Competition and Diversity in Generative AI

Manish Raghavan*

Abstract

Recent evidence, both in the lab and in the wild, suggests that the use of generative artificial intelligence reduces the diversity of content produced. As generative AI penetrates more markets and reaches widespread adoption, should we still expect to see content homogenization? And more broadly, is homogeneity "bad" for society? Our work begins with a simple observation: content producers do not act in isolation. Instead, they compete with one another (e.g., for customers or attention). We explore the impact competition has on both content diversity and overall social welfare. Through a formal game-theoretic model, we show that competitive markets select for diverse AI models, mitigating monoculture. We further show that a generative AI model that performs well in isolation (i.e., according to a benchmark) may fail to provide value in a competitive market. Our results highlight the importance of evaluating generative AI models across the breadth of their output distributions, particularly when they will be deployed in competitive environments. We validate our results empirically by using language models to play Scattergories, a word game in which players are rewarded for answers that are both correct and unique. Overall, our results suggest that homogenization due to generative AI is unlikely to persist in competitive environments.

1 Introduction

A growing body of literature on generative artificial intelligence reveals a surprisingly consistent stylized fact: when people use generative AI tools, the set of content they produce tends to be more homogeneous than content produced by more traditional means [4, 23, 51, 58, 69, 71, 86, 109, 111]. Across a wide range of domains including peer review [58], writing [69], digital art [111], and survey responses [109], access to generative AI tools (GAITs) leads to less diverse outcomes. Researchers refer to this phenomenon—where the use of similar or identical underlying AI tools lead to convergence in outcomes—as *algorithmic monoculture* [52] or *homogenization* [13]. Much of the empirical literature on the subject treats homogenization itself as the primary object of study, seeking to quantify and deeply understand it.

In contrast, our focus will be on the economic forces that impact homogenization. As AI tools become increasingly pervasive, (when) should we expect to see a loss of content diversity? And what societal consequences should we anticipate? We will explore these questions through the lens of competition between content producers, examining how their

*MIT Sloan School of Management & Department of Electrical Engineering and Computer Science. mragh@mit.edu

strategic interactions influence the homogenization of content. To do so, we must first understand the externalities that competing producers impose on one another.

Externalities in production. Homogenization is not inherently negative. There exist scenarios in which society benefits from conformity in content production. For example, suppose producers of educational math videos begin using GAITs to produce visuals. Students may find it easier to learn if visuals are consistent across producers. Here, homogeneity could *increase* students' welfare, overall engagement, and thus video producer welfare as well.

On the other hand, we could give a slightly different interpretation of this example where the opposite is true. Suppose, instead, that students have different learning styles, where some students learn best from visuals of style A while others learn best through style B. If all producers begin using a GAIT that tends to produce style A, students with a preference for style B will be underserved, and social welfare will *decrease*. In this example, production of a particular style yields diminishing marginal returns, both to society and to individual producers (assuming producer utility increases with consumption of their content).

Producers thus impose externalities on one another. In our first hypothetical scenario, externalities were positive, and we would expect both producers and society to benefit from homogeneity. In the second scenario, externalities were negative, and everyone benefits from diversity: if the production of similar content imposes negative externalities, then producers (and society) are better off with more diverse content.

In many of the domains discussed earlier—creative ideation, writing, art, etc.—content producers face negative externalities when creating similar content. There are several plausible mechanisms here. The demand for a particular micro-genre is finite, and as more producers crowd the space, their audiences may fragment. Prices could also fall as lookalikes enter the market, leading to reduced producer surplus. Abstracting away from the particular mechanisms at play, we will restrict our attention to *settings in which externalities are negative*, and conformity reduces producers' utilities.

Scattergories as a metaphor. Our goal in this work is to build a theoretical model of competition via GAITs. To do so, we introduce a prototypical example of generative competition: the word game Scattergories.[1] We will argue that Scattergories is a natural setting in which to study effects of competition on the use of GAITs, and we will build a theoretical model that captures and generalizes its key elements.

In Scattergories, players must come up with a word or phrase that begins with a given letter and matches a given category.[2] Critically, a player only scores a point if *no other player chose the same phrase*. For example, given the letter 'S' and the category "Fish," a player could respond with "Salmon" or "Swordfish." But a player might be better off choosing a less common answer like "Snapper" since others might be less likely to think of it. By design, players in Scattergories impose negative externalities on one another.

Suppose all players use the same GAIT to produce answers, meaning each player samples an answer and plays it without making any attempt to validate its correctness before doing

[1] A variant of Scattergories has also been proposed as an alternative to the Turing test in prior work [50].

[2] The validity of an answer is typically determined by some agreement within the group of players.

so.[3] What impact would we expect algorithmic monoculture to have? At first glance, we'd expect players to be far more likely to generate the same responses, leading to lower scores. A slightly more nuanced view of GAITs reveals additional complexity here. GAITs are not static. Users have partial control over the distribution from which a GAIT samples (for example, by temperature scaling [34] or tuning a prompt).

Players can thus increase the diversity of the distribution from which a GAIT samples, but should they? Diversity may come at a cost. AI tools are optimized to sample "better" responses with higher likelihood, so diversifying a tool's distribution will shift probability mass towards "worse" responses. (In the extreme, sampling from the maximally diverse uniform distribution is unlikely to produce coherent text, let alone valid Scattergories answers.) This raises a trade-off: diversifying a GAIT's output distribution reduces collisions with other players but also decreases the likelihood of sampling valid answers.

The present work: Modeling monoculture and competition in content production. Scattergories has several properties that make it an ideal starting point to model competition with generative AI.

- Producers who use the same tool are more likely to create similar content.
- Producers who create similar content impose negative externalities on one another.
- Producers can diversify a tool (increase its randomness) at the cost of reduced quality.

We take these ingredients and generalize them into a formal theoretical model in Section 3. After deriving a series of results on the relationship between competition and diversity (Sections 4 and 5), we will empirically test them by using language models to play Scattergories against one another (Section 6). In our model, producers generate content from one of K distinct types, where K might be very large. In Scattergories, K would represent the number of possible responses, valid or otherwise, to a (letter, category) pair.[4] A producer who generates content of type $k \in [K]$ derives utility that depends on (1) the value of that type of content $\mathbf{d}_k$ (e.g., 1 if it is a valid Scattergories answer and 0 otherwise) and (2) the total number of players C_k who produce content of the same type.

Players use GAITs to produce content. We model a GAIT as a permutation π on the K answer types: while a GAIT's distribution can be tuned, we assume that a user cannot change the ranking of answer types. Intuitively, under our assumption, players cannot induce a GAIT to sample a "worse" answer with higher likelihood than a "better" one. We discuss this assumption in greater detail in Section 3 and attempt to validate it in Section 6.1.

Although it is directly inspired by Scattergories, our model describes a broad range of settings involving the use of generative AI. For example, travellers have begun to use GAITs to plan vacations [103]. But because travellers compete for limited resources like hotel rooms and restaurant reservations at popular destinations, they impose negative externalities on one another. Algorithmic monoculture could result in the same recommended destinations or activities, and savvy travellers or GAIT developers might diversify their tools' output distributions to increase the likelihood of sampling less crowded (though potentially less

[3] We briefly discuss the impact that discretion in accepting/rejecting proposals might have in Section 7.

[4] In our experiments, K can be on the order of 10^{30}.

desirable) destinations. As another example, many applications in marketing involve competitive creative production. More generally, whenever two GAIT users would prefer that, holding quality fixed, they each received distinct generations, our formulation may prove relevant.

Our model reveals several insights, which we explore in Sections 4 and 5. At a high level, we find that when players face stronger competition, they generate more diverse content, although players' behavior at equilibrium is less diverse than would be socially optimal. Perhaps more surprisingly, we find that the relative quality of generative AI tools also depends on the strength of competition. The "best" tool in isolation can be suboptimal in the face of competition. Intuitively, a tool that produces a single high-quality answer may be less useful than a tool that can sample across a diverse set of answers, even if they are on average lower-quality. Beyond these, we include results on the price of anarchy, the limiting behavior of competition, and tools' shares in an overall market.

