q_k\right) = c(q_{k+1}) - c(q_k) \tag{5}$$

for all k.

Condition (5) has an intuitive interpretation. Consider a marginal seller with quality z_k; For this threshold to be optimal, the planner should be indifferent between assigning this seller to group k or $k+1$. If we move this seller from the lower group to the higher one, her expected quality (and price) increases from M_k to M_{k+1}, and her output changes from q_k to q_{k+1}. Because the mass of this marginal seller is infinitesimal, we can apply an envelope argument to ignore the induced changes in other sellers' outputs and in $Q(z)$, except through their first-order effect on the baseline price. The left-hand side of (5) shows the marginal *benefit* of this move 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 *cost*, namely the change in production cost when the marginal seller increases her output from q_k to q_{k+1}. At an optimum, these two effects must balance, yielding condition (5). 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

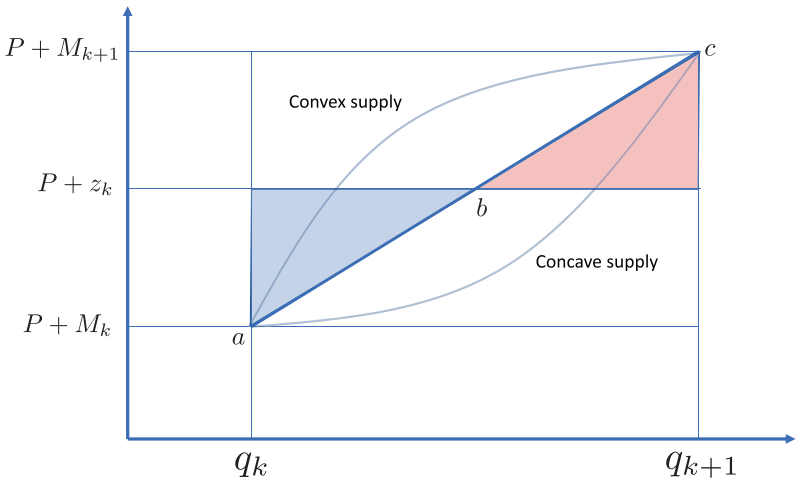

FIGURE 1 | Optimal threshold condition. *Notes:* [Colour figure can be viewed at wileyonlinelibrary.com]

14676451, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/joie.70025 by Maryam Saeedi - Carnegie Mellon University , Wiley Online Library on [12/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

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).[8] The area below the marginal cost function, that is, the supply function, between q_k and q_{k+1} equals the right-hand side of (5), whereas the area under the line $P + z_k$ equals the left-hand side of the equation. The difference between these two areas is

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

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 $z_k = \big(M_{k+1} + M_k\big)/2$. It also follows easily that for the convex supply case, $P + z_k$ must be higher so the two corresponding areas are equal, whereas the converse holds for the concave supply case. These results are summarized in the following lemma.

Lemma 2. *Let the thresholds $\boldsymbol{z} = \big(z_1, \ldots, z_{N-1}\big)$ maximize (3), and denote by $M_k = E\big(z|z_{k-1} \le z \le z_k\big)$ the corresponding conditional means. Then*

1. *$z_k = \big(M_k + M_{k+1}\big)/2$ if the supply function $S(p)$ is linear;*
2. *$z_k > \big(M_k + M_{k+1}\big)/2$ if the supply function is strictly convex in the interval $\big[M_k, M_{k+1}\big]$; and*
3. *$z_k < \big(M_k + M_{k+1}\big)/2$ if the supply function is strictly concave in the interval $\big[M_k, M_{k+1}\big]$.*

3.3 | The Linear Supply Benchmark

The linear supply function helps us simplify the profit function for the sellers. Consider the linear supply function, $S(p) = \alpha p$ for some constant $\alpha > 0$, implying a quadratic cost function $c(q) = q^2/(2\alpha)$. In this case, the profit function for a seller with expected quality z facing price $p = P(Q) + z$ is: $\pi(z) = [P(Q) + z]^2\alpha/2$.

Therefore, for any distribution G of observed qualities, total profits are

$$\begin{aligned}\Pi &= \frac{\alpha}{2}\int (P(Q) + z)^2 dG(z)\\ &= \frac{\alpha}{2}P(Q)^2 + \alpha P(Q)\overline{z} + \frac{\alpha}{2}\int z^2 dG(z),\end{aligned} \tag{6}$$

where $\overline{z} = \int z\, dG(z)$ is the mean expected quality, which equals the mean of the prior F and does not depend on G. In addition, when the supply function is linear, the equilibrium condition in Equation (2) reduces to

$$Q = \alpha \int (P(Q) + z)dG(z) = \alpha P(Q) + \alpha\overline{z},$$

As a result, Q (and thus consumer surplus) is independent of the information structure. Consequently, maximizing total profits as in Equation (6) is equivalent to maximizing $\int z^2\, dG(z)$. 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 $\boldsymbol{z} = \big(z_1, \ldots, z_{N-1}\big)$ are the ones that minimize*

$$\sum_{k=1}^{N}\int_{z_{k-1}}^{z_k} \big(z - M_k\big)^2 dF(z). \tag{7}$$

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

Problem (7) is precisely the standard k-means clustering objective used in statistics and machine learning: it minimizes the sum of squared deviations of the data points (here, seller qualities) from their group means. Proposition 2 shows that under linear supply, optimal simple ratings coincide with optimal k-means clusters in quality space. When F has a log-concave density, uniqueness follows from the results in MacQueen et al. [5].

This connection is practically important: it implies that platforms can compute optimal ratings using only the distribution of seller qualities and off-the-shelf clustering algorithms, without needing to know the full market structure.

3.4 | Supply Curvature and Threshold Ordering

Lemma 2 shows that, for a given pair of adjacent groups, the linear–supply thresholds lie between the optimal thresholds under strictly convex and strictly concave supply: with convex (concave) supply, the cutoff moves to the right (left) relative to the midpoint of the two group means. That argument applies locally to each pair of adjacent groups. We now extend this insight to a global ordering result: *all* optimal thresholds under convex (concave) supply are simultaneously higher (lower) than the k-means benchmarks from Proposition 2.

