Braess's Paradox of Generative AI

Boaz Taitler, Omer Ben-Porat

Technion—Israel Institute of Technology, Israel
boaztaitler@campus.technion.ac.il, omerbp@technion.ac.il

Abstract

ChatGPT has established Generative AI (GenAI) as a significant technological advancement. However, GenAI's intricate relationship with competing platforms and its downstream impact on users remains under-explored. This paper initiates the study of GenAI's long-term social impact resulting from the weakening network effect of human-based platforms like Stack Overflow. First, we study GenAI's revenue-maximization optimization problem. We develop an approximately optimal solution and show that the optimal solution has a non-cyclic structure. Then, we analyze the social impact, showing that GenAI could be socially harmful. Specifically, we present an analog to Braess's paradox in which all users would be better off without GenAI. Finally, we develop necessary and sufficient conditions for a regulator with incomplete information to ensure that GenAI is socially beneficial.

1 Introduction

ChatGPT has made Generative AI (GenAI) a household name, capturing the attention of the general public as the first fully operational GenAI (Naveed et al. 2023; Chang et al. 2023). Its widespread adoption stems from its capability to generate coherent texts and answer complex queries, showcasing human-like performance and even surpassing it (Katz et al. 2024). Beyond mere functionality, GenAI has revolutionized various applications, ranging from automating customer service interactions to facilitating creative writing endeavors. This transformative impact is underpinned by GenAI's strengths, which are attributed to an extensive corpus used during its training phase, encompassing many fields and languages.

In contrast to those benefits, there is a growing body of work raising various concerns regarding GenAI, such as biases (Abid, Farooqi, and Zou 2021), abusive and inappropriate content (Solaiman et al. 2023), and negative effects on users' behavior (Dvorak et al. 2024; Abbas, Jam, and Khan 2024). On top of those concerns, frequent training on high-quality data is a major issue (Chen, Zaharia, and Zou 2023). Updated training data is crucial to the quality of GenAI's answers, necessitating the continual incorporation of new data to reflect current trends and changes.

To illustrate, consider ChatGPT and Stack Overflow as competing platforms. ChatGPT utilizes the questions and answers on Stack Overflow as its training data, providing users with a quick answer tailored to their questions. Consequently, users shift towards ChatGPT and neglect Stack Overflow as a forum for their queries (as witnessed by recent work (del Rio-Chanona, Laurentsyeva, and Wachs 2023)). Over time, without additional training, the absence of fresh training data hampers ChatGPT's ability to provide accurate responses about new programming packages, while the diminished user activity on Stack Overflow affects both the volume and quality of answers available on the platform. This harmful combination negatively impacts the welfare of users, who suffer from outdated information and a lack of engagement in previously thriving community forums.

While the literature on ChatGPT and, more broadly, GenAI grows at an exponentially rapid pace, most of the research on ChatGPT since its release has focused on its performance (Frieder et al. 2024; Kocoń et al. 2023; Chen, Zaharia, and Zou 2023) and its applications across diverse domains (Kasneci et al. 2023; Liu et al. 2024). However, research has allocated scant attention, if any, to its social impacts concerning its intricate relationship with other competing platforms. Although GenAI is unanimously useful, its social impact prompts a critical research question: Is the existence of GenAI as Q&A platforms genuinely beneficial to its users, or would they be better off without it? Specifically, is the way we use GenAI socially beneficial in the long term?

Our contribution This paper initiates the study on GenAI's long-term social impact due to the weakening network effect of human-based platforms like Stack Overflow. We propose a sequential model that includes two platforms: A generative AI-powered service we term GenAI, and a human-based forum called Forum in which users can ask questions and answer questions of others. In our model, GenAI gains revenue from increased user traffic and incurs training and maintenance costs. GenAI chooses a *training scheme*: Strategically deciding in which rounds to train on fresh data. In contrast, Forum is a passive (i.e., non-strategic). A population of users utilizes GenAI and Forum. We assume that user utility from GenAI decreases between training rounds. We further assume that Forum has a network effect: User utility increases as the proportion of users using it rises. In our model, users choose according to a softmax function of the utilities from GenAI and Forum, allowing users to have different sensitivity levels to their utility.

First, we address GenAI's revenue-maximization task. We provide an efficient approximately optimal algorithm and show that the optimal training strategy has an unstable structure. Secondly, we examine the social impact. We say that a training scheme is *socially harmful* if users are better off in a world without GenAI, i.e., if users use Forum solely. We show the following surprising phenomena.

Theorem 1 (Braess's paradox of generative AI)**.** *The optimal training scheme of GenAI could be socially harmful.*

Our result resembles Braess's paradox (Braess 1968): A paradoxical behavior of transport networks, where adding an extra road can increase travel time. Analogously, even though generative AI provides high-quality service that beats human-based alternatives in the short term, it could lead to deteriorated welfare in the long term. This counter-intuitive observation stems from GenAI's negative impact on the network effect of Forum. After enough time without training, users who use GenAI receive low-quality service but are locked into GenAI, since Forum provides low-quality service as well without a strong user traffic. Since GenAI wishes to maximize its revenue, it could train sparsely and still gain most of the user traffic. We further analyze the Price of Anarchy (Roughgarden 2005; Koutsoupias and Papadimitriou 1999) and show that it is unbounded.

Thirdly, we adopt the point of view of a regulator. Theorem 1 implies that regulation could be required to ensure that social welfare is greater than the counterfactual welfare in a world without GenAI. Since the regulator is typically not informed about GenAI's training scheme, ensuring GenAI is socially beneficial poses a challenge. To address this, we develop necessary and sufficient conditions for social benefit and show that they are tight.

1.1 Related Work

Our work is inspired by the mass adoption of generative AI, contributing to an emergent line of work on foundation models and game theory (Laufer, Kleinberg, and Heidari 2024; Yao et al. 2024; Esmaeili et al. 2024; Conitzer et al. 2024; Dean et al. 2024b). This line of work includes challenges in the training process inspired by social choice theory (Conitzer et al. 2024) and mechanism design (Sun et al. 2024), ways to cut training costs by pooling (Huang, Vishwakarma, and Sala 2023), and revenue sharing between different actors (Laufer, Kleinberg, and Heidari 2024). In all of these cases, as argued by Dean et al. (2024b), planners should aim to understand societal impacts via mathematical modeling. Most relevant to our work is a recent work on content creation competition (Yao et al. 2024). It models a Tullock contest (Tullock 1980) between content creators, where some content creators use generative AI to create high-quality content, competing for users' engagement. The quality of the AI-generated content depends on the quality of the human-generated content. In a broader perspective, our work relates to social considerations in Machine Learning (see, e.g., (Caton and Haas 2024) for a recent survey). In this literature, social planners aim to restrict the output of a machine learning-based system to achieve long-term social welfare.