We validate our theoretical findings using Scattergories as a case study in Section 6.[5] We confirm our key theoretical results: (1) more competition leads to more diverse production at equilibrium, (2) equilibria are less diverse than socially optimal behavior, and (3) the best tool in isolation (i.e., the one most likely to produce a correct response) is not necessarily the best tool in a competitive setting. We find substantial heterogeneity across tools, and our results suggest that users of GAITs in competitive environments should evaluate their performance across their output distributions.

By taking a game-theoretic perspective on homogenization via generative AI, our model sheds light on how economic forces might influence homogenization going forwards. Our findings indicate the importance of measuring and optimizing for *competitive alignment*, the performance of a tool relative to its competitors. We discuss the implications of our findings for the development, evaluation, and deployment of generative AI in Section 7.

2 Related Work

Algorithmic monoculture and homogenization. Theoretical work on algorithmic monoculture has typically focused on prediction or allocation tasks, where multiple previously independent decision-makers begin to correlate their behavior via identical or similar algorithmic components [13, 22, 46, 52, 73, 74, 96]. In these settings, results often suggest that monoculture harms the welfare of decision-makers by reducing the overall information in a system [48, 52, 73]. Related work also considers the trade-off (or lack thereof) between accuracy and diversity [75].

Recent work has studied homogenization through generative AI empirically in a number of different settings [4, 23, 51, 58, 69, 71, 86, 109, 111]. Many of these studies argue implicitly or explicitly that similarity imposes negative externalities. Researchers have articulated objections to homogenization via AI beyond the purely economic considerations discussed in this work, including its effects on language [35, 57], creativity [17], culture [25, 62], and science [28, 63]. Others argue that we can and should use generative AI to explicitly introduce diverse and stochastic ideas [18, 98].

[5] Code to reproduce our experiments can be found at [REDACTED FOR REVIEW].

Alignment. A slightly different perspective on diversifying generative AI distributions comes from *alignment*, in which researchers seek to align output distributions from AI tools with some notion of the "right" distribution [88]. This research often indicates that the use of reinforcement learning with human feedback (RLHF) is responsible for misalignment between a model's output distribution and its training distribution, leading to a form of mode collapse [51, 69, 71, 104]. All of the tools we experiment with in Section 6 have been instruction-tuned, so we should expect some sort of "narrowing" of their distributions relative to what their pre-trained baselines might produce. Other work on distributional alignment seeks to modify training or fine-tuning processes to promote output diversity [87, 101, 105]. Simple strategies like temperature scaling [34], which we use in our experiments in Section 6, sometimes improve the calibration of output distributions [49, 86, 95].

Empirical studies on distributional alignment tend to differ from our work in their conception of what the "right" distribution is. Examples from the literature include reference comparisons to training distributions [104], events with known distributions (e.g., coin flips) [60], and public opinion [84]. In contrast, we make no assumptions about what the distribution "should" be; our analysis takes downstream economic consequences as the primary measure of interest. As we will see, it may be optimal for a tool to be quite *misaligned* as long as it fills a niche left open by existing tools. Despite this, the Scattergories setting we introduce could be an interesting testbed for future research on alignment.

Related game-theoretic models. Much of the prior work on competition in content production studies incentives for content production in platforms mediated by recommender systems [e.g., 8, 10, 11, 31, 65, 106, 108].[6] The goal of this work is often to design mechanisms (i.e., recommender systems) that induce desirable incentives for content producers.

More similar to ours is a recent line of work, inspired by Hotelling [40], modeling competition between producers who generate content represented in a vector space to meet consumer demand, also represented as vectors [2, 41, 42, 44]. In these models, consumers choose the content that most closely aligns with their preference vectors. Our model has some similar features: we could think of a GAIT's distribution over K discrete types to be a K-dimensional vector. However, our "choice model" is quite different. Demand is not fixed, and there is no cross-type competition. While our model already yields interesting results in this relatively simple model of competition, future work could extend it to a setting where similarity (and therefore competition) is continuous, not binary.

Recent research also models competition between humans and generative AI. Yao et al. [107] and Esmaeili et al. [26] study a setting where AI continually learns from past content. Taitler and Ben-Porat [90] analyze the impacts of generative AI on incentives for human content producers in a question-answering forum like StackOverflow. In contrast, all players in our model use generative AI tools.

Analogous models to ours appear in a number of settings. For example, models of spatial competition in the ride-sharing literature often feature discrete locations where drivers choose to operate, and in those locations drivers impose externalities on one another [14, 16, 21, 30]. Parimutuel or racetrack betting markets also share some features, where winning bettors

[6] Other related models study algorithmic competition in curation [43] and decision-making Jagadeesan et al. [45].

divide the pot in proportion to the sizes of their bets [36, 94].

Our work also has connections to more traditional game theory. We can interpret our model as a class of atomic congestion games [e.g., 55, 78] with parallel edges, particular cost functions, and restrictions on the strategy space: Each player must choose a mixed strategy that obeys a ranking constraint. The game we study turns to be a *valid utility game* [100], which allows us to leverage existing price of anarchy results. It is also a *potential game* [66, 78] (see Lemma A.8), so best-response dynamics converge.

3 Model

Formally, we study a symmetric n-player game. Each player uses a GAIT T to sample one of K types. Players have partial control over the distribution $\mathbf{p}$ from which T samples. A player can diversify T's output distribution, but we will assume that they cannot alter the distribution to produce outputs deemed "worse" by the GAIT with higher likelihood than those deemed "better." Let $\Delta(\pi)$ be the set of distributions over $[K]$ that are ordered according with respect to a permutation π, meaning if $\mathbf{p} \in \Delta(\pi)$, then $\mathbf{p}_{\pi_1} \geq \mathbf{p}_{\pi_2} \geq \cdots \geq \mathbf{p}_{\pi_K}$. Throughout Sections 3 to 5, we will make the following assumption.

Assumption 1. For a GAIT T and a given instance, there exists a permutation π such that T can only sample from a distribution $\mathbf{p} \in \Delta(\pi)$.[7]

When analyzing a single GAIT T, we will assume without loss of generality that $\pi = [1, 2, \ldots, K]$. Subject to the constraint that $\mathbf{p} \in \Delta(\pi)$, we impose no further restrictions—that is, we assume each producer can tune T to produce any distribution $\mathbf{p} \in \Delta(\pi)$. This pair of assumptions—T imposes a static ordering, and subject to that ordering, players have complete control—will both be violated in practice. GAITs are not as rigid as fixed rankings, and producers will not in practice be able to take advantage of the $K - 1$ degrees of freedom afforded by $\Delta(\pi)$. Nonetheless, as we will show in Section 6, the key results from our model continue to hold empirically.

Having chosen a distribution $\mathbf{p} \in \Delta(\pi)$, each player samples type $k \in [K]$ independently with probability $\mathbf{p}_k$. Each type k has a known nonnegative expected value $\mathbf{d}_k$, which is the utility a producer receives if she is the only producer who samples it. But because producers impose negative externalities on one another, producers receive utility that decreases in the number of competitors who sample the same type. In particular, let C_k be the number of players who produce type k. Then, each receives utility $\mathbf{d}_k / s(C_k)$, where $s(\cdot)$ is an increasing function. For example, if $s(x) = x$, then players receive an equal share of the overall value $\mathbf{d}_k$. Choosing $s(x) = x^\infty$ yields the scoring rule for Scattergories: $\mathbf{d}_k$ if the response is unique, and 0 otherwise. We refer to s as a *score function*, and it plays a role similar to a classical congestion function.