Proposition 3. *Suppose the quality distribution $F(z)$ has a log-concave density. Let $\big(z_1^L, \ldots, z_{N-1}^L\big)$ be the optimal thresholds for the linear supply case. The optimal vector of thresholds $\big(z_1, \ldots, z_{N-1}\big)$ for a convex (resp., concave) supply function is pointwise higher (resp., lower) than $\big(z_1^L, \ldots, z_{N-1}^L\big)$.*

This result has immediate practical value. The k-means thresholds from Proposition 2 depend only on the distribution of qualities, yet they provide a systematic lower (resp., upper) bound on the optimal thresholds whenever the supply function is convex (resp., concave). Platforms such as eBay or Airbnb can therefore compute the k-means thresholds using only the cross-sectional distribution of seller qualities, without knowing demand or cost parameters, and treat them as a starting benchmark. As we show in Section 4.4, these linear thresholds perform remarkably well as a heuristic even when the true supply function is not linear.

We can extend the comparison beyond the linear benchmark to rank thresholds across two arbitrary nonlinear supply functions, provided demand is perfectly elastic (i.e., P is exogenous). To formalize "more convex," we introduce the following order over supply functions.

14676451, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/joie.70025 by Maryam Saeedi - Carnegie Mellon University , Wiley Online Library on [12/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Definition. We call supply function S_B to be *more convex* than S_A if $S_B(p) = \phi(S_A(p))$ for some strictly increasing, convex function ϕ.

Thus S_B is obtained from S_A via a convex transformation of quantities. The next proposition shows that when the base price level is fixed, making supply more convex shifts all optimal thresholds to the right.

Proposition 4. *Suppose the quality distribution F has a log-concave density and demand is perfectly elastic (i.e., the base price P is exogenous). Let S_A and S_B be two supply functions such that S_B is more convex than S_A. Then the optimal thresholds for S_B are pointwise larger than for S_A: $z_k(S_B) \geq z_k(S_A)$ for all k.*

The intuition connects to the graphical analysis in Figure 1. The optimal threshold z_k equates the benefit of moving a marginal seller upward (the area under $P + z_k$) with the cost (the area under the supply curve between q_k and q_{k+1}). When supply is more convex, the area under the supply curve between any two quantities shifts upward—the "cost triangle" in Figure 1 grows. To restore equality, $P + z_k$ must also increase, which requires a higher threshold z_k. Under perfectly elastic demand, P is fixed, so this comparison is valid across supply functions. The supermodular structure ensures this local intuition extends to a global pointwise ordering across all thresholds simultaneously.

Together, Propositions 3 and 4 paint a consistent picture: as supply curvature increases, optimal thresholds shift monotonically to the right, and the linear (k-means) benchmark becomes a progressively weaker lower bound. Section 4.4 quantifies this relationship numerically.

3.5 | Quality Skewness and Certification Stringency

The optimal thresholds, as shown in Equation (7), depend on quality distribution F. For linear supply, we study how skewness in the quality distribution affects these optimal choices. We first analyze the simple case of a two-tier certification and show that the optimal threshold moves in the same direction as the skewness of the distribution: right-skewed distributions yield higher thresholds, while left-skewed distributions yield lower thresholds. We then extend this result and provide comparative statics for the full vector of thresholds under an appropriately skewness ordering.

Let us first provide some intuition. Consider a single certification threshold z^*. The planner faces a trade-off when deciding how stringent the certification cutoff should be. When assigning z^* to the upper group, there is an upward distortion of the supply of the seller at z^*, which depends on the gap $M_H - z^*$. This gap also measures the gain the seller receives from being pooled with higher-quality sellers. When assigning z^* to the lower group, there is a downward distortion of the supply of the seller at z^*, which depends on $M_L - z^*$. This distance also measures the loss of the marginal seller. Right (resp., left) skewness in the distribution $F(z)$ increases (resp., decreases) the upward distortion and decreases (resp., increases) the downward distortion, making it optimal to have more (resp., less) stringent certification threshold.

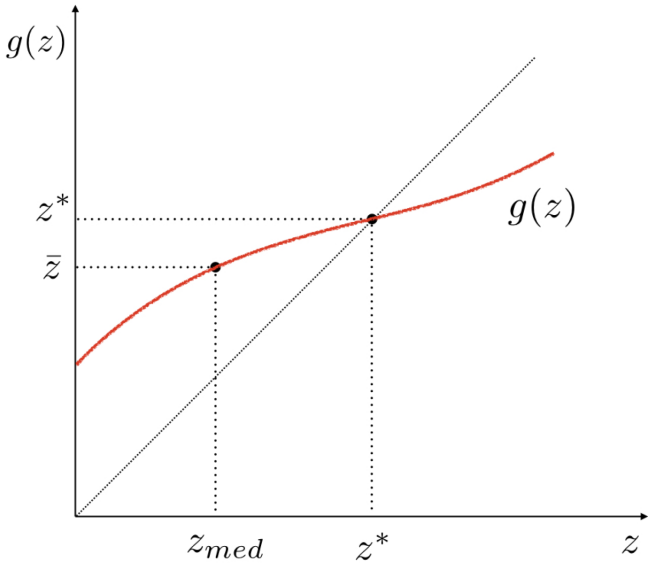

FIGURE 2 | z^* when the mean is greater than the median. *Notes:* [Colour figure can be viewed at wileyonlinelibrary.com]

Under linear supply, Proposition 2 implies that the optimal threshold solves

$$z^* = \frac{1}{2}\big(M_L(z^*) + M_H(z^*)\big). \tag{8}$$

This condition links the optimal threshold to the distribution's shape. First consider a symmetric distribution (i.e., where the median, z_{median}, equals the mean, $\bar{z}$). Given that any z^*, $F(z^*)M_L + (1 - F(z^*))M_H = \bar{z}$, setting the threshold $z^* = \bar{z} = z_{median}$ satisfies the above condition.

When the distribution is skewed, the threshold shifts in the same direction as the skewness. This can be proved as follows. Consider a right-skewed distribution where $\bar{z} > z_{median}$. Let $M_L(.)$ and $M_H(.)$ denote the conditional mean of the seller quality below and above any given value, respectively. Furthermore, denote $g(z) = \frac{1}{2}\big(M_L(z) + M_H(z)\big)$. Following Proposition 2, 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 $z = z_{median}$, $g(z) = \bar{z} > z$. Because the function $g(z)$ is increasing and continuous, the unique fixed point z^* must be to the right of z_{median}, and hence, $z^* > \bar{z}$, as illustrated in Figure 2. The result for the case of left skewness can be shown similarly. Consistently with these results, the empirical examples in Section 4.5 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 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,

14676451, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/joie.70025 by Maryam Saeedi - Carnegie Mellon University , Wiley Online Library on [12/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 | Optimal thresholds.