Our work also considers competition between platforms, a well-established topic in the economic literature (Rietveld and Schilling 2021; Karle, Peitz, and Reisinger 2020; Bergemann and Bonatti 2024). In the computer science literature, recent works study the equilibrium structure in competition between strategic cloud providers (Ashlagi, Tennenholtz, and Zohar 2010) and optimization algorithms (Immorlica et al. 2011). Other work studies competition in multi-learner settings (Dean et al. 2024a; Ginart et al. 2021; Ben-Porat and Tennenholtz 2019; Aridor et al. 2019). Additionally, a recent work analyzes how social welfare behaves in data-driven marketplaces (Jagadeesan, Jordan, and Haghtalab 2023), which is similar to our view of GenAI.

Finally, a crucial part of our model is Forum's network effect, which is inspired by the vast economic literature on this topic (Katz and Shapiro 1985; Rochet and Tirole 2003; Katz and Shapiro 1986; Feldman, Meir, and Tennenholtz 2013).

2 Model

General overview of the ecosystem The setting evolves over T discrete rounds, where in each round, a population of users chooses between GenAI and Forum for posing their questions. GenAI and Forum represent a Generative AI system and a Q&A forum, respectively. We take the perspective of GenAI, whose decisions are training times. In each round t, we denote by $x_t \in \{0, 1\}$ whether GenAI trains or not. We further let $\mathbf{x} = (x_1, \dots x_T)$ denote the *training scheme* of GenAI and always assume that GenAI trains on the first round, that is, $x_1 = 1$. Users' decisions are stochastic; we let p_t denote the proportion of users selecting GenAI in round t, with the complementary $1 - p_t$ selecting Forum. As we explain later, p_t depends on several elements including the training scheme $\mathbf{x}$, i.e., $p_t = p_t(\mathbf{x})$.

An instance of our problem is represented by the tuple $\langle r, c_m, c_{train}, R^c, R^s, \beta, p_1 \rangle$; we now elaborate on the components of the model.

GenAI The *revenue* of GenAI consists of three key components. Firstly, GenAI gets a reward of $r \in \mathbb{R}_+$ from users utilizing it, where r could represent direct payment, indirect income from advertising, etc. Secondly, GenAI pays a maintenance cost $c_m \in \mathbb{R}_+$. Thirdly, every time GenAI trains it incurs an additional cost of $c_{train} \in \mathbb{R}_+$. Altogether, the revenue of GenAI in round t is $v_t(\mathbf{x})$, where

$$v_t(\mathbf{x}) = p_t(\mathbf{x})r - c_m - x_t c_{train}.$$

We further let V denote the total revenue over the T rounds, namely, $V(\mathbf{x}) = \frac{1}{T}\sum_{t=1}^{T} v_t(\mathbf{x})$.

Users gain *utility* from using GenAI. We let $R^c : \mathbb{N} \to \mathbb{R}_+$ denote this utility, and assume that it decreases as the time since the last training increases. That is, $R^c(t - \tau)$ is the utility in round t if the last training of GenAI occurred in round $\tau < t$, which is decreasing in t. The gradual utility decrease captures several phenomena, such as outdated data like the identity of current Olympic medalists, or new technological developments that did not exist in its training data, like newly developed Python packages. Indeed, this assumption is well-supported empirically (Chen, Zaharia, and Zou 2023). Since user interaction with GenAI involves queries that yield varying degrees of relevance and timeliness, we think of R^c as the average utility users gain from GenAI. Other than being monotonically decreasing with respect to the last training time, we have no assumptions on the structure of R^c.

As we mentioned before, GenAI strategically picks a training scheme $\mathbf{x}$. For notational convenience, we define $\mathcal{T}(\mathbf{x}) = \{t | x_t = 1, t \in [T]\}$ to be the set of all training rounds in $\mathbf{x}$. Further, we let $\gamma_t = \gamma_t(\mathbf{x})$ denote the time between round t and its last training before that round, i.e., $\gamma_t(\mathbf{x}) = t - \max\{\mathcal{T}(\mathbf{x}) \cap [t]\}$. Therefore, we simplify our notations and use $R^c(\gamma_t(\mathbf{x}))$ to denote the user utility in round t.

Forum The revenue of Forum is also derived from the users who use it. We assume Forum is non-strategic, as its actions could be rather complex, and it is unclear whether or how such actions affect the ecosystem (see discussion in Section 6).[1] We denote by $R^s : [0,1] \to \mathbb{R}_+$ the average user utility from Forum in round t. We assume that R^s is time-invariant but is influenced by *network effects*, a well-established phenomenon empirically demonstrated in various studies (McIntyre and Srinivasan 2017; Katona, Zubcsek, and Sarvary 2011). That is, the utility R^s increases with the proportion of users $1 - p_t$ using Forum. To ease notation, we denote $R^s = R^s(p_t)$, which means R^s is decreasing as more users use GenAI. As before, $R^s(p_t)$ could have any general form, and we only assume it is a decreasing and differentiable function.

User behavior The *instantaneous welfare* in round t is the (weighted) average utility of users using GenAI and Forum, which is given by

$$u_t(\mathbf{x}) = p_t(\mathbf{x}) R^c(\gamma_t(\mathbf{x})) + (1 - p_t(\mathbf{x})) R^s(p_t(\mathbf{x})). \quad (1)$$

The (cumulative) *social welfare* $U : \{0,1\}^T \to \mathbb{R}_+$ is the sum of the instantaneous welfare over all rounds, $U(\mathbf{x}) = \sum_{t=1}^{T} u_t(\mathbf{x})$.

We now describe the proportions $(p_t)_t$. For $t = 1$, the proportion p_1 of users using GenAI is a model parameter. We let this quantity be determined exogenously as it could incorporate users' willingness to be early adopters, reflecting the population's appetite for innovation and novelty. For $t > 1$, we assume that user decisions are stochastic—they assign probabilities to GenAI and Forum based on the utilities the platforms yield. Further, users are Markovian: They base their decisions on the rewards from the preceding round $t-1$, but are otherwise independent of their past decisions. We formulate p_t as a softmax function of the utilities from GenAI and Forum from the last round,

$$p_t(\mathbf{x}) = \frac{\exp(\beta R^c(\gamma_{t-1}(\mathbf{x})))}{\exp(\beta R^c(\gamma_{t-1}(\mathbf{x}))) + \exp(\beta R^s(p_{t-1}(\mathbf{x})))}, \quad (2)$$

where the temperature β is the decision *sensitivity* parameter, determining how sensitive users are to their utility. For instance, $\beta \to \infty$ suggests that users are utility maximizers, whereas $\beta = 0$ indicates that users are indifferent to their utility and choose uniformly between GenAI and Forum.[2]

[1] For instance, Forum could sell data or develop its own GenAI. While working on this paper, it was reported that Stack Overflow partnered with OpenAI (Stack Overflow Blog 2024).

