Research Question 1: Information Asymmetry & Decision-Making Quality

How does pricing model transparency affect users' ability to make optimal consumption decisions, and what are the welfare implications of information gaps between providers and consumers?

Executive Summary

This research question investigates whether per-token pricing models enable or obstruct rational consumer decision-making in online services markets, particularly for AI LLMs. The central tension lies between the theoretical precision of usage-based pricing and the practical challenges of **bounded rationality**, **information asymmetry**, and **unpredictable consumption patterns**. Evidence suggests that while token-based pricing appears transparent, it creates significant cognitive burdens and decision-making failures that undermine consumer welfare.

Theoretical Framework

Information Economics & Asymmetric Information

Akerlof's foundational work on asymmetric information demonstrates that when sellers possess superior knowledge about product quality or costs, markets can fail to achieve efficient outcomes [1] [2]. In per-token pricing for AI services, **structural information asymmetries** emerge across multiple dimensions:

1. Hidden Complexity of Consumption

Research on LLM token pricing reveals a fundamental paradox: while per-token prices have declined dramatically (from €36 to €0.07 per million tokens), the **tokens-per-task** metric has increased by **100x** for reasoning models [3] [4]. This creates what practitioners call the "**LLM Cost Paradox**"—advertised price reductions mask exponentially rising total costs [3].

The true cost driver is not the per-token price but rather the **operational nuances** of how models tokenize and process requests [4] [5]. For instance, the same query in different languages can consume vastly different token quantities—up to **25 times more** for certain non-English languages due to tokenization inefficiencies [6] [7]. Users cannot predict these variations without deep technical knowledge.

2. Credence Goods Problem

Online AI services exhibit characteristics of **credence goods**—products where consumers cannot evaluate quality even after consumption [8] [9] [10] [11] [12]. In credence goods markets, three types of fraud are possible:

- **Undertreatment**: Providing insufficient service (using cheaper, less capable models)

- **Overtreatment**: Providing excessive service (unnecessarily verbose outputs)
- **Overcharging**: Charging for more than delivered (opaque token counting) [8] [9]

Expert sellers in credence markets possess **private information** about the optimal service level, creating opportunities for exploitation [10] [11]. Experimental evidence shows that price competition in credence goods markets can **increase fraud** rather than reduce it, as experts compete on margins rather than quality [9] [13].

In AI token pricing, providers control not just the price but the **unit of measurement itself** (tokenization algorithms), creating a double-asymmetry where users can neither verify consumption accuracy nor assess service appropriateness [14].

Bounded Rationality & Decision-Making Under Uncertainty

Herbert Simon's **bounded rationality** theory posits that humans employ simplified decision-making processes rather than optimizing across all available information [15] [16] [17]. In the context of token-based pricing, three constraints limit rational decision-making:

1. Computational Limitations

Users cannot accurately estimate token consumption for complex tasks. Research on on-demand service platforms demonstrates that both customers and providers struggle to gauge "congestion levels" (analogous to usage intensity), leading to **systematic estimation errors** [17]. The multinomial logit (MNL) framework models this as users being "incapable of accurately estimating" their true consumption, resulting in suboptimal purchasing decisions [17].

2. Information Processing Costs

Even when usage data is available, the **cognitive cost** of continuously monitoring and optimizing token consumption exceeds the benefit for most users. Platforms can exploit this by strategically manipulating information disclosure—providing "partial disclosure" when it increases profitability without significant loss of market coverage [1].

3. Temporal Discounting & Bill Shock

The separation between consumption (API calls) and payment (end-of-month billing) creates **temporal distance** that obscures cost accumulation. Industry reports indicate widespread "**bill shock**"—customers receiving unexpectedly high invoices when usage spikes [18] [19] [20] [21].

This aligns with behavioral economics research on **mental accounting**, where consumers categorize expenses into mental "buckets" that are treated as non-fungible [22] [23] [24]. A customer who budgets $100/month for "AI tools" experiences acute dissatisfaction when actual costs reach $400, even if the service delivered proportional value.

Empirical Evidence: Information Asymmetry in Token Pricing

Platform Information Disclosure Strategies

Research on online expert service (OES) platforms reveals how providers strategically manage information disclosure. A game-theoretic model incorporating consumer heterogeneity (sophisticated vs. naive) demonstrates that:

- When **marginal costs are intermediate**, platforms engage in **partial disclosure** to manipulate valuation beliefs and maximize profit [1]
- Counterintuitively, when naive consumers exist, platforms may **increase disclosure** compared to markets with only sophisticated consumers, because naive users misinterpret information in ways favorable to the platform [1]
- Platforms act as "designers of the disclosure framework," functioning as an additional player in the seller-consumer game beyond traditional two-sided market models [1]

For token-based pricing, this suggests providers may selectively disclose:

- **Simple metrics** (per-token price) while obscuring **complex metrics** (tokens-per-task, model efficiency variations)
- **Average costs** while downplaying **variability** and **tail risk** (95th percentile usage scenarios)
- **Gross consumption** while limiting **granular attribution** (which features/prompts consume the most tokens)