Distribution	Case	Mean/median	z^*	Certified
Pareto	$\alpha = 3$	1.19	2.73	5
	$\alpha = 4$	1.12	1.84	9
Exponential	All	1.45		20
$F(z) = z^\alpha$	$\alpha = 0.5$	1.32	0.41	36
$z \in [0, 1]$	$\alpha = 2$	0.94	0.62	62
Log-normal	$\sigma = 0.25$	1.03	1.09	36
$(\mu = 0)$	$\sigma = 1$	1.64	4.25	7

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

only a small fraction should get certified, 5% when the power parameter is 3, and 9% when the power parameter is 4.[9] 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 [25].

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 $\tilde{F}(g(x)) = F(x)$.

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 $\tilde{F}\left(x^2\right) = x$ 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.*

This proposition implies that for all k, $\tilde{F}\left(g\left(z_k\right)\right) = F\left(z_k\right) > F\left(l_k\right)$, 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 $\alpha = 2$.[10] Consistently with the previous proposition, the share of certified sellers is lower when $\alpha = 0.5$.

3.6 | Extension to Imperfect Competition

Up to now, we have assumed perfect competition. We now show that the k-means characterization in Proposition 2 extends to a canonical imperfect-competition model: Cournot competition with constant marginal cost.

Consider n firms indexed by $i = 1, \ldots, n$, each with expected quality z_i. Aggregate quantity is $Q = \sum_{i=1}^{n} q_i$. Inverse demand is $P(Q)$ as before, and each firm faces constant marginal cost c. Firm i chooses quantity q_i to maximize

$$\pi_i(q_i; Q, z_i) = [P(Q) + z_i - c]q_i.$$

Proposition 6. *Under Cournot competition with constant marginal cost, optimal rating thresholds coincide exactly with the k-means solution, independent of the number of firms n or the demand function $P(Q)$.*

The intuition for this equivalence is that, in this case, total quantity is once again independent of the specific information provided. Thus, maximizing total surplus is equivalent to maximizing total producer surplus. Additionally, given that quantity is independent of the information structure because of constant marginal cost, the only term in total producer surplus that depends on the information structure G is $\int z^2 \, dG(z)$, which is the same as the perfect competition with linear supply.

This equivalence result is powerful—it shows that our k-means characterization applies not just to perfect competition with linear supply, but also to the canonical model of imperfect competition. This broadens the applicability of our computational approach significantly.

4 | Performance of Simple Ratings

In this section, we evaluate welfare loss when the planner is restricted to *simple* rating systems—partitions with a small number of categories—rather than the unconstrained optimal information structure.

Throughout this section, we work in the *linear benchmark* introduced earlier: either (i) perfect competition with a linear supply function, or (ii) Cournot competition with constant marginal cost. As shown in Section 3, these two environments are equivalent for our purposes: in both cases (1) total quantity Q and hence consumer surplus are independent of the information structure, and (2) maximizing total surplus is equivalent to maximizing total producer surplus, which in turn is equivalent to maximizing $\int z^2 \, dG(z)$, where G is the distribution of expected qualities induced by the rating system.

We first formalize a welfare-performance measure for a generic rating system and relate it to a simple variance decomposition. We then use this representation to derive theoretical lower bounds for two-tier certification schemes, show numerically how quickly the performance of simple ratings converges to the full-information benchmark as we increase the number of categories, and numerically study how well the linear thresholds perform when the true supply function is nonlinear. The section closes with two empirical illustrations.

4.1 | Performance Measure

Total surplus is maximized with full information, as shown in Hopenhayn and Saeedi [6], Proposition 3(i). Our measure of performance of a simple rating system is the fraction of the gap between the full-information (unconstrained optimal) and no-information (no intervention) cases that is closed with this

14676451, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/joie.70025 by Maryam Saeedi - Carnegie Mellon University , Wiley Online Library on [12/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

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 recovered by the simple rating, using Equation (6). Under full information, total profits are $\frac{1}{2}P(Q)^2 + P(Q)\overline{z} + \frac{1}{2}\int z^2 dF(z)$. Therefore, the surplus gap with respect to the full-information case is

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

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

$$\overline{\Delta\Pi} = \frac{1}{2}\left(\int z^2 dF(z) - \overline{z}^2\right). \tag{9}$$

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

$$\begin{aligned}\Delta\Pi &= \frac{1}{2}\sum_{k=1}^{N}\int_{z_{k-1}}^{z_k}\left(z^2 - M_k^2\right)dF(z)\\ &= \frac{1}{2}\sum_{k=1}^{N}\int_{z_{k-1}}^{z_k}\left[\left(z - M_k\right)^2\right]dF(z).\end{aligned} \tag{10}$$

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 $[z_{k-1}, z_k]$. We are interested in seeing how much of the possible total surplus gain from information, as expressed in Equation (9), is captured by Equation (10), or, simply, the following ratio:

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

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.

4.2 | Theoretical Bounds

The simplest coarse rating scheme is a two-tier certification, widely used in practice. The next proposition provides a useful bound for the gains from certification that builds on the variance decomposition described earlier. 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 7. *When the supply function is linear, the relative performance of a two-tier setting satisfies*

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

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

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

Proof. From a well-known result in Stoyan and Daley [26 pp. 16–19], the conditions of this corollary imply that $cv_1 < 1$ and $cv_2 < 1$. Using the bound in Proposition 7 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 (Laplace), logistic, extreme value, and many others with some restriction on parameters (e.g., power function $F(z) = z^c$ for $c \geq 1$.) Laplace distribution achieves this bound when the threshold is set at the mean.[12]

Related bounds for two-sided matching problems can be found in McAfee [17], Hoppe et al. [18], and Shao [19]. The results of Wilson [16] imply that the losses from N-ratings are of order $1/N^2$.

This proposition, in addition to the corollary, has an important practical implication, for a wide range of distributions, certification can achieve at least half of the total possible surplus. Proposition 7 suggests that even using just the median as the certification threshold can be enough to reach this level. In the next section, we numerically show that these losses decrease rapidly as the number of signals increases. In Section 4.4, we show that even though our result here is for the linear supply case, or Cournot with fixed marginal cost, using the same thresholds for nonlinear supply achieves a significant share of surplus.