Assumptions Recall that we assume that R^c decreases with the time from the last training and that R^s decreases as more users rely on GenAI. Further, for our model to be realistic, we need to make sure that both GenAI and Forum could be attractive to users. Formally, we assume that there exists $t \in \mathbb{N}$ such that

$$R^c(t) < R^s(0) < R^c(0). \quad (3)$$

The left inequality implies that GenAI becomes inferior if enough time has passed since its last training, while the right inequality suggests that GenAI is better than Forum immediately after training.

Example 1. Consider an instance with GenAI parameters $r = 1$, $c_m = 0.6$, $c_{train} = 0.504$, utility functions $R^c(t) = 3 \cdot 0.5^t$, $R^s(p) = 1 - p$, and user behavior parameters $\beta = 1$, $p_1 = 1$. Let the horizon be $T = 20$, and consider the training scheme that only train at $t = 1$, i.e., $\mathbf{x}^0 = (x_1^0, \ldots, x_T^0) = (1, 0, \ldots, 0)$. At $t = 1$, the proportion is $p_1(\mathbf{x}^0) = p_1$, and the number of rounds from the last training round is $\gamma_1 = 0$; therefore,

$$u_1(\mathbf{x}^0) = p_1 R^c(0) + (1 - p_1) R^s(p_1) = 1 \cdot 3 \cdot 0.5^0 = 3.$$

Notice that if all the users were to use Forum, then the users' instantaneous welfare would have been $R^s(0) = 1$, which is lower than $u_1(\mathbf{x}^0)$. Thus, at $t = 1$, the presence of GenAI is socially beneficial. Next, GenAI's revenue in round 1 is

$$v_1(\mathbf{x}^0) = p_1 r - c_m - x_1^0 c_{train} = -0.104.$$

Moving on to round $t = 2$, we compute the proportion $p_2(\mathbf{x}^0)$. The proportion $p_2(\mathbf{x}^0)$ depends on $R^c(\gamma_1) = R^c(0)$ and $R^s(p_1(\mathbf{x}^0)) = R^s(1)$; therefore,

$$p_2(\mathbf{x}^0) = \frac{e^{\beta R^c(0)}}{e^{\beta R^c(0)} + e^{\beta R^s(1)}} = 0.95.$$

In the second round, $\gamma_2 = 1$ since $x_2^0 = 0$ and GenAI does not train. Consequently, we have all the necessary information to compute $u_2(\mathbf{x}^0)$ and $v_2(\mathbf{x}^0)$.

We highlight two additional training schemes. We let $\mathbf{x}^r$ represent the revenue-maximizing scheme, which can be shown to train in rounds $t \in \{1, 4, 7, 9, 12, 14, 17\}$. Additionally, let $\mathbf{x}^w$ represent the socially optimal scheme, training in every round. Figure 1 demonstrates the social welfare of the three schemes over time, and the *counterfactual* welfare $tR^s(0)$ obtained when users can only use Forum.

We observe several findings. First, the three schemes generate slightly higher welfare in the first round than the counterfactual due to our assumption in Inequality (3). Second, as we expect, $\mathbf{x}^w$ generates the highest welfare over time. Third, $\mathbf{x}^0$ provides higher welfare than the counterfactual till $t = 6$, and then they reverse. Namely, $\mathbf{x}^0$ makes users worse off. In that regard, notice that $\mathbf{x}^r$, the revenue-maximizing scheme, turns out to be better than the counterfactual in the long run.

[2] In practice, users can post their questions sequentially on both platforms; we model this process as two separate rounds.

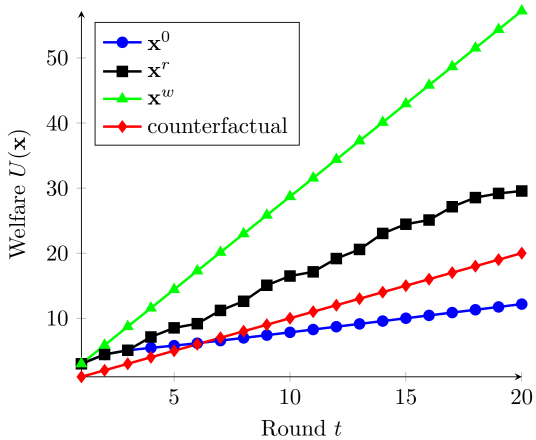


Figure 1: Social welfare over time in Example 1. The schemes $\mathbf{x}^0$, $\mathbf{x}^r$, and $\mathbf{x}^w$ are the no-training, revenue-maximizing, and welfare-maximizing schemes, respectively. The counterfactual welfare is the welfare users obtain in a world without GenAI.

3 GenAI's Revenue Maximization and Cyclic Training Schemes

In this subsection, we consider GenAI's revenue-maximization problem. We start by developing a dynamic programming-based approach to approximate the optimal revenue. Then, we highlight a class of simple training schemes—cyclic schemes, which train every fixed number of rounds. We show a surprising result: Cyclic schemes could be highly sub-optimal.

Recall that training schemes are binary vectors; hence, we can inefficiently find GenAI's revenue-maximizing scheme by considering all 2^T options. However, as we show next, there is an efficient algorithm that approximates the optimal revenue for a broad class of instances. We defer the full description of the algorithm to the appendix, and only write its formal guarantees.

Proposition 1. *Fix any instance such that R^s is L-Lipschitz satisfying $R^s(1) = 0$ and $\beta \cdot L < \frac{16}{7}$. For any $\varepsilon > 0$ such that $\varepsilon < \frac{16-7\beta L}{14\beta L e^{\beta L} T}$, there exists an algorithm that runs in $O\left(\frac{T^2}{\varepsilon}\right)$ time and returns a scheme $\mathbf{x}$ satisfying $V(\mathbf{x}) \geq \max_{\mathbf{x}'} V(\mathbf{x}') - \varepsilon r T$.*

Despite that Proposition 1 provides a way to approximate GenAI's revenue, such an approach has several weaknesses. First, from a practical standpoint, polynomial runtime can still be infeasible if the horizon T is large. Second, solutions obtained by dynamic programming can be uninterpretable, possibly hiding issues like large deviations in welfare over time. And third, GenAI's revenue may occasionally be negative for several consecutive rounds. Although we do not require a positive revenue balance over time in our model, this can be a weakness in real-world scenarios.

To that end, we examine a wide class of *cyclic training schemes*. As we show, cyclic training schemes offer certainty in both user utility and revenue. They create cycles in which utility and revenue can vary, but they become (asymptotically) constant over different cycles. Furthermore, cyclic training schemes could enjoy planned obsolescence (Bulow 1986; Lee and Lee 1998)—the practice of releasing new product versions at fixed intervals to encourage repeated purchases and maintain market share. The formal definition of a cyclic training scheme is as follows.

