

From Reviews to Real Output

Using sentiment analysis to measure service quality adjustment

John Lourenze Poquiz
jlp87@cam.ac.uk

Standard price deflation ignores quality change in services. Using 520,000+ Trustpilot reviews for UK Accommodation Services, we construct a sentiment-based quality index and adjust nominal output, revealing that official GDP significantly overstates pre-COVID volume growth

Why Quality Adjustment Matters

- GDP deflation separates price from volume, but volume must be adjusted for quality change (2008 SNA par 15.13)
- Services quality improvements are rarely captured in prices: a hotel opening a pool or a restaurant hiring a star chef improve output without a commensurate price rise
- Hedonic methods are well-developed for goods, where characteristics are measurable.

Sentiment Analysis as a Signal for Quality

- 520,464 Trustpilot reviews scraped for 1,073 UK Accommodation Services firms (2013–2024)
- We use **SentimentR**
 - Sentence-level analysis with negation shifters, normalised by word count
 - Balancing accuracy with reproducibility
- We take the average sentiment score for each establishment.

Quality adjustment

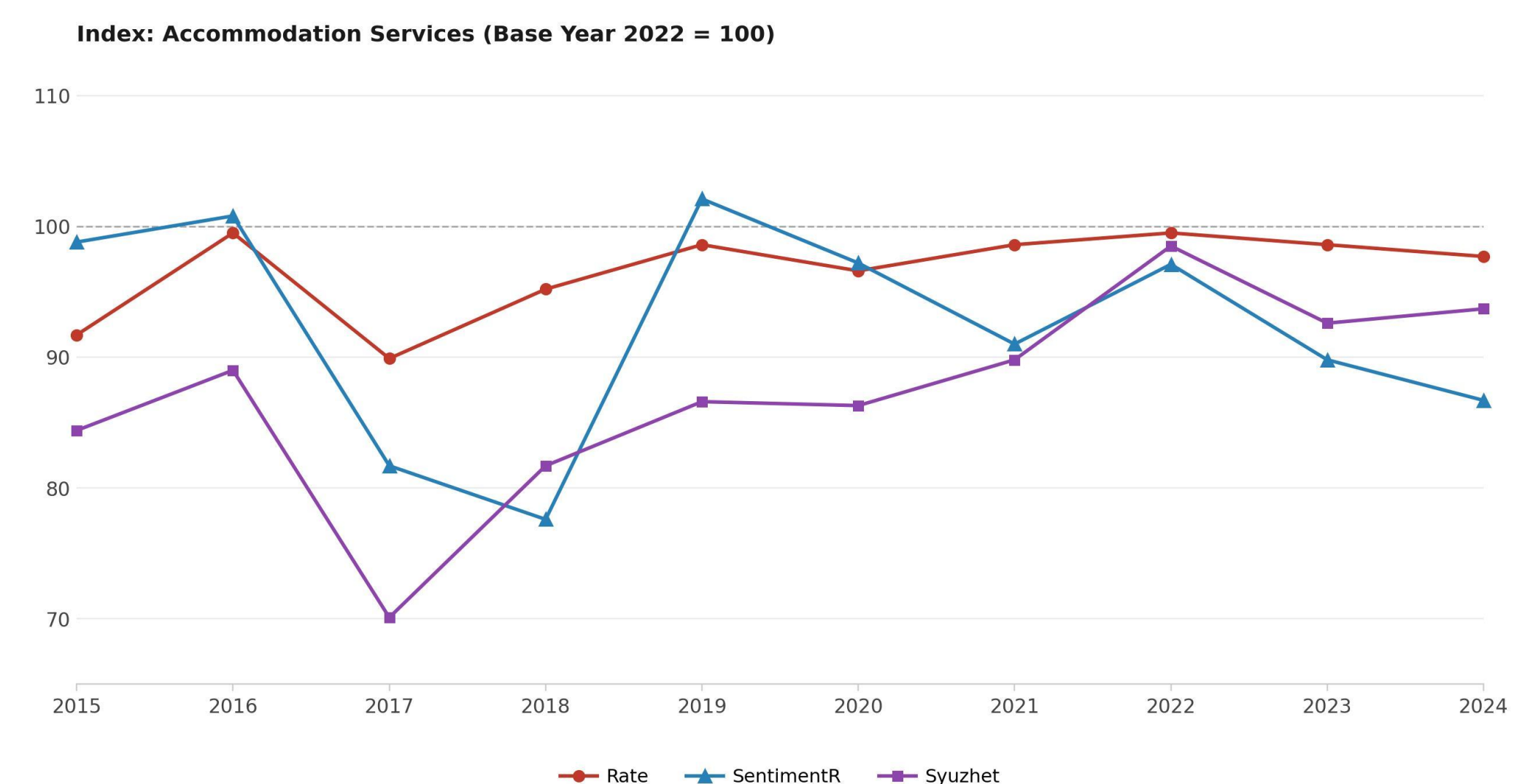
$$P_t^* = \frac{P_t}{\bar{S}_t}$$

- \bar{S}_t is a firm-level sentiment index (base year 2022 = 100), constructed from annual Trustpilot review scores

- P_t^* is the quality-adjusted deflator: observed price divided by the sentiment index

$$\tilde{Y}_t^r = \frac{Y_t^r}{P_t^*}$$

- \tilde{Y}_t^r is quality-adjusted real output, reflecting both quantity and perceived service quality



Comparing Published and Adjusted Growth Rate of Accommodation Services Industry Value Added

	Published		Adjusted	
	Total %	CAGR %	Total %	CAGR %
Pre-COVID (2015–2019)	16.11	3.81	4.90	1.20
COVID Impact (2018–2020)	-48.98	-28.57	-49.11	-28.66
Recovery (2020–2022)	91.67	38.45	92.29	38.67

Does it change the growth story?

- Pre-COVID CAGR drops from 3.81% to 1.20% after quality adjustment, official figures overstate growth by over 2.5 pp annually
- COVID shock and recovery patterns are preserved
- Adjusted and unadjusted series converge post-2020, as perceived quality stabilised after reopening