4.3 | Numerical Results

We now examine the numerical results for a variety of distribution functions often used in the economics literature. Table 2 reports the share of the total surplus gap that is closed with partitions of different sizes n when the supply is linear. 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 Corollary 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.

14676451, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/joie.70025 by Maryam Saeedi - Carnegie Mellon University , Wiley Online Library on [12/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 | Surplus gap.

Distribution	Case	% Surplus gap closed			
		$n = 2$	$n = 3$	$n = 5$	$n = 10$
Pareto	$\alpha = 3$	46	68	84	94
	$\alpha = 4$	54	74	89	97
Exponential	All	65	82	93	98
$F(z) = z^{\alpha}$	$\alpha = 0.5$	77	90	97	99
$z \in [0, 1]$	$\alpha = 2$	72	87	95	99
Log-normal	$\sigma = 0.25$	63	81	92	98
$(\mu = 0)$	$\sigma = 1$	55	75	89	97

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

The numerical results and the previous theoretical bounds suggest that in real world settings, we may not need very complicated mechanisms to attain most of the benefits from information provision. In this article, 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.

4.4 | Performance of Linear Thresholds for Nonlinear Supply

The thresholds derived above for the case of linear supply are easy to calculate and depend only on the distribution of qualities. We now ask how well these "linear" thresholds perform when the true supply function is nonlinear. To do so, we take the thresholds obtained under the linear benchmark (or, equivalently, Cournot competition with constant marginal cost) and apply them in environments where individual supply is given by

$$s(p) = p^{\theta}$$

for $\theta \in \left\{\frac{1}{4}, \frac{1}{2}, 2, 4\right\}$.

In line with Proposition 4, the surplus ratio falls as the curvature parameter increases: greater convexity pushes the optimal thresholds toward higher qualities, making the linear (k-means) thresholds a progressively poorer approximation. Neverthelese, for the moderate levels of curvature we consider, they still capture most of the optimal surplus.

For each distribution F considered earlier and each curvature parameter θ, we compute three objects: (i) the *optimal* thresholds that maximize total surplus given $s(p) = p^{\theta}$, and (ii) the *linear* thresholds obtained under $s(p) = p$ and (iii) the relative surplus gap closed using linear supply and compared to the optimal thresholds. Specifically, the relative gap is:

$$\frac{\overline{\Delta\Pi} - \Delta\Pi^{l}}{\overline{\Delta\Pi} - \Delta\Pi^{o}},$$

where, $\Delta\Pi^{l}$ is the surplus gap closed if the thresholds for linear supply are used, and $\Delta\Pi^{o}$ is the surplus gap closed using the optimal thresholds, $\overline{\Delta\Pi}$ is total possible surplus gain using full information as defined before.

Table 3 reports, for each (F, θ) pair, the fraction of firms that are certified under the optimal and linear thresholds, together with the relative surplus gap.[13] Remarkably, the performance of the linear thresholds is very close to the optimal ones in most cases, with the relative surplus gap typically above 95%. The only clear exceptions occur for Log-normal $(0, 1)$ when $\theta = 2$ and $\theta = 4$, and for the Exponential distribution when $\theta = 4$.

These three cases correspond to situations where supply is strongly convex and very elastic, and the distribution of qualities is right-skewed. As a result, an increase in thresholds leads to a large redistribution of output to certified firms and also an increase in total output. It is optimal then to certify a small fraction of firms that end up accounting for a large share of output. Note that in the log-normal case, when $\theta = 2$, it is optimal to certify only 0.6% of firms, while 7.4% are certified with the linear threshold. Moreover, in the linear case, the solution is not interior, and the market share of certified firms equals one. The same occurs when $\theta = 4$, where in both the optimal and linear case only certified firms produce, and it is optimal to certify a negligible fraction of firms. Similarly, for the exponential case with $\theta = 4$, it is optimal to certify only 0.2% of the firms, while 20.3% are certified with the linear threshold, and for both thresholds the market share of certified firms equals one.

Consistent with the comparative statics in Propositions 3 and 4, these are precisely the parameter configurations in which higher supply curvature pushes the optimal thresholds furthest to the right, so the fixed linear (k-means) thresholds provide the weakest approximation and the surplus gap is lowest. Overall, however, our results suggest that except in cases where supply is very elastic, and the quality distribution has a long right tail, linear thresholds perform very well relative to the optimal ones.

14676451, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/joie.70025 by Maryam Saeedi - Carnegie Mellon University , Wiley Online Library on [12/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 | Performance of linear supply thresholds on nonlinear supply functions.

	$\theta = 1/4$			$\theta = 1/2$		
	% certified		Relative surplus gap	% certified		Relative surplus gap
Distribution	**Optimal**	**Linear**	**Linear over optimal**	**Optimal**	**Linear**	**Linear over optimal**
Pareto $\alpha = 3$	15.3	4.9	97.6	8.2	4.9	97.5
Pareto $\alpha = 4$	12.5	8.7	97.9	11.3	8.7	99.1
Exponential	31.0	20.3	95.4	27.0	20.3	97.9
$F(z) = z^{1/2}$	41.1	36.0	98.7	38.9	36.0	99.6
$F(z) = z^2$	66.2	61.8	99.3	64.5	61.8	99.7
Log-normal $(0, 0.25)$	40.5	36.5	98.9	39.4	36.5	99.7
Log-normal $(0, 1)$	17.5	7.41	90.7	14.0	7.41	95.2
	$\theta = 2$			$\theta = 4$		
Pareto $\alpha = 3$	—	—	—	—	—	—
Pareto $\alpha = 4$	4.8	8.7	93.9	—	—	—
Exponential	9.8	20.3	88.9	0.4[a]	20.3[a]	33.4
$F(z) = z^{1/2}$	31.6	36.0	99.0	24.4	36.0	92.1
$F(z) = z^2$	57.3	61.8	99.2	49.7	61.8	94.8
Log-normal $(0, 0.25)$	30.8	36.5	99.0	20.0	36.5	91.7
Log-normal $(0, 1)$	0.6	7.4[a]	62.1	$\simeq 0$[a]	7.4[a]	0.4

[a]Market share of certified = 1.