Definition 1 ($k-$*cyclic* training scheme)**.** *We say that $\mathbf{x}$ is k-cyclic for $k \in \mathbb{N}$ if*

$$x_t = \begin{cases} 1 & t \pmod{k} = 1 \\ 0 & \textit{otherwise} \end{cases}.$$

Note that for all $k \in \mathbb{N}$, a k-cyclic scheme always trains in the first round $t = 1$, which is consistent with our definition of a scheme.[3] As the next proposition shows, cyclic training schemes are stable over time.

Proposition 2. *Fix an instance and $k \in \mathbb{N}$. There exists a sequence of constants $(q_i^\star)_{i=1}^k$ such that for every $i \in [k]$, it holds that*

$$\lim_{l,T\to\infty} p_{l\cdot k+i}(\mathbf{x}_T^k) = q_i^\star,$$

where $\mathbf{x}_T^k$ is the k-cyclic scheme for T rounds.

As an immediate corollary of this proposition, we get that

Corollary 1. *It holds that*

- $\lim_{l,T\to\infty} v_{l\cdot k+i}(\mathbf{x}_T^k) = q_i^\star r - c_m - \mathbb{1}_{i=1} c_{train}$
- $\lim_{l,T\to\infty} u_{l\cdot k+i}(\mathbf{x}_T^k) = q_i^\star R^c(i-1) + (1-q_i^\star) R^s(q_i^\star)$.

Corollary 1 ensures that both welfare and revenue are (asymptotically) constant in every step of every cycle. That is, they can vary inside a cycle but become constant in the same step i across different cycles.

Intuitively, one would expect cyclic schemes to be optimal, or at least approximately optimal. Indeed, optimizing the cycle length k to maximize revenue-per-cycle seems to maximize the revenue V, as the revenue-per-cycle is guaranteed to converge. Surprisingly, this is not the case even for "standard" instances like the one in Example 1.

Theorem 2. *Fix the instance in Example 1. There exists $T^\star \in \mathbb{N}$, such that for any $T > T^\star$ and any cyclic training scheme $\mathbf{x}$, it holds that*

$$\max_{\mathbf{x}' \in \{0,1,\}^T} V(\mathbf{x}') - V(\mathbf{x}) = \Theta(1).$$

Proof sketch of Theorem 2. The proof of the theorem is rather technical. To prove the theorem, we need to show that a non-cyclic scheme outperforms *all* cyclic schemes by a non-negligible quantity. First, we present several key lemmas that allow us to focus on cycles with bounded length. Then, we analyze the convergence rate of the proportions from Proposition 2 using *contraction* of local neighborhoods. This allows us to upper-bound the revenue-per-cycle for any

[3] In our formal statements, we address a more general definition of cyclic schemes, allowing the prefix and suffix of a cyclic scheme to behave arbitrarily. However, as long as their length is constant w.r.t. to T, all of our results still hold; thus, we focus on this simpler form here.

cycle length and thus the revenue. Finally, we provide another (non-cyclic) training scheme and use similar techniques to lower-bound its revenue. The proof is completed by showing that the lower bound of our non-cyclic scheme is better than the highest upper bound of cyclic schemes. $\square$

4 Social Impact

In this section, we explore societal implications, demonstrating that the presence of GenAI can be *socially harmful*, reducing social welfare compared to its absence. We begin with a formal definition of social harmfulness and then show a counter-intuitive phenomenon: Optimal GenAI policies could be socially harmful. We flash out the main issue behind social harmfulness, which is prolonged periods without training. Later, Subsection 4.1 analyzes the Price of Anarchy, quantifying the inefficiency due to strategic behavior.

Next, we formally define the notion of social benefit.

Definition 2 (Socially beneficial scheme)**.** *We say that a training scheme* $\mathbf{x}$ *is* socially beneficial *if*

$$U(\mathbf{x}) \geq T \cdot R^s(0). \tag{4}$$

Definition 2 compares the social welfare with and without the presence of GenAI. The left-hand side of Inequality (4) is the social welfare in the presence of GenAI under the scheme $\mathbf{x}$. The right-hand side represents the counterfactual social welfare in the absence of GenAI, where all users choose Forum, similarly to the counterfactual welfare in Example 1. Conversely, if Inequality (4) does not hold, we say that $\mathbf{x}$ is *socially harmful*. Interestingly,

Observation 1 (Braess's paradox of generative AI)**.** *There exist instances where the optimal training scheme* $\mathbf{x}$ *is socially harmful.*

Observation 1 suggests that using GenAI can lead to deteriorated welfare, a paradox analogous to Braess's paradox in traffic networks (Braess 1968). Indeed, since GenAI can potentially provide a better utility than Forum, why does it occur?

Two simultaneous forces cause this paradox. First, when users rely on GenAI, they decrease their usage of Forum. Since user utility from Forum depends on network effects, lower usage results in reduced utility. And this is where the second force enters: To keep its dominance, GenAI need not train much. After prolonged periods without training, its utility diminishes but is still better than the low utility from Forum since its network effect is weak.

4.1 Price of Anarchy

In this subsection, we explore another useful way to quantify social harmfulness. We adopt the Price of Anarchy (PoA), measuring the inefficiency due to strategic behavior (Koutsoupias and Papadimitriou 1999; Roughgarden 2005). In our context, let $\mathcal{X}$ be the set of revenue-maximizing training schemes. For any instance I, the PoA is defined as

$$\text{PoA}(I) = \frac{\max_{\mathbf{x}} U(\mathbf{x})}{\min_{\mathbf{x}\in\mathcal{X}} U(\mathbf{x})}.$$

The next proposition demonstrates that the PoA is unbounded.

Proposition 3. *For every* $M \in \mathbb{R}_+$*, there exists an instance* I *with* $\text{PoA}(I) > M$*.*

Proof sketch of Proposition 3. To prove the proposition, we focus on the terms $R^s(1)$ and $R^c(t)$ for large values of t. The former is the utility from Forum when all users use GenAI, while the latter is the utility from GenAI after t rounds without training. Consider an instance with parameters that satisfy $R^c(t) > 0$, $R^s(1) = 0$ and $\beta = \infty$. For this selection of sensitivity, the users are strategic, i.e., $p_t(\mathbf{x}) = 1$ if $R^c(\gamma_{t-1}) > R^s(p_{t-1}(\mathbf{x}))$. Due to our assumption in Equation (3), this property pushes users to choose GenAI with $p_t(\mathbf{x}) = 1$ for every t and every training scheme as $R^c(t) > R^s(1) = 0$. GenAI has no incentive to train, and therefore the no-training scheme $\mathbf{x}^0$ is the revenue-maximizing scheme, inducing $U(\mathbf{x}^0) = \sum_{t=1}^{T} R^c(t-1)$. On the other hand, training in every round maximizes the welfare; hence, $\max_{\mathbf{x}} U(\mathbf{x}) = TR^c(0)$. Overall, $\text{PoA}(I) = \frac{TR^c(0)}{\sum_{t=1}^{T} R^c(t-1)}$, which could be arbitrarily large if $\lim_{t\to\infty} R^c(t) = 0$. $\square$

5 Regulating Training Frequency

Despite that GenAI could be socially harmful, as the previous section demonstrates, there are socially beneficial training schemes (e.g., training every round). In this section, we aim to identify interventions to GenAI's scheme that could mitigate its harmful effects. We develop a set of requirements that could be imposed on GenAI by a regulator.