In summary, our game $\mathscr{G}(n, \mathbf{d}, s)$ proceeds as follows:

1. Each player $i \in [n]$ chooses a strategy $\mathbf{p}^{(i)} \in \Delta(\pi)$.

[7] Note that this ordering is not universal. T may have different permutations on different instances, and two different GAITs T and T' may produce different permutations on a given instance. In our model, algorithmic monoculture manifests as a shared ranking π over types across all who use T.

2. Each player i's content type is sampled $k_i \sim \mathbf{p}^{(i)}$ independently.

3. Each player i derives utility $\mathbf{d}_{k_i}/s(C_{k_i})$, where C_{k_i} is the number of players (including i) who sampled type k_i.

Score functions and social welfare. To fully specify the model, we must provide a class of score functions. Throughout this paper, we will frequently use score functions of the form $s_\gamma(x) \triangleq x^\gamma$ for $\gamma > 0$. (Observe that when $\gamma = 0$, there is no competition, and players behave independently from one another.) The functions s_γ have several desirable properties. $s_\gamma(1) = 1$, meaning if a single player samples type k, they face no competition. s_γ is increasing, so more competition implies higher congestion. And $1/s_\gamma$ is convex, so marginal congestion is decreasing in the amount of competition. Intuitively, the jump from no competition to a single competitor is significant, but if competition is already strong, one additional competitor makes less of a marginal difference.

Our definition of s_γ reveals a slight subtlety that will impact how we think about social welfare, reflecting modeling differences in Hotelling games. In one formulation (for example, Hotelling games on disconnected graphs [29, 53]), prices are fixed (or 0), but demand served is not. Social welfare is the total demand served, which is simply the sum of player utilities:[8]

$$\sum_{i\in[n]} \mathbf{d}_{k_i} \cdot \frac{1}{s_\gamma(C_{k_i})} = \sum_{k\in[K]} \mathbf{d}_k \frac{C_k}{s_\gamma(C_k)} = \sum_{k\in[K]} \mathbf{d}_k C_k^{1-\gamma}. \tag{1}$$

This maps nicely to the case where $\gamma \leq 1$. While society experiences diminishing marginal gains from items of type k, overall demand met for type k increases in C_k. We could microfound this, for example, with a consumer model in which consumers who want content of type k are still picky and reject each item of type k with some probability [e.g., 75]. Thus, more production of type k increases overall demand for type k met.

In another formulation (e.g., Hotelling games with endogeneous prices [40]), demand is fixed but prices are not. Players who sample the same type face (potentially imperfect) price competition, meaning the total surplus players derive from type k (i.e., total demand met $\times$ price per unit demand) *decreases* in C_k. We ignore transfers in our definition of social welfare, meaning social welfare is simply the overall demand met. Taking $\mathbf{d}_k$ to be the fixed demand for each k, social welfare is

$$\sum_{k\in[K]} \mathbf{d}_k \mathbb{1}\left[C_k > 0\right]. \tag{2}$$

This corresponds to the case where $\gamma \geq 1$. Here, we could microfound a model in which demand for type k is fixed at $\mathbf{d}_k$, and as C_k increases, price competition drives down producer surplus per unit of demand met. Perfect (Bertrand) price competition would imply that producer i's utility is $\mathbf{d}_{k_i}$ when $C_{k_i} = 1$ and 0 if $C_{k_i} > 1$, corresponding to the score function $s_\infty(x) = x^\infty$.

We have thus described two slightly different classes of score functions (s_γ for $\gamma \leq 1$ and $\gamma \geq 1$) with different definitions of social welfare (1) and (2) respectively. These definitions

[8] By convention, we let $s(0) = 1$ for any s to ensure that denominators are always nonzero.

coincide at $\gamma = 1$. In what follows, we build on them to provide a more general description of the family of score functions for which our results will hold. Let f be discrete-convex if for any $x \in \mathbb{N}$,

$$f(2+x) - f(1+x) \geq f(1+x) - f(x).$$

f is discrete-concave if $-f$ is discrete-convex. Let $\mathbb{N}_{>0}$ be the set of nonnegative integers. With this, we can define our class of score functions $\mathcal{S}$.

Definition 3.1 (Score function)**.** A score function $s \in \mathcal{S}$ satisfies

D.1 $s(x)$ is increasing on $\mathbb{N}_{>0}$.

D.2 $1/s(x)$ is discrete-convex on $\mathbb{N}_{>0}$.

D.3 $s(1) = 1$ and $s(x) > 1$ for some $x \in \mathbb{N}_{>0}$.

D.4 At least one of the following two conditions:

D.4(a) $x/s(x)$ is increasing and discrete-concave on $\mathbb{N}_{>0}$.

D.4(b) $x/s(x)$ is decreasing on $\mathbb{N}_{>0}$.

Let $\mathcal{S}_{\nearrow}$ be the set of functions s that satisfy **D.1** to **D.3** and **D.4(a)**. Let $\mathcal{S}_{\searrow}$ be the set of functions s satisfying **D.1** to **D.3** and **D.4(b)**. By definition, $\mathcal{S} = \mathcal{S}_{\nearrow} \cup \mathcal{S}_{\searrow}$.

Observe that together, **D.2** and **D.3** imply $s(2) > 1$. We include a discrete-concavity requirement in **D.4(a)** so that social welfare exhibits diminishing returns in C_k for any type k. By this definition, $s_\gamma \in \mathcal{S}_{\nearrow}$ for $\gamma \in (0, 1]$ and $s_\gamma \in \mathcal{S}_{\searrow}$ for $\gamma \geq 1$ (including $\gamma = \infty$). The only function in their intersection $\mathcal{S}_{\nearrow} \cap \mathcal{S}_{\searrow}$ is the identity function $s_1(x) = x$.

4 Competition under a Single Tool

We are now ready to analyze behavior under our model. We begin in this section with the case where all players have access to a single GAIT with ordering $\pi = [1;\ 2;\ \ldots\ K]$ and can choose any $\mathbf{p} \in \Delta(\pi)$. In Section 5, we will generalize to the case where players can choose from a set Π of available GAITs. First, we discuss our solution concept, symmetric Nash equilibria, in Section 4.1. Then, in Section 4.2, we study the *diversity* of equilibria, finding that stronger competition induces more diversity, but equilibria are less diverse than socially optimal behavior. We also characterize equilibria in the limit of infinite competition. We show in Section 4.3 that the *price of anarchy* [55] is bounded by 2, limiting the extent to which equilibria can be suboptimal. Throughout this section, we will assume $n \geq 2$, $\mathbf{d} \in \mathbb{R}^K_{\geq 0}$, and $s \in \mathcal{S}$.

4.1 Equilibria

Our focus in this section will be on symmetric behavior. Because $\mathscr{G}$ is symmetric and utilities are continuous, it is guaranteed to have a symmetric mixed Nash equilibrium [9, 33]. Every mixed strategy is equivalent to a pure strategy since $\Delta(\pi)$ is convex (see Lemma A.2). Thus, $\mathscr{G}$ is guaranteed to have a symmetric pure Nash equilibrium.