4.5 | Empirical Application

Finally, we illustrate our method in two empirically motivated applications. The goal of these exercises is not to provide fully realistic structural estimates of the underlying markets, but rather to show how the linear benchmark and the k-means solution can be used to design simple rating systems when only the cross-sectional distribution of qualities is observed.

In both applications, we interpret the results under the linear benchmark (or equivalently, Cournot with constant marginal cost). Thus, the thresholds we compute should be understood as optimal simple ratings in that benchmark and as a starting point for more detailed market-specific analyses.

4.5.1 | Medicare Advantage Providers

We first consider the market for Medicare Advantage (MA) plans. We use data, generously shared by Benjamin Vatter, on estimated quality scores for MA products across the United States. The details of how these quality measures are constructed are described in Vatter [4]; here we treat them as our quality variable z.

We proceed as follows. We pool all plans nationally into a single cross-sectional distribution of qualities, construct a smoothed approximation of this distribution using kernel density estimation (Figure 3), and apply our two-tier k-means procedure under linear supply (or Cournot with constant marginal cost) to compute the optimal single threshold.

The resulting optimal certification threshold is $z^* \approx 14.9$, slightly below the mean quality. This implies that roughly 66.5% of plans should be certified. Evaluating the performance of this two-tier rating relative to the full-information and no-information benchmarks, we find that it closes about 64% of the surplus gap.

The empirical quality distribution for MA plans is left-skewed: most plans have relatively high quality, with a longer lower tail. This is exactly the environment where our theory predicts that the optimal threshold should be below the mean, with a majority of plans certified and the rating primarily serving to screen out the lower tail.

4.5.2 | eBay Sellers' Ratings

Our second illustration uses data from Nosko and Tadelis [14] on eBay sellers' quality. They introduce a quality measure given by the percentage of transactions with positive feedback (as opposed to negative or no feedback). We treat this percentage as our quality index z.

Interpreting this variable as the ex-ante probability of a "good" experience, and in a simple two-state model where utility is affine in this probability, expected utility is an affine transformation of z. Thus, we can directly treat the distribution of this feedback measure as a distribution of quality levels in our framework. Based on this interpretation, we can use this distribution to calculate the optimal certification threshold, as we did for other distributions above.

14676451, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/joie.70025 by Maryam Saeedi - Carnegie Mellon University , Wiley Online Library on [12/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

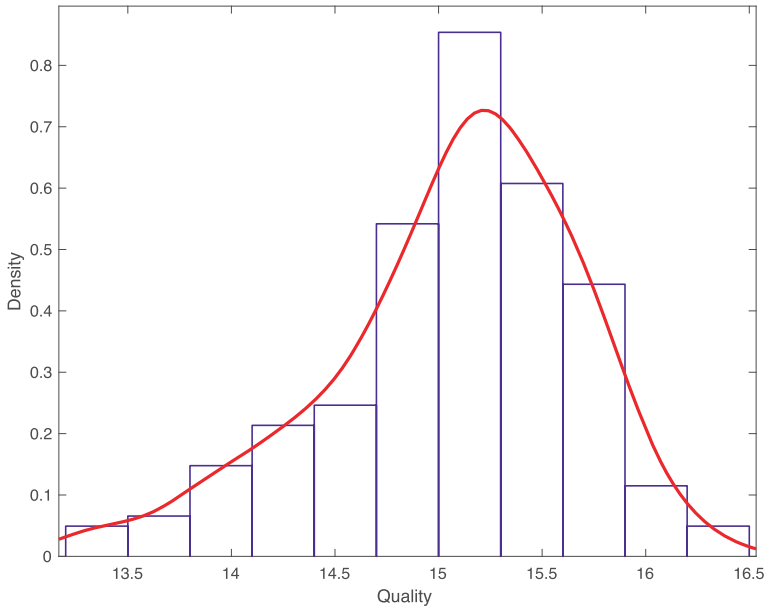

FIGURE 3 | Medicare advantage quality distribution. *Notes:* [Colour figure can be viewed at wileyonlinelibrary.com]

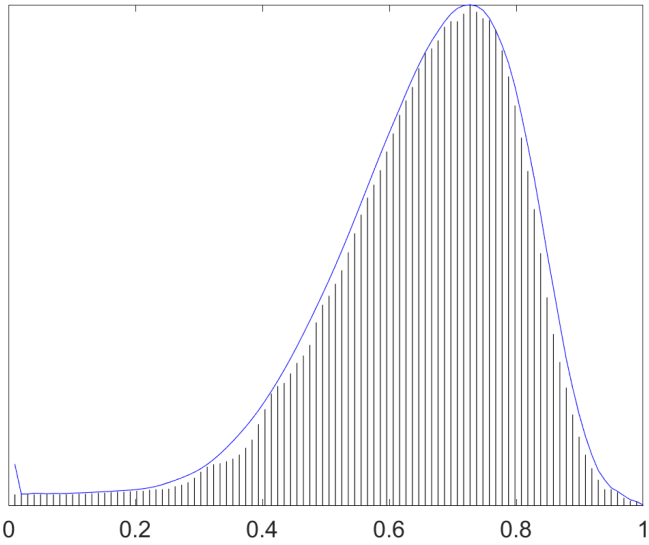

FIGURE 4 | Distribution of percent positive responses for eBay sellers. *Notes:* [Colour figure can be viewed at wileyonlinelibrary.com]

We proceed analogously to the MA case. Using the distribution of percent-positive feedback across sellers (Figure 4), we estimate a smooth approximation of F, compute the optimal two-tier certification threshold under the linear benchmark using our k-means method, and evaluate the resulting performance ratio.[14]

The results are again similar to the MA case. The distribution of seller quality is left-skewed: most sellers have very high feedback scores, with a lower tail of under-performers. The optimal two-tier rating certifies a large majority—more than 65%—of sellers, and the resulting certification closes about 63% of the surplus gap between full information and no information. Also in this case, the optimal rating primarily serves to screen out the relatively small group of low-quality sellers rather than to identify a narrow elite of top performers.

5 | Final Remarks

In this paper, we studied the optimal design of quality ratings in markets with limited signals. Ratings reallocate demand across producers, affecting both the average quality of goods consumed and average production costs. Optimal thresholds in a discrete rating system balance this trade-off and depend on market primitives, including the distribution of producer quality and the elasticity of supply. For instance, we show that when supply becomes more convex, the optimal thresholds pointwise shift to the right. Intuitively, in a simple two-tier certification system, a more elastic supply function calls for a higher threshold and a smaller share of certified sellers. We also show that skewness in the distribution of seller qualities matters: optimal thresholds shift in the direction of the skew.