The "Tokens Per Task" Deception

A critical finding in LLM pricing analysis is that **advertised per-token prices are misleading indicators of actual costs** [4] [5]. The cost structure has two components:

$$textTotalCost = textPriceperTokentimestextTokensperTask$$

While providers compete aggressively on the first term (reducing per-token prices), the second term—**tokens per task**—is:

1. **Not standardized** across providers
2. **Highly variable** across use cases (10x-100x variation)
3. **Controlled by providers** through model architecture and output verbosity
4. **Difficult for users to predict** ex ante [3] [4] [5]

This creates a **principal-agent problem** where providers can maintain or increase revenue while appearing to cut prices, by implementing models that consume more tokens per query. As one industry analysis notes: "The real number to watch is tokens per task—and that number is climbing faster than cost curves can keep up" [3].

Language Inequality & Structural Discrimination

A systematic analysis of OpenAI's pricing policy across 22 typologically diverse languages reveals profound **cross-linguistic unfairness**. The study demonstrates:

- Token requirements to convey identical information vary **up to 25-fold** across languages[6] [7]
- Languages with **non-Latin scripts** (e.g., Tamil, Thai, Arabic) consume significantly more tokens than English
- This creates **implicit price discrimination**, where non-English speakers pay substantially more for equivalent service
- The effect compounds for **low-resource languages**, exacerbating global digital inequality[6] [7]

From an information asymmetry perspective, this is particularly pernicious because:

1. Users cannot determine token efficiency for their language without technical knowledge of tokenization algorithms
2. Pricing appears "neutral" (same per-token rate) but produces **systematically unequal outcomes**
3. The inequality is **structural** (embedded in technical design) rather than transparent

This exemplifies how **technical complexity** itself becomes a mechanism for information asymmetry, allowing facially neutral pricing to obscure discriminatory effects.

Bounded Rationality in Usage-Based Pricing

Satisficing vs. Optimizing

Research on platform pricing with bounded rationality demonstrates that when customers and providers cannot accurately assess "congestion" (usage intensity), several patterns emerge:

1. Systematic Errors in Both Directions

Unlike traditional economic models assuming rational optimization, bounded rationality models using **multinomial logit (MNL)** frameworks show that users make errors in **both directions**—sometimes overestimating costs (leading to under-consumption) and sometimes underestimating (leading to over-consumption and bill shock)[17].

2. Platform Exploitation of Bounded Rationality

Critically, platforms can **benefit from user irrationality**. Simulation results indicate that when either customers' or providers' bounded rationality levels and service valuation are "either high or low," the platform's profits can increase[17]. This creates perverse incentives for platforms to:

- **Avoid clarity** in cost projection tools
- **Complicate** pricing structures to exceed users' cognitive processing capacity
- **Exploit** the gap between perceived and actual costs

3. High Bounded Rationality Can Benefit Users

Counterintuitively, the research finds that "high levels of bounded rationality in either customers or providers can lead to increased consumer surplus and/or labor welfare" under certain conditions[17]. This paradoxical result suggests that **imperfect information processing** can sometimes protect consumers from exploitation—if neither party can accurately optimize, predatory pricing strategies become ineffective.

Mental Accounting & Consumption Decisions

Thaler's **mental accounting** theory explains how consumers categorize money into distinct "mental accounts" that are treated as non-fungible despite economic equivalence[22] [24]. Research on windfall gains versus hard-earned money demonstrates:

- Consumers allocate "**windfall money**" (unexpected income) to hedonic consumption
- Consumers allocate "**hard-earned money**" to utilitarian consumption
- The **scarcity mindset** moderates these effects—under high scarcity, even windfall money is conserved[23] [25]

Applied to token-based pricing:

1. Budget Category Effects

A user who receives a corporate AI budget may treat it as "windfall" (discretionary funds) and consume tokens liberally. The same user paying personally may exhibit extreme caution, treating every API call as "hard-earned money" being depleted.

2. Subscription vs. Usage Mental Models

Subscription pricing creates a **sunk cost mental account**—the fee is paid and "gone," so usage feels "free" at the margin. Usage-based pricing keeps consumption in the **loss aversion** frame, where each additional token represents a tangible loss[26] [22].

This explains why enterprise customers often prefer **committed-use discounts** or **tiered plans with included tokens**—these structures convert usage-based costs into psychologically simpler subscription mental accounts[27] [28].

The Predictability Paradox

Industry analysis reveals a fundamental tension: **usage-based pricing feels fair but creates anxiety**[18] [19]. Customers appreciate paying only for what they use, yet:

- **Enterprise buyers** require **annual budget certainty** for procurement processes[19]
- **CFOs** struggle to forecast revenue when consumption is highly variable[19]
- **Development teams** fear experimentation will trigger unexpected costs, **inhibiting innovation**[29] [28]

This creates what behavioral economists call a "**satisfaction-dissatisfaction asymmetry**"—the fairness of usage-based pricing generates positive sentiment during signup, but **bill variability** generates disproportionate negative sentiment when invoices arrive[18] [20] [21].