The regulator's ability to intervene depends on its knowledge of the instance. We assume that the regulator has complete information on user-related parameters: The utility functions R^s, R^c and sensitivity β. If the regulator could demand GenAI to commit to a publicly known training scheme, it could verify that the scheme is socially beneficial. However, the regulator does not have access to any proprietary information belonging to GenAI and Forum. In particular, the regulator lacks knowledge of the scheme $\mathbf{x}$ and the resulting proportions $(p_t)_t$.

5.1 Welfare Between Two Consecutive Training Rounds

Our approach to ensure social benefit is to bound the maximal number of rounds between training. In this subsection, we explain why such a regulation is fundamental in guaranteeing social benefit. Let $\tau_1, \ldots, \tau_{|\mathcal{T}|}$ be the ordered elements of the training rounds $\mathcal{T}$, and let $\tau_{|\mathcal{T}|+1} = T + 1$. A sufficient condition for social benefit is that the inequality

$$\sum_{t=\tau_i}^{\tau_{i+1}-1} u_t(\mathbf{x}) \geq (\tau_{i+1} - \tau_i) \cdot R^s(0) \tag{5}$$

would hold for every $i \in \mathcal{T}$. To see why, fix a scheme $\mathbf{x}$. Notice that if Inequality (5) holds, then

$$U(\mathbf{x}) = \sum_{i=1}^{\mathcal{T}} \sum_{t=\tau_i}^{\tau_{i+1}-1} u_t(\mathbf{x}) \geq \sum_{i=1}^{\mathcal{T}} (\tau_{i+1}-\tau_i) \cdot R^s(0) = T \cdot R^s(0);$$

hence, $\mathbf{x}$ is socially beneficial by definition. As a result, ensuring Inequality (5) holds allows the regulator to ensure social benefit.

To that end, fix an arbitrary training round τ and let Δ denote the number of rounds between τ and the next training round (i.e., $\tau + \Delta \in \mathcal{T}$). Recall the definition of $u_t(\mathbf{x})$ in Equation (1). If the regulator knows $p_\tau(\mathbf{x})$, it could compute $u_\tau(\mathbf{x})$ as it only depends on $p_\tau(\mathbf{x})$ and the fact that $\gamma_\tau = 0$.[4]

The regulator could also compute $p_{\tau+t}(\mathbf{x})$ for any $t \in \{0, \dots, \Delta - 1\}$, due to the Markovian nature of the proportions; thus, it could compute $u_{\tau+t}(\mathbf{x})$ for those rounds as well. To sum, if the regulator knows $p_\tau(\mathbf{x})$, it could compute $\sum_{t=\tau}^{\tau+\Delta-1} u_t(\mathbf{x})$ and determine whether Inequality (5) holds. However, as noted before, the regulator cannot generally access the proportions $(p_t)_t$.

5.2 Bounding the Proportions

In this section, we develop non-trivial bounds on $\sum_{t=0}^{\Delta-1} u_{\tau+t}(\mathbf{x})$. Specifically, we use crude bounds on $p_\tau(\mathbf{x})$ for any scheme $\mathbf{x}$, and unravel the structure of the Markovian update of proportions to non-trivially bound $p_{\tau+t}(\mathbf{x})$ for any $t \in \{0, \dots, \Delta - 1\}$. The next technical lemma demonstrates the monotonicity of the induced proportions.

Lemma 1 (Monotonicity)**.** *Fix an instance and a training scheme $\mathbf{x}$, and let $(p_t)_{t=1}^T$ be the induced proportions. Further, let $(p'_t)_{t=1}^T$ be the induced proportion of the same instance except that the initial proportion is $p'_1 > p_1$. Then, for every $t \in [T]$ we have $p'_t > p_t$.*

Next, we introduce the auxiliary sequence $(q_t^\alpha)_{t=0}^\infty$ for every $\alpha \in [0, 1]$. We define q_t^α such that

$$q_t^\alpha = \begin{cases} \alpha & t = 0 \\ \frac{e^{\beta R^c(t-1)}}{e^{\beta R^c(t-1)} + e^{\beta R^s(q_{t-1}^\alpha)}} & t > 0 \end{cases}.$$

Note the resemblance between the recursive update of q and the definition of p_t in Equation (2). For instance, if $\alpha = p_\tau(\mathbf{x})$, we get $q_t^\alpha = p_{\tau+t}(\mathbf{x})$ for every $t \in \{0, \dots \Delta - 1\}$.

An immediate consequence of Lemma 1 is that

Corollary 2. *For every $t \in \{0, \dots \Delta - 1\}$, it holds that*

$$q_t^0 \le p_{\tau+t}(\mathbf{x}) \le q_t^1. \tag{6}$$

Notice that the left term of Inequality (6) is obtained by selecting $\alpha = 0$, while the right term is obtained by selecting $\alpha = 1$. Using the bounds in Corollary 2, we can sandwich the utility $u_{\tau+t}(\mathbf{x})$ for every $t \in \{0, \dots \Delta - 1\}$ by

$$\begin{aligned} \underline{u}_t &= q_t^0 R^c(t) + (1 - q_t^1) R^s(q_t^1), \\ \overline{u}_t &= q_t^1 R^c(t) + (1 - q_t^0) R^s(q_t^0). \end{aligned} \tag{7}$$

Indeed, the reader can verify that $\underline{u}_t \le u_{\tau+t}(\mathbf{x}) \le \overline{u}_t$. Equipped with Inequality (7), we are ready to give necessary and sufficient conditions.

Theorem 3. *Fix an instance and a scheme $\mathbf{x}$, and let Δ be the maximal number of rounds between two consecutive training rounds.*

- *Sufficient condition: If $\sum_{t=0}^{\Delta-1} \underline{u}_t \ge \Delta R^s(0)$ holds, then $\mathbf{x}$ is socially beneficial.*
- *Necessary condition: If $\mathbf{x}$ is socially beneficial, then $\sum_{t=0}^{\Delta-1} \overline{u}_t \ge \Delta R^s(0)$ holds.*

Theorem 3 provides the regulator with powerful tools. If the regulator is proactive, it can restrict GenAI to schemes that satisfy the sufficient condition. By doing so, it ensures that GenAI is socially beneficial, at the expense of a revenue decrease for GenAI. Alternatively, if the regulator is more passive and only wants to intervene when GenAI is guaranteed to be socially harmful, it could act as soon as the necessary condition is violated.