For the benchmark case of linear supply, we characterize optimal thresholds as the solution to a standard k-means clustering problem, providing a simple and easy-to-implement approach to the design of rating systems. This characterization allows us to derive bounds on the performance of rating systems as a function of the number of categories. In particular, we show theoretically that a simple certification mechanism achieves at least half of the surplus gains from the best feasible rating mechanism for a broad class of quality distributions, including those with log-concave densities. We then complement these results with numerical analysis. For example, for the exponential family of distributions, 65% of the surplus gains from full information can be captured with only two categories. Finally, we apply our method to two empirical settings to illustrate how it can be used to compute optimal thresholds and the associated welfare gains in practice.

From a practical perspective, our findings rationalize the widespread use of coarse ratings. While simple mechanisms sacrifice some information, the welfare loss is modest and diminishes rapidly with additional categories. Given the additional

14676451, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/joie.70025 by Maryam Saeedi - Carnegie Mellon University , Wiley Online Library on [12/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

advantages of simplicity, such as transparency for consumers, lower cognitive costs, and reduced implementation burden, our analysis supports continued reliance on coarse ratings across many market settings.

Acknowledgments

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.

Endnotes

[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 September, 2023.

[2]Moral hazard is a critical consideration in some markets, while adverse selection might play a more critical role in others, as suggested by Hui et al. [1] in an empirical study using eBay data. Optimal rating design with moral hazard and adverse selection is considered in Saeedi and Shourideh [2] in a simplified market environment. Shi et al. [3] and Vatter [4] also study optimal information disclosure in a model that includes moral hazard.

[3]Dranove and Jin [11] provide an excellent survey of the earlier papers on certification.

[4]We used this canonical model in a previous paper, Hopenhayn and Saeedi [6], to answer another question: what is the optimal solution to the optimal rating without considering the practical problem of simple ratings.

[5]This representation of the information structure is consistent with the approach followed in Ganuza and Penalva [22] and Gentzkow and Kamenica [23]. 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.

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

[7]Bergemann and Pesendorfer [24] find a similar result in the context of optimal information design in auctions.

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

[9]When $\alpha \leq 2$, the value of z^* is undefined, as total surplus is strictly increasing in z^* in all the support.

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

[11]Given the assumption of linear supply, the total quantity stays the same.

[12]Consider a Laplace distribution of $(0, b)$, we have: $\bar{z} = 0$, $M_1 = E(z|z < 0) = -b$; $M_2 = E(z|z > 0) = b$. Additionally $\int (z - \bar{z})^2 dF(z) = 2b^2$. Therefore, using the formula for the surplus gain for a two-tier system, we obtain: $\gamma = 0.5$.

[13]The empty cells correspond to cases where firm supply is not integrable as a result of a very elastic supply and long tails of the respective distributions.

[14]The data for the histogram comes directly from tab. 4 in Nosko and Tadelis [14].

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

[16]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 (A.12).

References

1. X. Hui, M. Saeedi, and N. Sundaresan, "Adverse Selection or Moral Hazard, an Empirical Study," *Journal of Industrial Economics* 66 (2018): 610–649.

2. M. Saeedi and A. Shourideh, "Optimal Rating Design Under Moral Hazard," 2026.

3. Z. Shi, K. Srinivasan, and K. Zhang, "Design of Platform Reputation Systems: Optimal Information Disclosure," *Marketing Science* 42 (2023): 500–520.

4. B. Vatter, *Quality Disclosure and Regulation: Scoring Design in Medicare Advantage* (Tech. Rep. 3, 2025).

5. J. MacQueen, "Some Methods for Classification and Analysis of Multivariate Observations," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1 (1967), 281–297.

6. H. Hopenhayn and M. Saeedi, "Optimal Information Disclosure and Market Outcomes," *Theoretical Economics* 18 (2023): 1317–1344.

7. M. Saeedi, "Reputation and Adverse Selection: Theory and Evidence From eBay," *Rand Journal of Economics* 50 (2019): 822–853.

8. D. W. Elfenbein, R. Fisman, and B. McManus, "Market Structure, Reputation, and the Value of Quality Certification," *American Economic Journal: Microeconomics* 7 (2015): 83–108.

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

10. G. Z. Jin and P. Leslie, "The Effect of Information on Product Quality: Evidence From Restaurant Hygiene Grade Cards," *Quarterly Journal of Economics* 118 (2003): 409–451.

11. D. Dranove and G. Z. Jin, "Quality Disclosure and Certification: Theory and Practice," *Journal of Economic Literature* 48 (2010): 935–963.

12. X. Hui, M. Saeedi, Z. Shen, and N. Sundaresan, "Reputation and Regulations: Evidence From eBay," *Management Science* 62 (2016): 3604–3616.

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

14. C. Nosko and S. Tadelis, *The Limits of Reputation in Platform Markets: An Empirical Analysis and Field Experiment* (Tech. rep., National Bureau of Economic Research, 2015).

15. X. Hui, Z. Liu, and W. Zhang, "From High Bar to Uneven Bars: The Impact of Information Granularity in Quality Certification," *Management Science* 69 (2023): 6109–6127.

16. R. Wilson, "Efficient and Competitive Rationing," *Econometrica: Journal of the Econometric Society* 57 (1989): 1–40.

17. R. P. McAfee, "Coarse Matching," *Econometrica* 70 (2002): 2025–2034.

18. H. C. Hoppe, B. Moldovanu, and E. Ozdenoren, "Coarse Matching With Incomplete Information," *Economic Theory* 47 (2011): 75–104.

19. R. Shao, "Generalized Coarse Matching," *Games and Economic Behavior* 100 (2016): 142–148.

20. Q. Lyu, W. Suen, and Y. Zhang, "Coarse Information Design," arXiv preprint, arXiv:2305.18020 (2023).

21. Y. C. Aybas and E. Turkel, "Persuasion With Coarse Communication," 2024.

14676451, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/joie.70025 by Maryam Saeedi - Carnegie Mellon University , Wiley Online Library on [12/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

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

23. M. Gentzkow and E. Kamenica, "A Rothschild-Stiglitz Approach to Bayesian Persuasion," *American Economic Review* 106 (2016): 597–601.