Welfare Implications & Market Failures

Consumer Welfare Effects

1. Under-Consumption Due to Uncertainty

When users cannot predict costs, **risk aversion** leads to **under-consumption** of valuable services. Economic theory predicts that uncertainty about prices creates an **option value** to waiting, depressing current demand[30] [31].

For AI services, this manifests as:

- Developers **hardcoding token limits** far below optimal levels
- Enterprises **restricting AI tool access** to avoid bill shock
- Individual users **forgoing valuable queries** to avoid crossing cost thresholds

Research on online shopping platforms demonstrates that **higher consumer information levels** generally increase platform profits and consumer surplus[32]. Conversely, information asymmetry in token pricing creates deadweight loss from foregone transactions.

2. Over-Consumption Due to Misestimation

Alternatively, **bounded rationality** can lead to **over-consumption** when users underestimate token usage. Post-consumption regret generates:

- **Churn** (cancellation after bill shock)[19]
- **Trust erosion** (perception of deceptive pricing)[33] [34]
- **Switching to competitors** (lock-in mitigation strategies)[35] [36]

3. Distributional Inequity

Information asymmetry effects are **not uniformly distributed**:

- **Sophisticated users** (with technical knowledge) can optimize consumption and negotiate better contracts
- **Naive users** pay higher effective prices through inefficient usage patterns[1]
- **Non-English speakers** face structural disadvantages from tokenization inefficiencies[6] [7]

This creates a **regressive pricing structure** where those with the least information pay the most per unit of value delivered—violating principles of **distributive justice**[37] [38] [39].

Provider Welfare & Market Dynamics

1. Revenue Optimization Through Opacity

Information asymmetry enables providers to **price discriminate** without explicit differential pricing. By controlling:

- **Tokenization algorithms** (which texts consume more tokens)

- **Model verbosity** (how many tokens are generated per query)
- **Default settings** (context window sizes, temperature parameters)

Providers can effectively charge different rates for equivalent services while maintaining a uniform nominal price [4] [5].

2. Adverse Selection in Model Choice

When users cannot accurately assess cost-performance trade-offs, they may:

- **Select suboptimal models** (overpaying for capability they don't need, or under-provisioning for critical tasks)
- **Default to premium models** as a risk-hedging strategy, even when cheaper alternatives would suffice
- **Avoid experimentation** with new models due to uncertainty about cost implications

This creates **market inefficiency**—resources are misallocated due to information failures rather than revealed preferences.

Systemic Market Failures

1. Race to Opacity

If information asymmetry enables profit extraction, competitive dynamics may create a "**race to opacity**" where:

- **Transparent pricing** becomes a competitive disadvantage (reveals actual costs, enabling informed shopping)
- Providers **complexify** pricing structures to obscure comparability
- **Standardization efforts** fail because no player has incentive to reduce information asymmetry

This mirrors findings in credence goods markets, where **price competition can increase fraud** because sellers compete on margins rather than transparency [9] [13].

2. Erosion of Trust & Market Collapse

Classic information economics predicts that severe information asymmetry can lead to **market collapse** (Akerlof's "lemons problem") [2]. In AI token pricing markets, this manifests as:

- **Reputation damage** to the pricing model itself (not just individual providers)
- **Regulatory intervention** to mandate transparency (as seen in subscription pricing) [26]
- **Shift to alternative models** (flat-rate, tiered, or outcome-based pricing) as trust erodes [29]

Research on online platforms shows that **information asymmetry reduction** through mechanisms like reviews, ratings, and transparency requirements can significantly improve market outcomes [40] [32].

Policy & Design Implications

Transparency Requirements

To mitigate information asymmetry, several interventions are supported by empirical evidence:

1. Standardized Cost Calculators

Platforms should provide **pre-transaction cost estimation tools** that account for:

- Historical usage patterns for similar tasks
- Language-specific token multipliers
- Model-specific efficiency ratings
- Confidence intervals (not just point estimates) [4] [5]

2. Real-Time Usage Monitoring

Bill shock is primarily a function of temporal delay between consumption and cost awareness. Real-time dashboards with:

- **Cumulative spend tracking**
- **Projected end-of-period costs**
- **Configurable alerts** at budget thresholds
  ...can convert usage-based pricing from a delayed-information problem to a continuous-feedback system [41] [18] [20].

3. Standardized Units of Measurement

The current situation where each provider controls tokenization algorithms creates incomparability. Industry-wide standards for:

- **Normalized task-based metrics** (e.g., "cost per 1000-word document summarization")
- **Benchmark suites** for cross-provider comparison
- **Language fairness indices** (token efficiency by language)
  ...would reduce information asymmetry while preserving competitive pricing [6] [7].