5.3 Access to Noisy Estimates

While Theorem 3 provides a practical approach, it does not quantify how far the bounds are from the actual utility. Specifically, the upper bound $\sum_{t=0}^{\Delta-1} \overline{u}_t$ and the lower bound $\sum_{t=0}^{\Delta-1} \underline{u}_t$ could be distant from the actual utility $\sum_{t=0}^{\Delta-1} u_{\tau+t}(\mathbf{x})$. This subsection provides a remedy. From here on, we limit our attention to instances where R^s is L-Lipschitz satisfying $R^s(1) = 0$ and $\beta L < \frac{16}{7}$.

Recall that we have previously assumed that the regulator is not informed about the proportions at all. However, in some cases, the regulator might have access to noisy estimates of the proportions. For instance, the regulator could survey the population and form a confidence interval on the proportion. Next, we shall assume that for every $t \in [T]$, the regulator can construct an estimate $\hat{p}$ such that $|p_t(\mathbf{x}) - \hat{p}| < \varepsilon$ for $\varepsilon > 0$. Furthermore, we assume ε is known to the regulator, i.e., it knows the size of the confidence interval.

In such a case, we can complement our monotonicity result (Lemma 1). The next theorem shows that the Markovian update of proportions forms a contraction mapping.

Theorem 4. *Fix any instance and scheme $\mathbf{x}$. Let $\varepsilon > 0$ such that $\varepsilon < \frac{16 - 7\beta L}{14\beta L e^{\beta L}}$. There exists $\gamma = \gamma(\beta, L, \epsilon)$, $\gamma \in (0, 1)$ such that if $|p_\tau(\mathbf{x}) - q_0^\alpha| < \varepsilon$, then for every $t \in \{0, \dots \Delta - 1\}$ it holds that $|p_{\tau+t}(\mathbf{x}) - q_t^\alpha| < \varepsilon \cdot \gamma^t$.*

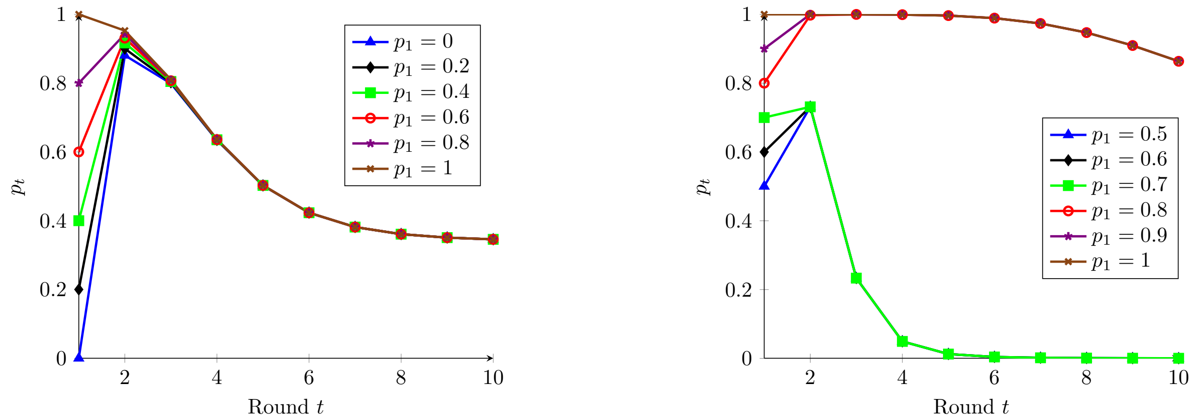
Theorem 4 implies that we expect the upper (or lower) bound q_t^α to approach the actual proportion $p_{\tau+t}$ at an exponential rate.

Example 2. Recall the instance in Example 1 and the no-training scheme $\mathbf{x}^0$. Figure 2a demonstrates contraction for $p_t(\mathbf{x})$ assuming different initial proportions p_1. After three rounds, the resulting $p_t(\mathbf{x}^0)$ is indistinguishable.

In contrast, consider the instance $R^c(t) = 1.1 \cdot 0.8^t$, $R^s(p) = \frac{1}{1 + e^{-100(0.8 - p)}}$, $\beta = 10$. Figure 2b examines how $p_t(\mathbf{x}^0)$ changes over time for various selections of p_1. This instance violates the condition $\beta L > \frac{16}{7}$, resulting in divergence. Specifically, for $p_1 \in [0, 0.7]$, the proportion in later rounds convergence to 0. For $p_1 \in [0.8, 1]$, the proportion converges to 1. Importantly, two close initial proportions in the range $(0.7, 0.8)$ could converge to different proportions in later rounds.

From here, we use Theorem 4 and the regulator's information $\hat{p}$ and ε to provide much tighter necessary and sufficient conditions. For every $t \in \{0, \dots, \Delta - 1\}$, instead of using the crude lower bound q_t^0 on $p_{\tau+t}(\mathbf{x})$ from Corollary 2, we use the more tighter bound $q_t^{\hat{p}-\varepsilon}$. Indeed, by the way

[4] To verify such a requirement is fulfilled, the regulator must be informed when GenAI trains. This is a realistic assumption that is also studied in recent research (Chen, Zaharia, and Zou 2023).



(a) Demonstrating contraction for the instance in Example 1.

(b) Demonstrating expansion for the instance in Example 2.

Figure 2: Demonstrating contraction and expansion of long-term proportions with varying initial proportions p_1. In Figure 2a, strong contraction applies. The proportion p_3 for $t = 3$ is almost invariant to p_1. In Figure 2b, the initial proportion determines whether the long-term proportion converges to 0 or 1, showcasing that even an ε-close estimate can be unfruitful to the regulator.

we define the auxiliary sequence q, if $\hat{p} - \varepsilon \leq p_\tau(\mathbf{x})$ then $q_0^{\hat{p}-\varepsilon} \leq p_\tau(\mathbf{x})$; thus, Lemma 1 ensures that $q_t^{\hat{p}-\varepsilon} \leq p_{\tau+t}(\mathbf{x})$ for every $t \in \{0, \dots, \Delta - 1\}$. We form a similar upper bound $q_t^{\hat{p}+\varepsilon} \geq p_{\tau+t}(\mathbf{x})$ since $\hat{p} + \varepsilon \geq p_\tau(\mathbf{x})$. To enhance readability, we use the shorter notation $\underline{q}_t$ for the lower bound, i.e., $\underline{q}_t = q_t^{\hat{p}-\varepsilon}$, and $\overline{q}_t = q_t^{\hat{p}-\varepsilon}$ for the upper bound. Therefore, we know that for every $t \in \{0, \dots, \Delta - 1\}$, $\underline{q}_t \leq p_{\tau+t}(\mathbf{x}) \leq \overline{q}_t$. Using these bounds on the proportion, we improve the previously proposed bounds on the instantaneous welfare from Equation (7),

$$\begin{aligned} \underline{u}_t^\varepsilon &= \underline{q}_t R^c(t) + (1 - \overline{q}_t) R^s(\overline{q}_t), \\ \overline{u}_t^\varepsilon &= \overline{q}_t R^c(t) + (1 - \underline{q}_t) R^s(\underline{q}_t). \end{aligned} \tag{8}$$

As before, we have $\underline{u}_t^\varepsilon \leq u_{\tau+t}(\mathbf{x}) \leq \overline{u}_t^\varepsilon$. Using Theorem 4, we show the following proposition.