24. D. Bergemann and M. Pesendorfer, "Information Structures in Optimal Auctions," *Journal of Economic Theory* 137 (2007): 580–609.

25. W. Van Zwet, *Convex Transformations of Random Variables, Math* (Centrum, 1964).

26. D. Stoyan and D. J. Daley, *Comparison Methods for Queues and Other Stochastic Models* (John Wiley & Sons, Inc., 1983).

27. D. Mease and V. N. Nair, "Unique Optimal Partitions of Distributions and Connections to Hazard Rates and Stochastic Ordering," *Statistica Sinica* 16, no. 4 (2006): 1299–1312.

28. P. Milgrom and J. Roberts, "Rationalizability, Learning, and Equilibrium in Games With Strategic Complementarities," *Econometrica: Journal of the Econometric Society* 58 (1990): 1255–1277.

Appendix A

Proofs

Proof of Proposition 1

Consider a partition of the set of sellers into sets $S_1, \ldots, S_N$. 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 [6], this results in higher total surplus.

Proof of Lemma 1

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 $q_1, \ldots, q_N$. In particular, this implies that $\partial Q / \partial z_k = f(z_k)(q_k - q_{k+1})$. Given that $M_k = \int_{z_{k-1}}^{z_k} z dF(z) / (F(z_k) - F(z_{k-1}))$, 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 Lemma 2

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

Lemma 3. *The optimal thresholds satisfy the following condition:*

$$\frac{z_k - M_k}{M_{k+1} - M_k} S(p_k) + \frac{M_{k+1} - z_k}{M_{k+1} - M_k} S(p_{k+1}) = \frac{\int_{p_k}^{p_{k+1}} S(p) dp}{p_{k+1} - p_k}, \quad \text{(A.1)}$$

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

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

Substituting in (5) and rearranging gives

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

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

We use the expression found in Lemma 3. Equation (A.1) equates the expected value of $S(p)$ under two lotteries. The left hand side lottery has weights $\alpha = (z_k - M_k)/(M_{k+1} - M_k)$ on price p_k and $(1 - \alpha)$ 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

$$z_k - M_k = M_{k+1} - z_k. \quad \text{(A.2)}$$

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 (A.2), although 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:

$$\begin{aligned}
TS(z) &= \int_0^Q P(x)dx + \sum_{k=1}^N [F(z_k) - F(z_{k-1})] \\
&\quad \left[M_k(P + M_k) - \frac{1}{2}\left((P + M_k)^2\right)\right] \quad \text{(A.3)} \\
&= \int_0^Q P(x)dx + \sum_{k=1}^N [F(z_k) - F(z_{k-1})]\left[\frac{1}{2}M_k^2 - \frac{1}{2}P^2\right]. \quad \text{(A.4)}
\end{aligned}$$

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

$$\sum_{k=1}^N [F(z_k) - F(z_{k-1})](M_k - \bar{z})^2, \quad \text{(A.5)}$$

where $\bar{z} = \sum_{k=1}^N [F(z_k) - F(z_{k-1})]M_k$ is the mean seller quality, which is independent of the partition. The above expression is the variance between partitions. Because total variance is fixed, maximizing (A.5) is equivalent to minimizing (7). Uniqueness of the thresholds is guaranteed when the distribution has a log-concave density, as shown in Mease and Nair [27].

Proof of Proposition 3

We use the following properties of distributions with log-concave densities (see Lemma 1 in Mease and Nair [27]):

$$\mathbb{E}(z | s \leq z \leq s + d) - s \text{ is decreasing in } s \text{ for } d > 0 \text{ and} \quad \text{(A.6)}$$

$$s - \mathbb{E}(z | s - d \leq z \leq s) \text{ is increasing in } s \text{ for } d > 0, \quad \text{(A.7)}$$

and these properties are preserved when conditioning on intervals.

Lemma 4. *Suppose F is a distribution with log-concave density, and let $m(a, b) = E_F(z | a \leq z \leq b)$. Suppose the vector of thresholds $\{l_k\}_{k=1}^{N-1}$ satisfies*

$$l_k - m(l_{k-1}, l_k) = m(l_k, l_{k+1}) - l_k, \quad \text{(A.8)}$$

14676451, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/joie.70025 by Maryam Saeedi - Carnegie Mellon University , Wiley Online Library on [12/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

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

$$z_k - m(z_{k-1}, z_k) > m(z_k, z_{k+1}) - z_k. \tag{A.9}$$

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

To prove Lemma 4, we first use the following:

Claim. Under the assumptions of Lemma 4, 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

$$\begin{aligned} z_k - m(z_{k-1}, z_k) &> m(z_k, z_{k+1}) - z_k \\ &\geq m(z_k, z_k + l_{k+1} - l_k) - z_k \\ &\geq m(l_k, l_{k+1}) - l_k \\ &= l_k - m(l_{k-1}, l_k). \end{aligned} \tag{A.10}$$

The first inequality follows from (A.9), the second one, from the monotonicity of m, the third, from (A.6), and the last, from (A.8). 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

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

where the first inequality follows from condition (A.7), and the second one, from the monotonicity of m. This inequality contradicts (A.10), 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 4. 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 $k = 1, \ldots, h$. For $k = 1$, the claim would imply that $z_0 < l_0$, which cannot be true if the distribution had a lower bound, because 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 2 and equations (A.8) and (A.10) hold, so Lemma 4 proves the proposition.

Proof of Proposition 4

In order to prove this proposition, we need to first state two lemmas. The first establishes a local ordering of conditional expectations based on the convexity of supply. The second establishes the monotonicity of the threshold response function.