Cognitive Simplification Strategies

Given bounded rationality constraints, pricing structures should minimize cognitive load:

1. Hybrid Models

Evidence suggests **committed-use discounts** and **tiered plans with included tokens** reduce anxiety while maintaining usage-based fairness [27] [28]. These convert open-ended liability into **bounded risk**.

2. Defaults & Safeguards

Soft caps (alerts), **hard caps** (automatic cessation), and **opt-in expansion** can protect boundedly rational users from costly errors while preserving flexibility[18] [20].

3. Simplified Mental Models

Rather than exposing users to raw token counts, providers could offer **task-based abstractions**:

- "50 monthly reports" instead of "500,000 tokens"
- "$0.10 per query" instead of "$0.002 per 1000 tokens"

This aligns pricing presentation with users' mental models of value[5] [42].

Regulatory Considerations

1. Mandatory Pre-Transaction Disclosure

Analogous to **Truth in Lending** requirements for financial products, regulators could mandate:

- **Good faith estimates** of typical costs for described use cases
- **Worst-case scenario disclosures** (95th percentile usage)
- **Historical volatility** information (how much bills have varied for similar customers)

2. Fairness Audits

Given evidence of structural discrimination (language-based pricing inequality), regulators might require:

- **Disparate impact analysis** of tokenization algorithms
- **Fairness certification** for pricing models
- **Remediation requirements** when systematic bias is detected[6] [7]

3. Right to Explanation

Drawing on **AI ethics** frameworks, users could have a right to:

- **Itemized billing** showing token consumption by request
- **Explanation** of why certain requests consumed unexpected token quantities
- **Contestation mechanisms** for billing disputes[14]

Research Gaps & Future Directions

Methodological Needs

1. Field Experiments

Most current evidence comes from laboratory experiments (credence goods) or observational data (platform pricing). **Randomized controlled trials** varying disclosure levels, cost calculator availability, and billing frequency would provide causal evidence on information asymmetry mitigation strategies.

2. Cross-Platform Comparisons

Systematic comparison of token consumption for **identical tasks** across OpenAI, Anthropic, Google, and other providers would quantify the magnitude of opacity and incomparability in current markets [43] [44] [45].

3. Longitudinal Studies

How do users' **cost prediction accuracy** and **satisfaction** evolve over time with usage-based pricing? Do users learn to estimate costs accurately, or does bounded rationality persist?

Theoretical Extensions

1. Dynamic Information Asymmetry

Current models largely assume static information structures. **Dynamic models** could explore:

- How **learning** affects information asymmetry over time
- Whether **network effects** in information (users sharing cost experiences) reduce asymmetry
- How **strategic disclosure** by providers evolves in response to user learning

2. Multi-Dimensional Asymmetry

Most models focus on **single dimensions** of asymmetry (quality OR cost OR appropriateness). Token pricing involves **compound asymmetries**:

- Users don't know optimal service level (credence goods problem)
- Users don't know how to measure consumption (tokenization opacity)
- Users don't know comparative pricing (cross-provider incomparability)

Models integrating these multiple asymmetries would better capture market realities.

3. Behavioral Welfare Economics

Traditional welfare analysis assumes rational agents. **Behavioral welfare economics** frameworks could assess:

- Whether **bounded rationality** implies consumers are **better off** with simpler (but less "fair") subscription pricing
- How to weigh **ex ante preferences** (desire for usage-based fairness) against **ex post outcomes** (bill shock and regret)
- Whether **paternalistic** interventions (mandatory caps, defaults) improve welfare despite restricting choice

Conclusion

Per-token pricing for online services, particularly AI LLMs, presents a **paradox of transparency**. While theoretically precise and fair (charging for actual consumption), the model creates **profound information asymmetries** that undermine rational decision-making:

1. **Hidden complexity**: Tokens-per-task variability dwarfs per-token price differences, but users cannot predict consumption
2. **Bounded rationality**: Cognitive limitations prevent optimal usage decisions even when information is available
3. **Strategic opacity**: Providers benefit from information asymmetry, creating adverse incentives for transparency
4. **Structural inequality**: Technical design choices (tokenization) create discriminatory outcomes masked by neutral pricing

The welfare implications are substantial:

- **Under-consumption** from uncertainty (foregone value)
- **Over-consumption** from misestimation (bill shock, regret)
- **Distributional inequity** (sophisticated users benefit, naive users pay premium)
- **Market failures** (adverse selection, erosion of trust)

Addressing these challenges requires **multi-layered interventions**:

- **Technical**: Standardized metrics, real-time monitoring, cost calculators
- **Behavioral**: Simplified mental models, defaults, hybrid pricing structures
- **Regulatory**: Mandatory disclosure, fairness audits, contestation rights

The central policy question is not whether token-based pricing is inherently superior to alternatives, but rather: **Under what conditions can information asymmetries be sufficiently mitigated to enable welfare-enhancing usage-based pricing?** Current evidence suggests those conditions do not yet exist in AI services markets, implying that the theoretical benefits of usage-based pricing may not materialize in practice without substantial structural reforms.