Proposition 4. *Fix any instance and scheme $\mathbf{x}$. Let $\varepsilon > 0$ such that $\varepsilon < \frac{16-7\beta L}{28\beta L e^{\beta L}}$. There exists $\gamma = \gamma(\beta, L, \epsilon)$, $\gamma \in (0, 1)$ such that for every $t \in \{0, \dots, \Delta - 1\}$, it holds that*

$$\begin{aligned} 0 &\leq u_{\tau+t}(\mathbf{x}) - \underline{u}_t^\varepsilon \leq 2\varepsilon\gamma^t \left(R^c(t) + 2L\right), \\ 0 &\leq \overline{u}_t^\varepsilon - u_{\tau+t}(\mathbf{x}) \leq 2\varepsilon\gamma^t \left(R^c(t) + 2L\right). \end{aligned}$$

The next theorem combines all the results of this section.

Theorem 5. *Fix any instance and scheme $\mathbf{x}$. Let $\varepsilon > 0$ such that $\varepsilon < \frac{16-7\beta L}{28\beta L e^{\beta L}}$. There exists $\gamma = \gamma(\beta, L, \epsilon)$, $\gamma \in (0, 1)$ such that*

$$\sum_{t=0}^{\Delta-1} \overline{u}_t^\varepsilon - \sum_{t=0}^{\Delta-1} \underline{u}_t^\varepsilon \leq 4\varepsilon \sum_{t=0}^{\Delta-1} \gamma^t \left(R^c(t) + 2L\right). \tag{9}$$

The theorem assists the regulator in several ways. First, imagine a proactive regulator wishes to enforce the sufficient condition of Theorem 3 using the improved bounds $\underline{u}_t^\varepsilon$, namely to require that $\sum_{t=0}^{\Delta-1} \underline{u}_t^\varepsilon \geq \Delta R^s(0)$. In such a case, the regulator could argue that the margin between $\sum_{t=0}^{\Delta-1} \underline{u}_t^\varepsilon$ and the actual welfare $\sum_{t=0}^{\Delta-1} u_{t\tau+t}(\mathbf{x})$ is small; hence, this sufficient condition is not too stringent. Second, imagine a passive regulator that intervenes only if it *knows* that GenAI is socially harmful, that is, if $\sum_{t=0}^{\Delta-1} \overline{u}_t^\varepsilon < \Delta R^s(0)$. The passive regulator is guaranteed that if it does not intervene, the extent to which GenAI could be harmful is less than the right-hand-side of Inequality (9).

6 Discussion and Future Work

This paper initiates the research on the dynamics of the competition between generative AI and human-based platforms. After introducing the formal model, we studied GenAI's revenue-maximization problem. We have proposed an approximately optimal algorithm, and showed that optimal schemes are not cyclic. Then, we analyze social welfare. We demonstrated a Braess's paradox-like phenomena, where despite that GenAI is initially socially beneficial, its impact on Forum's network effects leads to deteriorated welfare. We also showed an infinite Price of Anarchy. Finally, we developed tools that could assist regulators to ensure that GenAI is socially beneficial in the long term without the need to use any proprietary information privately known to GenAI.

We see considerable scope for future work. From a technical perspective, some of our results apply to sub-classes of instances. For example, the statements in Subsection 5.3 are relevant in cases where $\beta L \leq \frac{16}{7}$. We suspect that this is a by-product of our analysis and the techniques we use to show contraction. Future work could overcome this limitation. More conceptually, our model assumes that Forum is non-strategic. Future work could relax this assumption, modeling Forum as a strategic player, e.g., a data seller, relating to works on information markets (Bergemann and Bonatti 2019; Chen et al. 2018; Chen, Xu, and Zheng 2020). In such a case, Forum could strategically decide what to sell, when to sell, and at which price, thereby affecting the welfare dynamics.

Acknowledgements

This research was supported by the Israel Science Foundation (ISF; Grant No. 3079/24).

References

Abbas, M.; Jam, F. A.; and Khan, T. I. 2024. Is it harmful or helpful? Examining the causes and consequences of generative AI usage among university students. *International Journal of Educational Technology in Higher Education*, 21(1): 10.

Abid, A.; Farooqi, M.; and Zou, J. 2021. Persistent anti-muslim bias in large language models. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, 298–306.

Aridor, G.; Liu, K.; Slivkins, A.; and Wu, Z. S. 2019. The perils of exploration under competition: A computational modeling approach. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, 171–172.

Ashlagi, I.; Tennenholtz, M.; and Zohar, A. 2010. Competing schedulers. In *Proceedings of the AAAI conference on artificial intelligence*, volume 24, 691–696.

Ben-Porat, O.; and Tennenholtz, M. 2019. Regression equilibrium. In *Proceedings of the 2019 ACM Conference on Economics and Computation*, 173–191.

Bergemann, D.; and Bonatti, A. 2019. Markets for information: An introduction. *Annual Review of Economics*, 11(1): 85–107.

Bergemann, D.; and Bonatti, A. 2024. Data, competition, and digital platforms. *American Economic Review*, 114(8): 2553–2595.

Braess, D. 1968. Über ein Paradoxon aus der Verkehrsplanung. *Unternehmensforschung*, 12: 258–268.

Bulow, J. 1986. An economic theory of planned obsolescence. *The Quarterly Journal of Economics*, 101(4): 729–749.

Caton, S.; and Haas, C. 2024. Fairness in machine learning: A survey. *ACM Computing Surveys*, 56(7): 1–38.

Chang, Y.; Wang, X.; Wang, J.; Wu, Y.; Yang, L.; Zhu, K.; Chen, H.; Yi, X.; Wang, C.; Wang, Y.; et al. 2023. A survey on evaluation of large language models. *ACM Transactions on Intelligent Systems and Technology*.

Chen, L.; Zaharia, M.; and Zou, J. 2023. How is ChatGPT's behavior changing over time? *arXiv preprint arXiv:2307.09009*.

Chen, Y.; Immorlica, N.; Lucier, B.; Syrgkanis, V.; and Ziani, J. 2018. Optimal data acquisition for statistical estimation. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, 27–44.

Chen, Y.; Xu, H.; and Zheng, S. 2020. Selling information through consulting. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, 2412–2431. SIAM.

Conitzer, V.; Freedman, R.; Heitzig, J.; Holliday, W. H.; Jacobs, B. M.; Lambert, N.; Mossé, M.; Pacuit, E.; Russell, S.; Schoelkopf, H.; Tewolde, E.; and Zwicker, W. S. 2024. Social Choice for AI Alignment: Dealing with Diverse Human Feedback. *CoRR*, abs/2404.10271.