Utility and social welfare. Let $X(n,p)$ denote a binomial random variable with those parameters. Given that the $n-1$ other players choose the strategy $\mathbf{p}$, player i's expected utility for strategy $\mathbf{p}'$ is

$$U(\mathbf{p}', \mathbf{p}; n, \mathbf{d}, s) \triangleq \sum_{k\in[K]} \mathbf{p}'_k \mathbf{d}_k \mathbb{E}\left[\frac{1}{s(C_k)} \mid k_i = k\right] = \sum_{k\in[K]} \mathbf{p}'_k \mathbf{d}_k \mathbb{E}\left[\frac{1}{s(1+X(n-1,\mathbf{p}_k))}\right].$$

When it is clear from context, we will omit the $(n, \mathbf{d}, s)$. In a slight abuse of notation, we will sometimes write $U(\mathbf{p}, \mathbf{p})$ as simply $U(\mathbf{p})$. $\mathbf{p}$ is a symmetric equilibrium for $\mathscr{G}(n, \mathbf{d}, s)$ if and only if $U(\mathbf{p}, \mathbf{p}; n, \mathbf{d}, s) \geq U(\mathbf{p}', \mathbf{p}; n, \mathbf{d}, s)$ for all $\mathbf{p}' \in \Delta(\pi)$. We next define expected social welfare for symmetric strategies, taking care to account for the different interpretations discussed in Section 3. In the case where $s \in \mathcal{S}_{\nearrow}$, following (1),

$$W(\mathbf{p}; n, \mathbf{d}, s) \triangleq \sum_{k\in[K]} \mathbf{d}_k \mathbb{E}\left[\frac{C_k}{s(C_k)}\right] = \sum_{k\in[K]} \mathbf{d}_k \mathbb{E}\left[\frac{X(n,\mathbf{p}_k)}{s(X(n,\mathbf{p}_k))}\right] = \sum_{i\in[n]} U(\mathbf{p}; n, \mathbf{d}, s).$$

When $s \in \mathcal{S}_{\searrow}$, following (2),

$$W(\mathbf{p}; n, \mathbf{d}, s) \triangleq \sum_{k\in[K]} \mathbf{d}_k \mathbb{E}\left[\mathbb{1}\left[C_k > 0\right]\right] = \sum_{k\in[K]} \mathbf{d}_k \Pr[X(n,\mathbf{p}_k) > 0] = \sum_{k\in[K]} \mathbf{d}_k \left(1 - (1-\mathbf{p}_k)^n\right).$$

Again, we will omit $(n, \mathbf{d}, s)$ when it is clear from context. Note that W no longer depends on s when $s \in \mathcal{S}_{\searrow}$, so welfare-maximizing strategies are identical for all $s \in \mathcal{S}_{\searrow}$.

Characterizing equilibria with decreasing d. We begin by analyzing the special case where $\mathbf{d}$ is decreasing in k, i.e., π perfectly ranks outputs by their expected value.

Assumption 2. $\mathbf{d}_k$ is weakly decreasing in k.

Under Assumption 2, we will show that Assumption 1 is no longer binding for symmetric strategies; intuitively, symmetric behavior should put more probability mass on better responses, so any symmetric equilibrium or socially optimal strategy $\mathbf{p}$ should be decreasing in k as well. Define

$$u_s(n,p) \triangleq \mathbb{E}\left[\frac{1}{s(1+X(n,p))}\right].$$

When it is clear from context, we will drop the subscript s and just write $u(n,p)$. With this, we can characterize equilibria:

Lemma 4.1. *Under Assumption 2, $\mathscr{G}(n, \mathbf{d}, s)$ has a unique symmetric equilibrium $\mathbf{p} = \mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s)$, which is a pure strategy equilibrium. Let $k^* = \max\{k : \mathbf{p}_k > 0\}$. $\mathbf{p}$ satisfies the following properties:*

1. $\mathbf{p} \in \Delta(\pi)$.
2. *For all $k \leq k^*$,*

$$\mathbf{d}_k u(n-1, \mathbf{p}_k) = U(\mathbf{p}).$$

3. *For all* $k > k^*$, $U(\mathbf{p}) \geq \mathbf{d}_k$.

This holds even if we remove Assumption 1 (i.e., if players can choose $\mathbf{p} \notin \Delta(\pi)$*).*

We defer all proofs in this section to Appendix A.1. For intuition, observe that $\mathbf{d}_k u(n-1, \mathbf{p}_k)$ is precisely a player's utility conditioned on sampling type k. Lemma 4.1 simply shows that conditional utility is equal across all types sampled with positive probability, which suffices to show that $\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s)$ is unique. We can find $\mathbf{p}_{\text{EQ}}$ via a simple water-filling algorithm: Beginning with $\mathbf{p}_k = 0$ for all k, increase $\mathbf{p}_1$ until a player's expected utility conditioned on sampling type 1 is equal to $\mathbf{d}_2$. Then, increase $\mathbf{p}_1$ and $\mathbf{p}_2$ simultaneously until their conditional expected utilities reach $\mathbf{d}_3$, and so on until the sum of probabilities is 1. Still under Assumption 2, we can show a similar result for socially optimal symmetric strategies, which we denote $\mathbf{p}_{\text{OPT}}(n, \mathbf{d}, s)$.[9]

Lemma 4.2. *Under Assumption 2,* $\mathscr{G}(n, \mathbf{d}, s)$ *has a unique symmetric optimal strategy* $\mathbf{p}^* = \mathbf{p}_{\text{OPT}}(n, \mathbf{d}, s)$*. Let* $c = \frac{1}{n}\|\nabla W(\mathbf{p}^*)\|_\infty$*, and let* $k^* = \max\{k : \mathbf{p}^*_k > 0\}$*.* $\mathbf{p}^*$ *satisfies the following properties:*

1. $\mathbf{p}^* \in \Delta(\pi)$.
2. *If* $s \in \mathcal{S}_\nearrow$*, for all* $k \leq k^*$*,*

$$\frac{1}{n}\left(\nabla W(\mathbf{p}^*)\right)_k = \frac{\partial}{\partial p}\mathbf{d}_k p u_s(n-1, p)\Big|_{p=\mathbf{p}^*_k} = c.$$

3. *If* $s \in \mathcal{S}_\searrow$*, for all* $k \leq k^*$*,* $\mathbf{d}_k(1-\mathbf{p}^*_k)^{n-1} = c$.
4. *For all* $k > k^*$*,* $c \geq \mathbf{d}_k$.

This holds even if we remove Assumption 1 (i.e., if players can choose $\mathbf{p} \notin \Delta(\pi)$*).*

Whereas equilibrium strategies equalize *conditional* utility for each k with nonzero $\mathbf{p}_k$, optimal strategies equalize *marginal* utility for each k with nonzero $\mathbf{p}^*_k$.

Removing the assumption that d is decreasing. Assumption 2 requires π to be perfect: it correctly ranks the K types. We cannot expect GAITs to be perfect in general. To what extent do Lemmas 4.1 and 4.2 hold for imperfect GAITs, where we no longer make Assumption 2?

Our main result in this subsection shows that, somewhat surprisingly, Assumption 1 effectively yields Assumption 2 *without loss of generality*. While a GAIT may not be perfect, there is a strongly equivalent "alternate reality" in which it is.

Theorem 4.3. *For every* $\mathbf{d} \in \mathbb{R}^K_{\geq 0}$*, there is some* $\tilde{\mathbf{d}} \in \mathbb{R}^K_{\geq 0}$ *such that* $\tilde{\mathbf{d}}$ *is decreasing in* k*, and for all* $n \geq 2$*,* $s \in \mathcal{S}$*,*

1. $\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s) = \mathbf{p}_{\text{EQ}}(n, \tilde{\mathbf{d}}, s)$*, and per-player utilities are equal.*

[9] Allowing for asymmetric optimal strategies can sometimes strictly increase social welfare; we ignore these for now to provide a meaningful comparison with symmetric equilibria.

2. $\mathbf{p}_{\text{OPT}}(n, \mathbf{d}, s) = \mathbf{p}_{\text{OPT}}(n, \tilde{\mathbf{d}}, s)$*, and per-player utilities are equal.*

By Lemmas 4.1 and 4.2, $\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s)$ *and* $\mathbf{p}_{\text{OPT}}(n, \mathbf{d}, s)$ *are both unique.*

Theorem 4.3 tells us that even when $\mathbf{d}$ is not decreasing, under Assumption 1, symmetric equilibrium and optimal strategies are exactly what they would be in an alternate game given by $\tilde{\mathbf{d}}$. Because they are unique, we can restrict ourselves without loss of generality to the decreasing $\mathbf{d}$ case for the remainder of this section. In other words, by making Assumption 1, we can also make Assumption 2 in our analysis, allowing us to apply Lemmas 4.1 and 4.2.