References

Citations are embedded throughout the document using bracketed numbers corresponding to the source IDs from the research phase.

[46] [47] [48] [49] [50] [51] [52] [53] [54] [55] [56] [57] [58] [59] [60] [61] [62] [63] [64] [65] [66] [67] [68] [69] [70] [71] [72] [73] [74] [75] [76] [77] [78] [79] [80] [81] [82] [83] [84] [85] [86] [87] [88] [89] [90] [91] [92] [93] [94] [95] [96] [97] [98] [99] [100] [101] [102] [103] [104] [105] [106] [107] [108] [109] [110] [111] [112] [113] [114] [115] [116] [117] [118] [119] [120] [121] [122] [123] [124] [125] [126] [127] [128] [129] [130] [131] [132] [133] [134] [135] [136] [137] [138] [139] [140] [141] [142] [143] [144] [145] [146] [147] [148] [149] [150] [151] [152] [153] [154] [155] [156] [157] [158] [159] [160] [161] [162] [163] [164] [165] [166] [167] [168] [169] [170] [171] [172] [173] [174] [175] [176] [177] [178] [179] [180] [181] [182] [183] [184] [185] [186] [187] [188] [189] [190] [191] [192] [193] [194] [195] [196] [197] [198] [199] [200] [201] [202] [203] [204] [205] [206] [207] [208] [209] [210] [211]

⁂

1. https://www.tandfonline.com/doi/full/10.1080/01605682.2023.2269212
2. https://www.investopedia.com/terms/a/asymmetricinformation.asp
3. https://www.ikangai.com/the-llm-cost-paradox-how-cheaper-ai-models-are-breaking-budgets/
4. https://www.finops.org/wg/genai-finops-how-token-pricing-really-works/
5. https://kinde.com/learn/billing/billing-for-ai/ai-token-pricing-optimization-dynamic-cost-management-for-llm-powered-saas/
6. https://arxiv.org/pdf/2305.13707.pdf
7. https://aclanthology.org/2023.emnlp-main.614.pdf
8. https://www.econtribute.de/RePEc/ajk/ajkdps/ECONtribute_258_2023.pdf
9. https://www.econstor.eu/bitstream/10419/171619/1/wp-13-176_rev.pdf
10. https://arxiv.org/abs/2509.06069
11. https://www.econstor.eu/bitstream/10419/238225/1/2020-01.pdf
12. https://pure.mpg.de/pubman/item/item_3031522_6/component/file_3501197/2019_03online.pdf
13. https://www.hec.ca/finance/Fichier/Mimra2013.pdf
14. https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID3358724_code2714802.pdf?abstractid=3108333&mirid=1&type=2
15. https://psychotricks.com/bounded-rationality/
16. https://thedecisionlab.com/biases/bounded-rationality
17. https://www.sciencedirect.com/science/article/abs/pii/S0377221724006362
18. https://stripe.com/resources/more/pricing-flexibility-in-ai-services
19. https://flexprice.io/blog/why-ai-companies-have-adopted-usage-based-pricing
20. https://schematichq.com/blog/why-usage-based-billing-is-taking-over-saas
21. https://www.younium.com/blog/usage-based-pricing
22. https://www.behavioraleconomics.com/mental-money-the-psychology-of-subscription-payment-options/
23. https://pmc.ncbi.nlm.nih.gov/articles/PMC10361766/
24. https://bear.warrington.ufl.edu/brenner/mar7588/Papers/thaler-mktsci1985.pdf
25. https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2023.1162916/full
26. https://instituteofinterneteconomics.org/behavioral-economics-of-subscription-pricing/
27. https://stripe.com/resources/more/token-pricing-how-it-works-and-how-to-make-the-most-of-it
28. https://techforward.io/the-token-economy-why-usage-based-ai-pricing-is-both-a-blessing-and-a-trap/
29. https://www.getmonetizely.com/articles/genai-pricing-models-from-tokens-to-outcomes
30. https://arxiv.org/pdf/2311.00846.pdf
31. https://www.sciencedirect.com/science/article/abs/pii/S1059056022003112
32. https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2022.821979/full
33. https://verpex.com/blog/marketing-tips/price-discrimination-online-the-fairness-of-personalized-pricing-based-on-user-data
34. https://www.sciencedirect.com/science/article/abs/pii/S0278431924003268
35. https://onlinelibrary.wiley.com/doi/10.1111/jems.12612
36. https://conference.iza.org/DATA_2022/stenzhorn_e32647.pdf