Dean, S.; Curmei, M.; Ratliff, L.; Morgenstern, J.; and Fazel, M. 2024a. Emergent specialization from participation dynamics and multi-learner retraining. In *International Conference on Artificial Intelligence and Statistics*, 343–351. PMLR.

Dean, S.; Dong, E.; Jagadeesan, M.; and Leqi, L. 2024b. Accounting for AI and Users Shaping One Another: The Role of Mathematical Models. *arXiv preprint arXiv:2404.12366*.

del Rio-Chanona, M.; Laurentsyeva, N.; and Wachs, J. 2023. Are large language models a threat to digital public goods? evidence from activity on stack overflow. *arXiv preprint arXiv:2307.07367*.

Dvorak, F.; Stumpf, R.; Fehrler, S.; and Fischbacher, U. 2024. Generative AI Triggers Welfare-Reducing Decisions in Humans. *arXiv preprint arXiv:2401.12773*.

Esmaeili, S. A.; Bhawalkar, K.; Feng, Z.; Wang, D.; and Xu, H. 2024. How to Strategize Human Content Creation in the Era of GenAI? *arXiv preprint arXiv:2406.05187*.

Feldman, M.; Meir, R.; and Tennenholtz, M. 2013. Competition in the presence of social networks: how many service providers maximize welfare? In *International Conference on Web and Internet Economics*, 174–187. Springer.

Frieder, S.; Pinchetti, L.; Griffiths, R.-R.; Salvatori, T.; Lukasiewicz, T.; Petersen, P.; and Berner, J. 2024. Mathematical capabilities of chatgpt. *Advances in Neural Information Processing Systems*, 36.

Ginart, T.; Zhang, E.; Kwon, Y.; and Zou, J. 2021. Competing AI: How does competition feedback affect machine learning? In *International Conference on Artificial Intelligence and Statistics*, 1693–1701. PMLR.

Huang, T.-H.; Vishwakarma, H.; and Sala, F. 2023. Train'n trade: foundations of parameter markets. *Advances in Neural Information Processing Systems*, 36: 28478–28490.

Immorlica, N.; Kalai, A. T.; Lucier, B.; Moitra, A.; Postlewaite, A.; and Tennenholtz, M. 2011. Dueling algorithms. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, 215–224.

Jagadeesan, M.; Jordan, M. I.; and Haghtalab, N. 2023. Competition, alignment, and equilibria in digital marketplaces. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, 5689–5696.

Karle, H.; Peitz, M.; and Reisinger, M. 2020. Segmentation versus agglomeration: Competition between platforms with competitive sellers. *Journal of Political Economy*, 128(6): 2329–2374.

Kasneci, E.; Seßler, K.; Küchemann, S.; Bannert, M.; Dementieva, D.; Fischer, F.; Gasser, U.; Groh, G.; Günnemann, S.; Hüllermeier, E.; et al. 2023. ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences*, 103: 102274.

Katona, Z.; Zubcsek, P. P.; and Sarvary, M. 2011. Network effects and personal influences: The diffusion of an online social network. *Journal of marketing research*, 48(3): 425–443.

Katz, M. L.; and Shapiro, C. 1985. Network externalities, competition, and compatibility. *The American economic review*, 75(3): 424–440.

Katz, M. L.; and Shapiro, C. 1986. Technology adoption in the presence of network externalities. *Journal of political economy*, 94(4): 822–841.

Katz, U.; Cohen, E.; Shachar, E.; Somer, J.; Fink, A.; Morse, E.; Shreiber, B.; and Wolf, I. 2024. GPT versus Resident Physicians—A Benchmark Based on Official Board Scores. *NEJM AI*, 1(5): AIdbp2300192.

Kocoń, J.; Cichecki, I.; Kaszyca, O.; Kochanek, M.; Szydło, D.; Baran, J.; Bielaniewicz, J.; Gruza, M.; Janz, A.; Kanclerz, K.; et al. 2023. ChatGPT: Jack of all trades, master of none. *Information Fusion*, 101861.

Koutsoupias, E.; and Papadimitriou, C. 1999. Worst-case equilibria. In *Annual symposium on theoretical aspects of computer science*, 404–413. Springer.

Laufer, B.; Kleinberg, J.; and Heidari, H. 2024. Fine-Tuning Games: Bargaining and Adaptation for General-Purpose Models. In *Proceedings of the ACM on Web Conference 2024*, 66–76.

Lee, I. H.; and Lee, J. 1998. A theory of economic obsolescence. *The Journal of Industrial Economics*, 46(3): 383–401.

Liu, J.; Xia, C. S.; Wang, Y.; and Zhang, L. 2024. Is your code generated by chatgpt really correct? rigorous evaluation of large language models for code generation. *Advances in Neural Information Processing Systems*, 36.

McIntyre, D. P.; and Srinivasan, A. 2017. Networks, platforms, and strategy: Emerging views and next steps. *Strategic management journal*, 38(1): 141–160.

Naveed, H.; Khan, A. U.; Qiu, S.; Saqib, M.; Anwar, S.; Usman, M.; Barnes, N.; and Mian, A. 2023. A comprehensive overview of large language models. *arXiv preprint arXiv:2307.06435*.

Rietveld, J.; and Schilling, M. A. 2021. Platform competition: A systematic and interdisciplinary review of the literature. *Journal of Management*, 47(6): 1528–1563.

Rochet, J.-C.; and Tirole, J. 2003. Platform competition in two-sided markets. *Journal of the european economic association*, 1(4): 990–1029.

Roughgarden, T. 2005. *Selfish routing and the price of anarchy*. MIT press.

Solaiman, I.; Talat, Z.; Agnew, W.; Ahmad, L.; Baker, D.; Blodgett, S. L.; Daumé III, H.; Dodge, J.; Evans, E.; Hooker, S.; et al. 2023. Evaluating the Social Impact of Generative AI Systems in Systems and Society. *arXiv preprint arXiv:2306.05949*.

Stack Overflow Blog. 2024. Stack Overflow and OpenAI Partner to Strengthen the World's Most Popular Large Language Models. *Stack Overflow Blog*. Https://stackoverflow.co/company/press/archive/openai-partnership.

Sun, H.; Chen, Y.; Wang, S.; Chen, W.; and Deng, X. 2024. Mechanism Design for LLM Fine-tuning with Multiple Reward Models. *arXiv preprint arXiv:2405.16276*.

Tullock, G. 1980. Efficient Rent Seeking. In Buchanan, J. M.; Tollison, R. D.; and Tullock, G., eds., *Toward a Theory of the Rent-Seeking Society*, 97–112. College Station: Texas A and M University Press.

Yao, F.; Li, C.; Nekipelov, D.; Wang, H.; and Xu, H. 2024. Human vs. Generative AI in Content Creation Competition: Symbiosis or Conflict? *arXiv preprint arXiv:2402.15467*.