We briefly provide some intuition for what the equivalent game $\tilde{\mathbf{d}}$ looks like. Suppose $\mathbf{d} = [1;\ 0;\ 1]$. Ideally, a player could choose a distribution like $\mathbf{p}' = [1/2;\ 0;\ 1/2]$, but of course, $\mathbf{p}' \notin \Delta(\pi)$. Under Assumption 1, $\mathbf{p}$ must respect the constraint $\mathbf{p}_2 \geq \mathbf{p}_3$. In this example, it is never optimal or rational to set $\mathbf{p}_2 > \mathbf{p}_3$, since a player's utility strictly increases if they transfer mass from $\mathbf{p}_2$ to $\mathbf{p}_3$ or $\mathbf{p}_1$. Thus, for any equilibrium or socially optimal strategy, $\mathbf{p}_2 = \mathbf{p}_3$.

The final step of the proof is to observe that if $\mathbf{p}_2 = \mathbf{p}_3$, then $\mathbf{d} = [1;\ 0;\ 1]$ is effectively the same as $\tilde{\mathbf{d}} = [1;\ 1/2;\ 1/2]$. More generally, we can replace sequences of (weakly) increasing values of $\mathbf{d}$ by their means, since $\mathbf{p}_{\text{EQ}}$ and $\mathbf{p}_{\text{OPT}}$ will assign equal probabilities to all of those elements anyways. Iteratively applying the argument yields the desired result. This procedure is often known as the "Pool Adjacent Violators" algorithm [7] and is commonly used for isotonic regression [5, 15, 99].

A simple example. To build intuition, consider an example with $n = 2$ players and $K = 2$ types with values $\mathbf{d} = [3;\ 2]$. In the absence of competition, each player would choose the distribution $\mathbf{p} = [1;\ 0]$, since type 1 is higher-quality than type 2. But with the score function $s_1(x) = x$, these strategies would yield utility $3/2$ per player. They would be better off diversifying: there is unmet demand $\mathbf{d}_2 = 2$, and by shifting probability mass from $\mathbf{p}_1$ to $\mathbf{p}_2$, players could increase their utilities.

In fact, the unique symmetric pure Nash equilibrium for this example is $\mathbf{p} = [4/5;\ 1/5]$: by Lemma 4.1, the symmetric equilibrium satisfies

$$\begin{aligned}
\mathbf{d}_1 \mathbb{E}\left[\frac{1}{s(1 + X(n-1, \mathbf{p}_1))}\right] &= \mathbf{d}_2 \mathbb{E}\left[\frac{1}{s(1 + X(n-1, \mathbf{p}_2))}\right] \\
3\left(1 - \mathbf{p}_1 + \frac{\mathbf{p}_1}{2}\right) &= 2\left(1 - \mathbf{p}_2 + \frac{\mathbf{p}_2}{2}\right) \\
\mathbf{p}_1 &= \frac{4}{5}.
\end{aligned}$$

Competition between 2 players has diversified equilibrium behavior. And with more competition (as n increases), we might expect even more diverse behavior. On the other hand, the symmetric socially optimal strategy is $\mathbf{p}^* = [3/5;\ 2/5]$, which by Lemma 4.2 satisfies

$$\begin{aligned}
\mathbf{d}_1(1 - \mathbf{p}_1^*)^{n-1} &= \mathbf{d}_2(1 - \mathbf{p}_2^*)^{n-1} \\
3(1 - \mathbf{p}_1^*) &= 2(1 - \mathbf{p}_2^*) \\
\mathbf{p}_1^* &= \frac{3}{5}.
\end{aligned}$$

Thus, while competition induces diversity, socially optimal behavior is even more diverse.

4.2 Diversity of equilibria

In isolation, we should not expect diverse behavior. A single person playing Scattergories by themselves has no need to choose obscure (and potentially incorrect) answers; they simply need a single valid answer. But with more competition, a player must sample from a more diverse distribution in order to avoid collisions with others. Our main results in this section formalize this intuition. We show in Theorems 4.5 and 4.6 that more competitive environments do indeed induce more diverse behavior. On the other hand, we show in Theorem 4.7 that the equilibrium distribution $\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s)$ is less diverse than the optimal distribution $\mathbf{p}_{\text{OPT}}(n, \mathbf{d}, s)$. We then discuss limiting behavior of competition.

First, we must define "diversity" more precisely. All of our results will use a particularly strong notion of the relative diversity of distributions known as *majorization*:

Definition 4.4 (Majorization)**.** For distributions $\mathbf{p}, \mathbf{q} \in \Delta(\pi)$, $\mathbf{p}$ majorizes $\mathbf{q}$ if

$$\sum_{\ell=1}^{k} \mathbf{p}_\ell \geq \sum_{\ell=1}^{k} \mathbf{q}_\ell. \qquad (\forall k \in [K])$$

We write this as $\mathbf{p} \succ \mathbf{q}$.

Majorization provides a useful notion of the relative diversity of discrete distributions. By definition, if $\mathbf{p} \succ \mathbf{q}$, then $\phi(\mathbf{p}) \leq \phi(\mathbf{q})$ for any Schur-concave function ϕ [e.g., 68]. (And similarly, $\mathbf{p} \succ \mathbf{q}$ implies $\phi(\mathbf{p}) \geq \phi(\mathbf{q})$ for any Schur-convex function ϕ.) Standard measures of diversity, such as Shannon [85], Rényi [77], and Tsallis [97] entropies, are Schur-concave. As a result, $\mathbf{p} \succ \mathbf{q}$ implies $\mathbf{p}$ has lower entropy than $\mathbf{q}$. Similarly, measures of inequality like the Gini coefficient [32] are Schur-convex, so we can also interpret $\mathbf{p} \succ \mathbf{q}$ to say that $\mathbf{p}$ is more "unequal" than $\mathbf{q}$. We will present our results in terms of majorization, and intuitively, we can understand "$\mathbf{p}$ majorizes $\mathbf{q}$" ($\mathbf{p} \succ \mathbf{q}$) as "$\mathbf{p}$ is less diverse than $\mathbf{q}$." We begin by showing that increasing the number of players n increases equilibrium diversity. All proofs in this section are deferred to Appendix A.2.

Theorem 4.5 (More players $\Longrightarrow$ more diverse equilibria)**.** *For all $n \geq 2$, $\mathbf{d} \in \mathbb{R}^K_{\geq 0}$, $s \in \mathcal{S}$,*

$$\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s) \succ \mathbf{p}_{\text{EQ}}(n+1, \mathbf{d}, s).$$

We provide a sketch of the proof here. Consider the following thought experiment: what if, in the $n+1$ player game $\mathscr{G}(n+1, \mathbf{d}, s)$, all players played $\mathbf{p} = \mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s)$? We show that if they did, per-type expected utility $\mathbf{d}_k u(n, \mathbf{p}_k)$ would be *increasing* in k (for k such that $\mathbf{p}_k > 0$): playing the n-player equilibrium strategy in the $n+1$-player game would over-exploit the types with high $\mathbf{d}_k$ and under-exploit those with lower $\mathbf{d}_k$.

By Lemma 4.1, under the $n+1$-player equilibrium $\mathbf{p}_{\text{EQ}}(n+1, \mathbf{d}, s)$, types should have equal conditional utility. So in order to transform $\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s)$ into $\mathbf{p}_{\text{EQ}}(n+1, \mathbf{d}, s)$, we have to decrease probabilities on the high-$\mathbf{d}_k$ types and raise them on the low-$\mathbf{d}_k$ types. This seems like it should lead to a more diverse distribution, since we are lowering the larger probabilities and raising the smaller probabilities, and indeed, we show that this implies $\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s) \succ \mathbf{p}_{\text{EQ}}(n+1, \mathbf{d}, s)$.[10]

[10] Intuitively, we might think the same holds for $\mathbf{p}_{\text{OPT}}$; Figure 1 shows that this is not the case.