37. https://www.nature.com/articles/s41598-022-19792-3
38. https://www.frontiersin.org/journals/sociology/articles/10.3389/fsoc.2022.883999/full
39. https://onlinelibrary.wiley.com/doi/abs/10.1111/poms.13369
40. http://www.emerald.com/jeim/article/36/6/1533-1555/205340
41. https://www.prompts.ai/en/blog/managing-token-level-costs-ai
42. https://www.getmonetizely.com/articles/understanding-token-based-pricing-for-agentic-ai-systems-a-new-paradigm-in-ai-economics
43. https://www.solvimon.com/pricing-guides/openai-versus-anthropic
44. https://www.aipricingcomparison.com/text-generation-api-pricing-calculator
45. https://anotherwrapper.com/tools/llm-pricing
46. https://onlinelibrary.wiley.com/doi/10.1002/mde.4472
47. https://www.frontiersin.org/articles/10.3389/fmars.2025.1601322/full
48. https://arxiv.org/abs/2410.13090
49. https://www.mdpi.com/0718-1876/19/2/61
50. http://ledger.pitt.edu/ojs/ledger/article/download/226/214
51. https://arxiv.org/pdf/2101.06210.pdf
52. https://www.frontiersin.org/articles/10.3389/fphy.2021.631659/pdf
53. https://arxiv.org/pdf/2307.16874.pdf
54. https://arxiv.org/pdf/2208.10271.pdf
55. https://dl.acm.org/doi/pdf/10.1145/3649318
56. https://www.rairo-ro.org/articles/ro/pdf/2022/01/ro210226.pdf
57. https://backend.orbit.dtu.dk/ws/files/290548890/2022_TMG_PartC_1_.pdf
58. https://arxiv.org/pdf/2410.19107.pdf
59. https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID4437801_code2617082.pdf?abstractid=3613261&mirid=1
60. https://skywork.ai/skypage/en/Technical-Barriers-to-Entry:-Challenges-in-North-American-AI-Model-Localization-and-Implementation/1950070050110914560
61. https://dl.acm.org/doi/fullHtml/10.1145/3497701.3497733
62. https://www.getmonetizely.com/articles/how-to-design-effective-pricing-models-for-network-effects-and-platform-businesses
63. https://competition-bureau.canada.ca/en/how-we-foster-competition/education-and-outreach/consultation-artificial-intelligence-and-competition-what-we-heard
64. https://en.wikipedia.org/wiki/Two-sided_market
65. https://www.sciencedirect.com/science/article/abs/pii/S0304405X23000715
66. https://www.mercatus.org/research/working-papers/data-really-barrier-entry-rethinking-competition-regulation-generative-ai
67. http://arxiv.org/pdf/2502.16363.pdf
68. https://competitionpolicyinternational.com/assets/Uploads/Autumn2014Schmalensee.pdf
69. https://www.nfx.com/post/network-effects-manual
70. https://www.linkedin.com/pulse/global-ai-llm-market-critical-analysis-through-five-uzwyshyn-ph-d--is3dc
71. https://academic.oup.com/jeea/article-pdf/1/4/990/10312916/jeea0990.pdf

72. https://www.tandfonline.com/doi/full/10.1080/13241583.2024.2393933
73. https://ieeexplore.ieee.org/document/10172155/
74. https://www.multiresearchjournal.com/arclist/list-2024.4.6/id-4262
75. https://elibrary.imf.org/openurl?genre=journal&issn=1018-5941&volume=2023&issue=027
76. https://ieeexplore.ieee.org/document/10380440/
77. https://ijsab.com/volume-32-issue-1/6438
78. https://www.mdpi.com/2071-1050/13/17/9762/pdf
79. https://www.frontiersin.org/articles/10.3389/fgwh.2022.696529/full
80. https://jaesj.journals.ekb.eg/article_398940.html
81. https://www.mdpi.com/1911-8074/17/4/133
82. https://fepbl.com/index.php/csitrj/article/view/577
83. https://www.mdpi.com/2071-1050/13/16/8996
84. https://www.mdpi.com/2071-1050/13/16/8996/pdf
85. https://www.mdpi.com/2071-1050/12/1/49
86. https://www.mdpi.com/2079-9292/5/4/65/pdf?version=1475056683
87. https://www.adb.org/sites/default/files/publication/939786/source-multilateral-platform-sustainable-infrastructure.pdf
88. https://www.mdpi.com/1911-8074/17/4/133/pdf?version=1711119244
89. https://www.frontiersin.org/articles/10.3389/fpsyg.2022.821979/pdf
90. https://arxiv.org/pdf/2208.04710.pdf
91. https://www.mdpi.com/2071-1050/12/9/3893/pdf
92. https://www.bci.ca/adapting-risk-models-for-todays-infrastructure-investment-opportunities/
93. https://ctu.ieee.org/blog/2023/02/03/solutions-to-the-digital-divide-moving-toward-a-more-equitable-future/
94. https://www.pwc.com/gx/en/industries/tmt/digital-infrastructures-defining-moment-on-climate.html
95. https://daijobu.ai/2025/05/19/millions-of-tokens-the-invisible-unit-of-measurement-shaping-modern-ai/
96. https://martinhilbert.net/CheapEnoughWD_Hilbert_pre-print.pdf
97. https://www.ey.com/en_sg/media/podcasts/moneymultiple/2025/06/unlocking-value-in-growing-digital-infrastructure
98. https://www.getmonetizely.com/articles/should-your-ai-agent-use-token-based-or-subscription-pricing
99. https://arxiv.org/pdf/2402.09697.pdf
100. https://premierscience.com/wp-content/uploads/2024/11/pjcs-24-356-.pdf
101. https://www.sciencedirect.com/science/article/pii/S0148296323001157
102. https://www.linkedin.com/pulse/provisioned-capacity-ai-beginners-guide-dedicated-vs-asaf-liveanu-i1mhe
103. https://www.brookings.edu/articles/fixing-the-global-digital-divide-and-digital-access-gap/
104. https://www.semanticscholar.org/paper/cef1b80bd30d8c31beb37bc73cf1f15a37962008
105. https://ieeexplore.ieee.org/document/10825591/
106. https://academic.oup.com/jamiaopen/article/doi/10.1093/jamiaopen/ooaf055/8161131
107. https://arxiv.org/abs/2508.19008
108. https://mededu.jmir.org/2025/1/e67244

