ECONOMIC LENS

Issue 25

6 September 2026

Beyond the Rating: Making Better Use of Online Reviews

Are Online Ratings Enough?

Ratings have become one of the most common ways of measuring consumer experience. From hotels and restaurants to airlines and tourist attractions, millions of consumers leave ratings every day, creating a continuously expanding source of information on the quality of goods and services.

Platforms have also moved beyond a single overall rating, offering subscores for dimensions such as cleanliness, service, value for money, or food quality. But a score can tell us that there is a problem without telling us what is causing it or what needs to change.

A restaurant, for example, may receive a low food-quality score, but its reviews may reveal that customers repeatedly complain about meat, specifically the fact that it is overcooked . That distinction matters: identifying the specific problem allows the business to determine whether it needs to change its cooking process, standards, or even its meat supplier.

The same limitation matters for policymakers. Ratings can signal that performance is deteriorating across a sector, destination, or service, but they do not necessarily show what is driving that deterioration or where policy intervention should be directed. Written reviews can provide this missing information, but analyzing them manually takes considerable time and may require multiple experts with different areas of expertise. The challenge becomes even greater as the volume of reviews grows and feedback is provided across different languages and dialects.

The result is an information gap: consumers are providing detailed accounts of their experiences, but much of that information is lost when feedback is reduced to numerical ratings and predefined categories.

This lens introduces a new methodology for turning large volumes of unstructured public feedback into evidence for business and policy decisions, reducing the time and resources required to effectively identify and diagnose problems.

From Ratings to the Reasons Behind Them

Recent advances in Large Language Models (LLMs) make it possible to systematically analyze the information contained in written reviews at scale. Using few-shot prompting, the model can be guided to identify the specific problems described by reviewers, recognize different ways of describing the same issue, and separate multiple issues raised in a single review.

This becomes particularly valuable when reviews are numerous or written across different languages and dialects, as manually analyzing this volume of feedback can require significant time and multiple experts with different areas of expertise. Rather than relying only on predefined rating categories, the analysis can organize reviews around the issues people actually report, allowing businesses and policymakers to identify recurring problems that may not be visible in the ratings alone.

The approach does not replace ratings or surveys. It complements them by adding the information behind the number:

Rating → **How much dissatisfaction exists**

Subscore → **Where the concern lies**

Review text → **What is specifically happening**

Large-scale analysis → **Which problems recur and how widely**

This gives businesses and policymakers a clearer basis for determining what needs to change, where intervention is needed, and what form that intervention should take.

From a Declining Rating to a Specific Intervention

To demonstrate what can be gained by looking beyond the rating, we analyzed passenger reviews from Tripadvisor for an anonymous European airline over time.

1. What the Rating Tells Us

The airline's average passenger rating fell from 3.27 in 2019 to around 1.6 by 2024. The decline clearly indicates a deterioration in passenger satisfaction. But the rating alone cannot tell a business or policymaker what is driving that deterioration, where the problems are occurring, or where action should begin.

2. What the Reviews Reveal

Table 1 shows how written reviews provide a first layer of explanation by identifying the issues most frequently reported by passengers through the number of mentions:

Table 1: Frequency of most mentioned issues

Problem described by passengers	Mentions
Flight disruptions	690
Rude flight attendants	591
Poor food quality	555
Poor communication during delays	536
Lack of amenities	442
Unresponsive crew	346
Lost baggage	335

However, this is only the first layer which identifies where the main problems lie, but it does not yet show precisely what is driving each problem. The methodology can drill down into any broad category and identify the specific issues reported by passengers. For example, the 690 mentions classified under "Flight Disruptions" can be further separated into distinct issues:

A- Flight Delays and cancellations

B- Poor Communication regarding delays

C- Unexplained cancellations

D- Missed connections

E-Excessive flight delays

F-Inadequate pre-Flight communication

G- Unclear announcements

The analysis can be further refined to identify where and when specific problems occur. For example, recurring complaints about flight delays could be traced to particular routes, airports, or periods, while complaints about baggage handling could be linked to specific locations or stages of the passenger journey. This level of detail allows decision-makers to distinguish between widespread problems and problems concentrated in particular locations or periods, and to direct resources and interventions accordingly.

This changes the question from "Why is the rating falling?" to "What exactly is going wrong, and where should action be taken?" A broad finding such as "flight disruptions" does not by itself indicate an appropriate response. Identifying the underlying issue allows decision-makers to distinguish between operational failures, communication failures, and other specific problems—and to direct intervention accordingly.

From Diagnosis to Intervention

The same methodology can be applied across sectors where online reviews are available, including tourism, hospitality, restaurants, retail, transportation, and other consumer-facing services. In each case, it can move analysis beyond overall ratings to identify the specific issues affecting consumer experience and inform targeted interventions.

For businesses, this provides a basis for moving beyond a declining score and identifying the specific changes required. For policymakers, the implications are broader because public interventions ultimately affect people's welfare and can influence the performance and reputation of entire sectors.

In areas such as airports and tourism, for example, service quality can shape a visitor's experience from the moment they enter a country. Airports are often a tourist's first point of contact with a country, making the quality of that experience part of the country's overall tourism experience and international image. At the macro level, persistent problems in these services can therefore affect tourism demand, visitor spending, sector performance, and ultimately economic activity.

The value of this methodology is not simply to measure dissatisfaction. It is to identify what is driving it, allowing businesses and policymakers to move from detecting a problem to targeting the appropriate intervention—and to track whether conditions improve over time.