Theorem 4.5 tells us that increased competition incentivizes more diverse behavior. We show that a similar result holds when increased competition takes the form of a stronger score function s. In particular, we consider score functions of the form $s_\gamma(x) = x^\gamma$.

Theorem 4.6 (Greater externalities $\implies$ more diverse equilibria)**.** *For all $n \geq 2$, $\mathbf{d} \in \mathbb{R}^K_{\geq 0}$, $\gamma_0 \leq \gamma_1$,*

$$\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s_{\gamma_0}) \succ \mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s_{\gamma_1}).$$

The proof idea behind Theorem 4.6 is somewhat similar to that of Theorem 4.5: we ask what would happen if players in $\mathscr{G}(n, \mathbf{d}, s_{\gamma_1})$ were to play the strategies $\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s_{\gamma_0})$, finding that they would again place too much probability mass on types with high values of $\mathbf{d}_k$.

Taken together, Theorems 4.5 and 4.6 provide some intuition on when we should expect to see more diverse vs. more homogeneous behavior from producers using GAITs. In settings where competitors do not impose strong negative externalities on one another, behavior should be relatively homogeneous. Suppose, for example, teachers use a GAIT to generate problem sets. We might expect to see homogeneity across teachers if the value of a homework problem is unaffected by other teachers' behavior. But if some teachers begin publishing solution manuals to their problem sets online, a student may be able to use another teacher's solutions for their homework. Similar assignments now impose negative externalities (from the teachers' perspective). As a result, at equilibrium, teachers should diversify their use of GAITs to avoid generating similar problem sets.

Although competition induces diversity, equilibria are still more homogeneous than would be socially optimal:

Theorem 4.7 (Optimal strategies are more diverse than equilibria)**.** *For all $n \geq 2$, $\mathbf{d} \in \mathbb{R}^K_{\geq 0}$, $s \in \mathcal{S}$,*

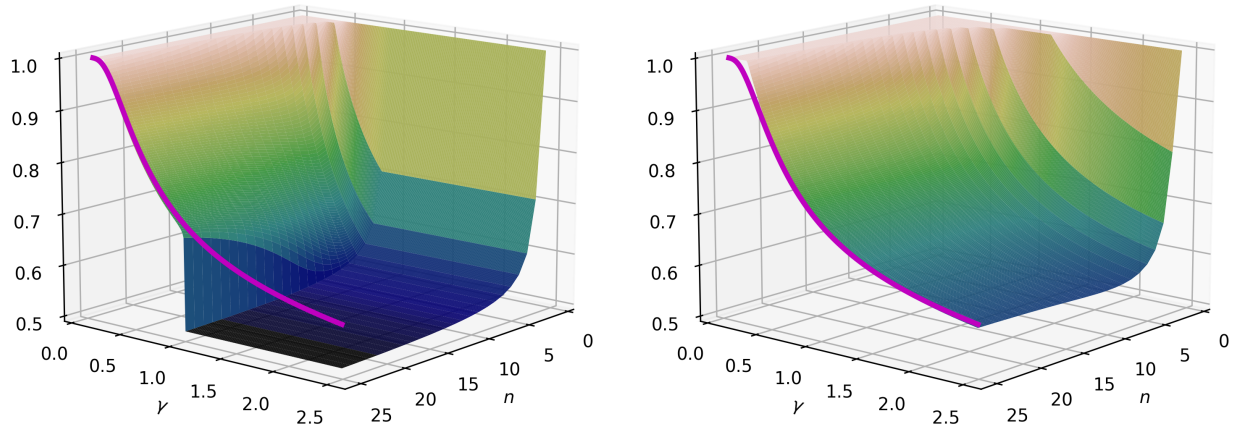
$$\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s) \succ \mathbf{p}_{\text{OPT}}(n, \mathbf{d}, s).$$

Again, the proof considers the hypothetical in which players play $\mathbf{p}_{\text{OPT}}$, showing that they could increase their own utility (at the expense of social welfare) by placing more probability mass on high-$\mathbf{d}_k$ types. To summarize, we have shown so far that competition induces diversity (Theorems 4.5 and 4.6), but equilibria are less diverse than would be socially optimal (Theorem 4.7).

Competition in the limit. As a brief aside, we comment on the effects of competition in the limit. For the score functions s_γ, we consider how $\mathbf{p}_{\text{EQ}}$ and $\mathbf{p}_{\text{OPT}}$ relate as n grows, finding a surprising discontinuity. For $\gamma < 1$, $\mathbf{p}_{\text{EQ}}$ and $\mathbf{p}_{\text{OPT}}$ approach the same distribution $\mathbf{p}^\dagger(\gamma)$, which is proportional to $\mathbf{d}_k^{1/\gamma}$. But for $\gamma \geq 1$, they diverge: $\mathbf{p}_{\text{EQ}}$ still approaches $\mathbf{p}^\dagger$, but $\mathbf{p}_{\text{OPT}}$ approaches the uniform distribution. Intuitively, we can explain this qualitative shift as follows. When $\gamma < 1$, there is no limit to social welfare. More production of type k results in greater welfare, though with diminishing returns. But for $\gamma \geq 1$, social welfare is bounded by $\sum_k \mathbf{d}_k$, and maximizing social welfare amounts to minimizing the chance that any type fails to get produced. Note that this discontinuity is not simply an artifact of our two definitions of social welfare for $\mathcal{S}_\nearrow$ and $\mathcal{S}_\searrow$, since it arises even if we restrict ourselves to $\mathcal{S}_\nearrow$ (i.e., for $\gamma < 1$ vs. $\gamma = 1$). We state the result formally below. See Figure 1 for further intuition.

Lemma 4.8 (Strategies as $n \to \infty$)**.** *Assume without loss of generality that* $\mathbf{d} \in \mathbb{R}^K_{\geq 0}$ *is weakly decreasing. Let* $\mathbf{p}^{(n)} = \mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s_\gamma)$*, and let* $\mathbf{q}^{(n)} = \mathbf{p}_{\text{OPT}}(n, \mathbf{d}, s_\gamma)$*. Let* $\mathbf{p}^\dagger(\gamma)_k \triangleq \mathbf{d}_k^{1/\gamma} / \sum_{k' \in [K]} \mathbf{d}_{k'}^{1/\gamma}$*. Then,*

- $\left\{\mathbf{p}^{(n)}\right\}_{n\to\infty}$ *approaches* $\mathbf{p}^\dagger(\gamma)$*.*
- *If* $\gamma < 1$*,* $\left\{\mathbf{q}^{(n)}\right\}_{n\to\infty}$ *approaches* $\mathbf{p}^\dagger(\gamma)$*.*
- *If* $\gamma \geq 1$ *(or more generally, for any* $s \in \mathcal{S}_\searrow$*),* $\left\{\mathbf{q}^{(n)}\right\}_{n\to\infty}$ *approaches the uniform distribution.*



(a) $\mathbf{p}_{\text{OPT}}(n, \mathbf{d}, s_\gamma)_1$ for $\mathbf{d} = [5;\ 2]$.

(b) $\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s_\gamma)_1$ for $\mathbf{d} = [5;\ 2]$.

Figure 1: For further intuition on Lemma 4.8, we show how $\mathbf{p}_{\text{OPT}}(n, \mathbf{d}, s_\gamma)_1$ and $\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s_\gamma)_1$ respectively vary with both n and γ for the instance given by $\mathbf{d} = [5;\ 2]$. For $\gamma < 1$, both $\mathbf{p}_{\text{OPT}}$ and $\mathbf{p}_{\text{EQ}}$ converge to $\mathbf{p}^\dagger(\gamma)$, shown in magenta. For $\gamma \geq 1$, $\mathbf{p}_{\text{OPT}}$ converges to the uniform distribution, with a discontinuity at $\gamma = 1$. Further, while $\mathbf{p}_{\text{EQ}}$ is monotone in both n and γ (Theorems 4.5 and 4.6), $\mathbf{p}_{\text{OPT}}$ is not. Increasing n can lead to a *less* diverse symmetric socially optimal strategy.