109. https://ieeexplore.ieee.org/document/11170906/
110. https://www.mdpi.com/0718-1876/17/4/63/pdf?version=1664164923
111. https://ej-ai.org/index.php/ejai/article/view/82
112. https://www.cureus.com/articles/350635-preparing-for-vascular-surgery-board-certification-a-comparative-study-using-large-language-models
113. https://theaspd.com/index.php/ijes/article/view/10923
114. https://ascopubs.org/doi/10.1200/JCO.2025.43.16_suppl.e21598
115. http://arxiv.org/pdf/2503.18129.pdf
116. https://arxiv.org/pdf/2502.07736.pdf
117. https://arxiv.org/html/2410.17950
118. https://arxiv.org/pdf/2407.10834.pdf
119. https://arxiv.org/pdf/2402.11754.pdf
120. http://arxiv.org/pdf/2406.06565.pdf
121. https://arxiv.org/pdf/2409.01666.pdf
122. https://arxiv.org/abs/2403.06150
123. https://langtail.com/llm-price-comparison
124. https://www.sciencedirect.com/science/article/pii/S0899825623000726
125. https://jurnal.unikal.ac.id/index.php/hk/article/view/3664
126. https://www.tandfonline.com/doi/full/10.1080/09537325.2022.2088342
127. https://link.springer.com/10.1007/s10479-021-04036-w
128. https://www.tandfonline.com/doi/full/10.1080/07421222.2023.2229122
129. https://www.ssrn.com/abstract=5218518
130. https://breached.company/red-sea-cable-cuts-the-hidden-crisis-threatening-global-internet-infrastructure/
131. https://journals.sagepub.com/doi/10.1177/10591478241305333
132. https://journal.uinjkt.ac.id/index.php/etikonomi/article/view/33892
133. https://www.semanticscholar.org/paper/f61fed43aa07694fa1df0a4ead140ed1ac39a4bf
134. https://www.semanticscholar.org/paper/e7fbad668e5b950d901cf706b8b300b2a28958c6
135. http://arxiv.org/pdf/1007.4586.pdf
136. https://linkinghub.elsevier.com/retrieve/pii/S0148296322005689
137. https://arxiv.org/pdf/1904.05656.pdf
138. http://www.scholink.org/ojs/index.php/ibes/article/download/4410/4994
139. https://arxiv.org/pdf/2303.13295.pdf
140. https://nottingham-repository.worktribe.com/preview/943810/Cred_SubEval_EJ-Style.pdf
141. https://link.springer.com/10.1007/s10257-025-00702-9
142. https://downloads.hindawi.com/journals/jam/2023/4456931.pdf
143. https://aisel.aisnet.org/ecis2018_rp/147/
144. https://conjointly.com/blog/willingness-to-pay/
145. https://journals.sagepub.com/doi/10.1177/20539517211069632
146. https://www.quantilope.com/resources/how-to-conduct-pricing-research-using-conjoint-analysis