Lemma 5. *Let* S_A *and* S_B *be strictly increasing supply functions where* S_B *is more convex than* S_A. *For any fixed interval of qualities* $[m_l, m_h]$ *and fixed base price* P, *the supply-weighted average quality is strictly higher under* S_B *than under* S_A:

$$\frac{\int_{m_l}^{m_h} m \, dS_B(P+m)}{S_B(P+m_h) - S_B(P+m_l)} \geq \frac{\int_{m_l}^{m_h} m \, dS_A(P+m)}{S_A(P+m_h) - S_A(P+m_l)}$$

Proof. Let $w_A(m)$ and $w_B(m)$ denote the probability density functions for quality m induced by the supply responsiveness over the interval. Specifically, for $i \in \{A, B\}$, the density is given by:

$$w_i(m) = \frac{S_i'(P+m)}{S_i(P+m_h) - S_i(P+m_l)}$$

We analyze the likelihood ratio $\frac{w_B(m)}{w_A(m)}$. Using the definition $S_B(p) = \phi(S_A(p))$, we have:

$$S_B'(P+m) = \phi'(S_A(P+m)) \cdot S_A'(P+m).$$

Let K be the ratio of the normalization constants, which is independent of m:

$$K = \frac{S_A(P+m_h) - S_A(P+m_l)}{S_B(P+m_h) - S_B(P+m_l)}.$$

The likelihood ratio is:

$$\begin{aligned} \frac{w_B(m)}{w_A(m)} = \frac{S_B'(P+m)}{S_A'(P+m)} \cdot K &= \frac{\phi'(S_A(P+m))S_A'(P+m)}{S_A'(P+m)} \\ &\cdot K = K \cdot \phi'(S_A(P+m)). \end{aligned}$$

Since S_A is increasing and ϕ is convex (so ϕ' is increasing), the term $\phi'(S_A(P+m))$ is nondecreasing in m. This implies that the distribution of qualities under S_B dominates that under S_A in the likelihood ratio order (and thus First-Order Stochastic Dominance). Therefore, the expected quality under S_B is strictly higher. □

Lemma 6. *Let* $T(z; S)$ *be the mapping defined by the system of first-order conditions for supply* S, *such that the* k*-th component* $T_k(z; S)$ *is the optimal threshold* z_k *given the adjacent thresholds* z_{k-1} *and* z_{k+1}. *The mapping* $T(z; S)$ *is monotone increasing in* z *for any strictly increasing supply function* S.

Proof. Using Equation (5), and letting $p = P + m$, we have:

$$\begin{aligned} P + z_k &= \frac{c(q_{k+1}) - c(q_k)}{q_{k+1} - q_k} \\ &= \frac{\int_{P+M_k}^{P+M_{k+1}} (P+m) \, dS(P+m)}{S(P+M_{k+1}) - S(P+M_k)}, \end{aligned}$$

The term P on the RHS can be factored out of the expectation:

$$P + z_k = P + \frac{\int_{M_k}^{M_{k+1}} m \, dS(P+m)}{S(P+M_{k+1}) - S(P+M_k)}.$$

Thus:

$$z_k = E_S[m | M_k \leq m \leq M_{k+1}],$$

where the expectation is taken with respect to the supply-induced measure on quality m.

We now consider the monotonicity: If the vector z increases, the neighboring conditional means $M_k(z_{k-1}, z_k)$ and $M_{k+1}(z_k, z_{k+1})$ increase. This shifts the integration interval $[M_k, M_{k+1}]$ to the right. For any distribution, shifting the conditioning bounds upward strictly increases the conditional expectation. Thus, the best response z_k must increase. □

Now, using the above lemmas, we prove Proposition 4. Let z^A and z^B denote the vectors of optimal thresholds for supply functions S_A and S_B, respectively. These thresholds are the fixed points of the respective best-response mappings: $z^A = T(z^A; S_A)$ and $z^B = T(z^B; S_B)$.

This optimization problem can be characterized as a pseudo-supermodular game. We view each threshold z_k as a "player" whose strategy is to choose a position on the real line. The game exhibits strategic complementarities: as established in Lemma 6, if the neighboring thresholds (players $k-1$ and $k+1$) choose higher values, the optimal response for player k is to also choose a higher value.

The proof relies on the comparative statics result for supermodular games (specifically, Theorem 6 in Milgrom and Roberts [28]). We require two conditions:

14676451, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/joie.70025 by Maryam Saeedi - Carnegie Mellon University , Wiley Online Library on [12/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

1. **Ordered Mappings:** For any fixed vector z (and thus fixed means M), we compare the best responses. From the simplified first-order condition derived in Lemma 6:

$$T_k(z;S) = E_S[m|M_k \le m \le M_{k+1}].$$

By Lemma 5, since S_B is more convex than S_A, the conditional expectation (RHS) is strictly higher under S_B than under S_A for the same bounds. Thus, $T(z;S_B) \ge T(z;S_A)$ pointwise.

2. **Monotonicity:** The mappings $T(\cdot;S_A)$ and $T(\cdot;S_B)$ are monotone increasing. This is established in Lemma 6.

Let z^* be the fixed point. Since $T(\cdot;S_B)$ pointwise dominates $T(\cdot;S_A)$ and both are monotone, standard fixed-point comparison theorems (e.g., Milgrom and Roberts [28]) imply that the fixed points are ordered:

$$z^B \ge z^A.$$

Proof of Proposition 5

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

Lemma 7. *Let $g(z_1), \ldots, g(z_{N-1})$ be the optimal thresholds for $\tilde{F}$. Let $M_k = m(z_{k-1}, z_k) = E_F(z_{k-1} \le z \le z_k)$. Then $z_k - M_k > M_{k+1} - z_k$.*

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

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

□

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 (A.8) follows from the necessary condition for optimal thresholds, and (A.10) follows from the previous lemma.

Proof of Proposition 6

The first order conditions are:

$$P'(Q)q(z) + P(Q) + z = c. \tag{A.11}$$

Summing over the FOC of all sellers we get:

$$P'(Q)Q + nP(Q) + n\bar{z} = nc, \tag{A.12}$$

where $\bar{z}$ is the average quality of sellers in the market, which is assumed to be exogenous.[15] Thus Q is pinned down by Equation (A.12) independently of the cross-sectional distribution of z (apart from its mean $\bar{z}$).[16] Furthermore, from Equation (A.11) it follows that

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

therefore, supply is linear in z, given Q. Substituting for quantity in the profit function we get:

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

Adding profits of all the sellers, we can write total profit as following

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

Also note that consumer surplus depends on Q, which is independent of the information structure. Thus, maximizing total surplus is equivalent to maximizing total profits. The only term in the total profit function that varies with information structure is: $\int z^2 dG(z)$. Therefore, maximizing total profit and total surplus is equivalent to maximizing $\int z^2 dG(z)$, as in the case of perfect competition with linear supply (see Equation (6)).

Proof of Proposition 7

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

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

where the second equality follows from

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

and similarly for cv_2. From the above inequality,

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

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

14676451, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/joie.70025 by Maryam Saeedi - Carnegie Mellon University , Wiley Online Library on [12/03/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License