While the proof is quite involved, the idea is simple: for sufficiently large n, $X(n, p)$ concentrates around np, so $u_{s_\gamma}(n, p) \approx (np)^{-\gamma}$. Another interesting limiting case is where $\gamma = \infty$, i.e., players only derive utility when their sampled types are unique (which is how Scattergories is intended to be played). In this case, equilibria coincide with optimal strategies under any $s \in \mathcal{S}_\searrow$.

Lemma 4.9 (Equivalence when $\gamma = \infty$)**.** *For all* $n \geq 2$*,* $\mathbf{d} \in \mathbb{R}^K_{\geq 0}$*,* $s \in \mathcal{S}_\searrow$*,*

$$\mathbf{p}_{\text{OPT}}(n, \mathbf{d}, s) = \mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s_\infty).$$

It turns out that $\gamma = \infty$ also yields a bound on the diversity of equilibria. No $s \in \mathcal{S}$ yields an equilibrium that is more diverse than the one induced by s_∞.

Lemma 4.10 ($\gamma = \infty$ yields the most diverse equilibria)**.** *For all $n \geq 2$, $\mathbf{d} \in \mathbb{R}^K_{\geq 0}$, $s \in \mathcal{S}$,*

$$\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s) \succ \mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s_\infty).$$

4.3 Price of anarchy

Optimal behavior is more diverse than equilibrium behavior, and while they may coincide in the limit, equilibria will in general be suboptimal. We conclude this section by showing that they cannot be *too* suboptimal: the price of anarchy [55, 80] is at most 2. In fact, this is true even if we allow for asymmetric behavior. Let $\mathbf{P} \in \Delta(\pi)^n$ be a $K \times n$ matrix in which each column i is player i's strategy. Social welfare is then defined as before, where each player independently samples from potentially different distributions. We show in Appendix A.3 that $\mathscr{G}$ is a *valid utility game* [100], which directly yields the following result. We also give a family of instances for which it is tight.

Theorem 4.11 (Price of anarchy)**.** *The price of anarchy for $\mathscr{G}$ is 2. Formally, let $\mathbf{P}$ be any equilibrium of $\mathscr{G}(n, \mathbf{d}, s)$, and let $\mathbf{P}^*$ be any socially optimal strategy matrix. Then,*

$$W(\mathbf{P}^*) \leq 2W(\mathbf{P}).$$

There exists a family of instances $\{\mathscr{G}(n, \mathbf{d}^{(n)}, s)\}_{n\to\infty}$ for which this is asymptotically tight.

Thus, while selfish behavior leads to lower diversity, its negative consequences on social welfare are bounded. Having analyzed the setting in which all players use a single GAIT, we next generalize to the case where players can choose a GAIT.

5 Competition across Tools

In this section, players strategically choose a GAIT, represented by a permutation $\pi^{(j)}$ over $[K]$, from a set of choices $\Pi = \{\pi^{(1)}, \pi^{(2)}, \ldots, \pi^{(J)}\}$. Subject to their choice of GAIT, each player chooses a distribution $\mathbf{p} \in \Delta(\pi^{(j)})$ that obeys the ordering given by $\pi^{(j)}$. This strictly generalizes the setting described in Section 4, in which Π was a singleton set. Our results here will not be as general as before; indeed, symmetric pure strategy equilibria need not exist in this setting. Our goal will instead be to provide illustrative examples of how competition across tools might unfold, which will ground our empirical results in Section 6.2.

Analyzing market share. Our primary focus in this section will be on *market share at equilibrium*. How many players choose to use each GAIT? We will restrict our attention to pure partially symmetric equilibria, in which all players who choose $\pi^{(j)}$ also choose the same distribution $\mathbf{p}^{(j)}$. Given a set of GAITs Π and values $\mathbf{d}$, a pure partially symmetric equilibrium can be specified by $\{m_j\}_{j=1}^J$, the number of players who use GAIT j, and $\{\mathbf{p}^{(j)}\}_{j=1}^J$, their respective strategies. Let $\mathcal{E}(n, \Pi, \mathbf{d})$ be set of all such pure partially symmetric equilibria. With this, we can define market share.

Definition 5.1. Given an instance $(n, \Pi, \mathbf{d})$, the market share of GAIT j is

$$\mathcal{M}_j(n, \Pi, \mathbf{d}) \triangleq \{m : \exists \text{ an equilibrium in } \mathcal{E}(n, \Pi, \mathbf{d}) \text{ such that } m_j = m\}.$$

In other words, $\mathcal{M}_j$ is the set of possible market shares for tool j. Define $\overline{\mathcal{M}}_j(n, \Pi, \mathbf{d}) \triangleq \max \mathcal{M}_j(n, \Pi, \mathbf{d})$ and $\underline{\mathcal{M}}_j(n, \Pi, \mathbf{d}) \triangleq \min \mathcal{M}_j(n, \Pi, \mathbf{d})$ to be the maximal and minimal (respectively) market shares that tool j can have at any equilibrium in $\mathcal{E}$.

Intuitively, we might expect market share to be related to the quality of a GAIT: more players should choose a "better" tool over a worse one. Because tools may take a range of market shares across different equilibria, we need to be careful in how we formalize this. Let $\Pi = \{\pi^{(1)}, \pi^{(2)}\}$. Tool 1 is "better" than tool 2 in isolation if $\underline{\mathcal{M}}_1(1, \Pi, \mathbf{d}) = 1$ (and therefore $\overline{\mathcal{M}}_2(1, \Pi, \mathbf{d}) = 0$). This means that a single player facing no competition strictly prefers tool 1 to tool 2. Under this condition, we might expect that with n players, GAIT 1's market share would be at least as high as 2's. Conservatively, we can express this as $\overline{\mathcal{M}}_1(n, \Pi, \mathbf{d}) \geq \underline{\mathcal{M}}_2(n, \Pi, \mathbf{d})$, meaning that the "best" equilibrium for tool 1 yields market share at least as large as the "worst" equilibrium for tool 2.

Our first result in this setting shows that this is not the case: it is possible for a GAIT that appears better in isolation to get strictly less market share than a worse one as competition (i.e., the number of players n) increases.

Lemma 5.2 (The dominant GAIT depends on the strength of competition)**.** *There exists* $(n, \Pi = \{\pi^{(1)}, \pi^{(2)}\}, \mathbf{d})$ *such that* $\underline{\mathcal{M}}_1(1, \Pi, \mathbf{d}) = 1$ *and* $\overline{\mathcal{M}}_1(n, \Pi, \mathbf{d}) < \underline{\mathcal{M}}_2(n, \Pi, \mathbf{d})$.

All proofs for this section can be found in Appendix B. Lemma 5.2 shows that GAIT quality is not one-dimensional. A benchmark might suggest that $\pi^{(1)}$ is better than $\pi^{(2)}$, but in a competitive marketplace, $\pi^{(2)}$ might turn out to have an edge. We next show conditions under which a tool gets completely locked out of the market.