147. https://www.productfocus.com/willingness-to-pay-the-hidden-engine-behind-effective-pricing/
148. https://arxiv.org/pdf/2105.01441.pdf
149. https://labs.adaline.ai/p/token-burnout-why-ai-costs-are-climbing
150. https://www.abacademies.org/articles/pricing-strategies-in-a-digital-economy-a-microeconomic-perspective-17498.html
151. https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-025-12321-8
152. https://www.tamarly.ai/blog-2-1/melvines-ai-analysis-12-understanding-tokens-and-the-costs-of-large-language-models-llms-for-enterprises
153. https://www.sciencedirect.com/science/article/abs/pii/S0925527321001225
154. https://ijsi.in/wp-content/uploads/2025/07/18.02.024.20251003.pdf
155. https://cloudwars.com/ai/enterprise-ai-minute/breaking-down-token-based-pricing-for-generative-ai-large-language-models-llms/
156. https://ojs.apspublisher.com/index.php/amit/article/download/391/300/769
157. https://www.renascence.io/journal/bounded-rationality-customers-simplified-decision-making-processes
158. https://www.mdpi.com/0718-1876/20/3/201
159. http://www.emerald.com/jeim/article/34/5/1429-1451/216071
160. https://ieeexplore.ieee.org/document/9934060/
161. https://www.ijecs.in/index.php/ijecs/article/view/4447
162. https://noonomy-journal.ru/images/3_1_2024/3_1_10.pdf
163. https://www.mdpi.com/2071-1050/16/19/8545
164. https://ieeexplore.ieee.org/document/10502110/
165. https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-024-10777-8
166. https://link.springer.com/10.1007/978-3-031-43185-2_10
167. https://journals.sagepub.com/doi/10.1177/10946705231173116
168. https://www.allsocialsciencejournal.com/search?q=SER-2025-3-064&search=search
169. http://arxiv.org/pdf/2404.00311.pdf
170. https://www.iiakm.org/ojakm/articles/2023/OJAKM_Volume11_2pp1-24.php
171. http://arxiv.org/pdf/1506.06648.pdf
172. https://journals.sagepub.com/doi/pdf/10.1177/10946705231173116
173. https://www.mdpi.com/2071-1050/14/19/11954/pdf?version=1663838570
174. http://arxiv.org/pdf/2503.21448.pdf
175. https://www.mdpi.com/2227-7072/6/4/87/pdf
176. https://www.mdpi.com/2071-1050/13/24/13701/pdf?version=1639475248
177. https://arxiv.org/html/2407.05484v1
178. https://sevenfigureagency.com/implementing-value-based-pricing-in-digital-agencies/
179. http://www.tandfonline.com/doi/abs/10.1080/00207543.2014.922707
180. https://easydigitaldownloads.com/blog/value-based-pricing-for-digital-products-and-services/
181. https://www.ibbaka.com/ibbaka-market-blog/pricing-patterns-for-generative-ai
182. https://sevenfigureagency.com/implementing-value-based-pricing-for-digital-agencies/

183. https://socsc.ktu.lt/index.php/Social/article/view/14247/7540
184. https://www.willingnesstopay.com/webinar/agentic-ai-pricing-4-of-6-ai-pricing-models---part-2-tokens-credit-systems
185. https://www.salesforce.com/sales/cpq/value-based-pricing/
186. https://journals.umcs.pl/h/article/download/1742/1357
187. https://www.youtube.com/watch?v=ZHIwPwAPzlA
188. http://mecs-press.org/ijieeb/ijieeb-v15-n3/v15n3-3.html
189. https://ges.jvolsu.com/index.php/en/component/attachments/download/1848
190. https://www.semanticscholar.org/paper/974665e62c139c842cb12359ea08e20222904f10
191. https://www.semanticscholar.org/paper/2875e7d26ede8bbc59a0d4bc2e187d369e72a15c
192. https://www.semanticscholar.org/paper/62d07091dfd8deb3f688b7599e563e04534ab415
193. https://journals.sagepub.com/doi/10.1016/j.ausmj.2019.07.002
194. https://www.emerald.com/insight/content/doi/10.1108/JIDE-08-2021-0004/full/pdf?title=the-achilles-tendon-of-dynamic-pricing-the-effect-of-consumers-fairness-preferences-on-platforms-dynamic-pricing-strategies
195. https://www.ccsenet.org/journal/index.php/ibr/article/download/66540/36058
196. https://www.mdpi.com/0718-1876/18/3/60/pdf?version=1688609520
197. https://ejbe.org/EJBE2021Vol14No28p107-J-GOTMARE.pdf
198. https://asistdl.onlinelibrary.wiley.com/doi/10.1002/pra2.2015.145052010043
199. https://journals.sagepub.com/doi/10.1177/21582440241293304
200. https://www.tandfonline.com/doi/pdf/10.1080/1331677X.2020.1844587?needAccess=true
201. https://www.fiegenbaum.solutions/en/blog/dramatic-drop-ai-token-prices-opportunities-challenges-sustainability
202. https://papers.ssrn.com/sol3/Delivery.cfm/5251923.pdf?abstractid=5251923&mirid=1
203. https://journals.sagepub.com/doi/10.3233/ISU-240230
204. https://www.untaylored.com/post/demystifying-the-lock-in-business-model-a-comprehensive-explanation
205. https://www.sciencedirect.com/science/article/pii/S305070062500043X
206. https://www.ewadirect.com/proceedings/aemps/article/view/25070
207. https://pubsonline.informs.org/doi/10.1287/mnsc.2023.4917
208. https://onlinelibrary.wiley.com/doi/10.1002/mde.4460
209. https://pubsonline.informs.org/doi/10.1287/mnsc.2022.4530
210. https://s-lib.com/en/issues/eiu_2025_05_v8_a22/
211. https://www.mdpi.com/0718-1876/20/4/286