Lemma 5.3 (Strict domination is possible)**.** *Suppose* $s(x) < \infty$ *for* $x \in \mathbb{N}_{>0}$*, and suppose* $\exists\ \hat{k}$ *such that* $\mathbf{d}_{\hat{k}} = 0$*. For any* $\pi^{(1)}$ *such that* $\hat{k}$ *is not ranked first by* $\pi^{(1)}$*, there exists* $\pi^{(2)}$ *such that for all* $\Pi \supseteq \{\pi^{(1)}, \pi^{(2)}\}$*,* $\overline{\mathcal{M}}_2(n, \Pi, \mathbf{d}) = 0$ *for all* n*.*

Lemma 5.3 is a bit dense, but the intuition is simple. As long as there's a 0-value type that $\pi^{(1)}$ doesn't rank first, we can find a strictly worse GAIT $\pi^{(2)}$ that never gets any market share. The construction is fairly simple: we obtain $\pi^{(2)}$ from $\pi^{(1)}$ by simply taking the 0-value element $\hat{k}$ and putting it at the first position of $\pi^{(2)}$. Our final result states that the perfect GAIT weakly captures the entire market against *all possible* competition. On the other hand, for any imperfect tool $\pi^{(1)}$, there exists some other tool $\pi^{(2)}$ that exploits $\pi^{(1)}$'s imperfection.

Lemma 5.4 (Only the perfect tool weakly dominates)**.** *If* $\pi^{(1)}$ *is a perfect ordering according to* $\mathbf{d}$*, i.e.,* $\mathbf{d}_{\pi_1^{(1)}} \geq \mathbf{d}_{\pi_2^{(1)}} \geq \cdots \geq \mathbf{d}_{\pi_K^{(1)}}$*, then* $\overline{\mathcal{M}}_1(n, \Pi, \mathbf{d}) = n$ *for all* $\Pi \ni \pi^{(1)}$*. Otherwise, there exist* $\pi^{(2)}$ *and* n_0 *such that for all* $n \geq n_0$*,* $\overline{\mathcal{M}}_1(n, \{\pi^{(1)}, \pi^{(2)}\}, \mathbf{d}) < n$*.*

As a result, a strategic tool developer need not attempt to create a perfectly aligned tool; instead, they might seek "competitive alignment" by filling a niche left open by existing tools. A tool might, for example, prioritize a particular visual or linguistic style even if it isn't what the majority of users want simply because existing tools fail to provide it. This concludes our theoretical exploration. Next, we validate our results empirically by using language models to play Scattergories.

6 Case Study: Scattergories

Our empirical setup is simple: n players each use a GAIT to generate a single response to a Scattergories instance given by a (letter, category) pair. Each player can strategically choose a temperature τ at which to sample. Formally, the GAIT outputs a logit ℓ_t for each possible token t in its vocabulary $\mathcal{V}$, and at temperature τ, we sample token t with probability proportional to $\exp(\ell_t/\tau)$. Higher temperatures induce more diverse distributions at the individual token level. In the special case where $\tau = 0$, the GAIT deterministically produces the token with the highest logit, $\arg\max_{t\in\mathcal{V}} \ell_t$.

Because Scattergories requires complete words or phrases, answers typically consist of multiple tokens. We generate tokens sequentially until we encounter a `STOP` token. (Most GAITs have a built in end-of-message token used to signal that generation is complete.) We additionally treat all tokens containing non-alphanumeric characters as `STOP` tokens. In all of our experiments, we use score functions of the form $s_\gamma(x) = x^\gamma$. Although the original rules of Scattergories call for $\gamma = \infty$, we use other values because equilibrium and optimal strategies coincide under s_∞ (see Lemma 4.9).

Data and language models. The categories we use for our games come from a free online Scattergories game.[11] We randomly sample 25 instances, each of which is a (letter, category) pair, and we average across them in all of our results. A full list of our sample can be found in Table 2. Our experiments compare six GAITs, each of which is a relatively small open-source language model: gemma-2-2b-it (gemma2), Llama-3.1-8B-Instruct (llama3.1), Llama-3.2-1B-Instruct (llama3.2), Mistral-7B-Instruct-v0.3 (mistral), Nemotron-Mini-4B-Instruct (nemotron), and Phi-3.5-mini-instruct (phi3.5). In addition to these, we use a seventh language model, Qwen2.5-7B-Instruct (qwen2.5), to validate answers. See Table 1 for more details.

While relying on language models to evaluate one another comes with several downsides [19, 102, 110], our setting has a number of mitigating factors that make it a more reasonable choice. Even with human players, Scattergories works via self-evaluation. Evaluation is fairly simple: each evaluation is a binary response to a short answer. We do not seek to make strong claims about the correctness of the evaluations made by qwen2.5, although manual spot-checking reveals that its judgements are reasonable. Instead, we view our results as confirmation that there exists a reasonable setting in which our theoretical results hold. We validate answers using a separate language model to minimize the risk of self-preferencing [54, 59, 70]. One downside we do not address in this work is the fact that we treat restatements of the same answer (e.g., "swordfish" and "sword fish") as distinct. For both answer generation and validation, we include two in-context examples to improve performance (see Figure 13 for the full prompts).

Even in this limited setting where answers are short sequences of tokens, computing the distribution of answers for a given (language model, temperature, instance) combination is still non-trivial. Ideally, we would simply write down a GAIT's output distribution as a function of temperature τ. Indeed, if we only considered single-token answers, this would be straightforward: each token t is sampled with probability $\propto \exp(\ell_t/\tau)$. If a GAIT's

[11] `https://swellgarfo.com/scattergories/`

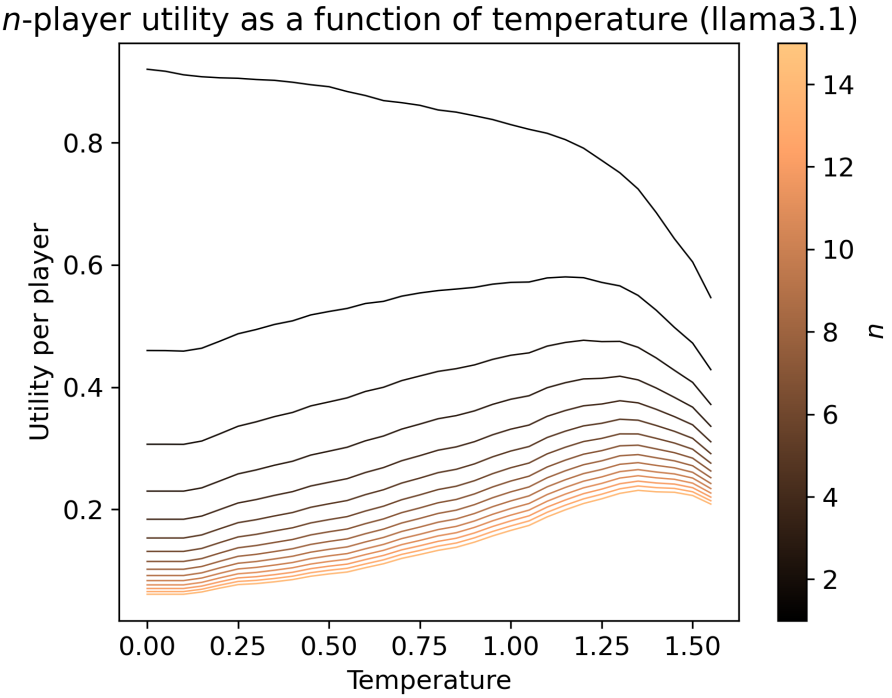


Figure 2: Per-player utility as a function of τ for different values of n. The highest line corresponds to $n = 1$, and the lowest corresponds to $n = 15$. $\gamma = 1.0$.

vocabulary contains $|\mathcal{V}|$ tokens, we need $\Theta(|\mathcal{V}|)$ space just to write down. For multi-token answers, this becomes prohibitive, requiring $\Theta(|\mathcal{V}|^d)$ time and space, where d is the maximum answer length.[12]

Instead, we take a sampling-based approach: we can simulate games of Scattergories by sampling completions from a GAIT (at each temperature τ) for each player. While this approach is more tractable, it comes with its own disadvantages. We generate samples independently for each τ under consideration, meaning we must take temperatures at discrete points (we choose intervals of 0.05). With enough samples (we choose 2,000 for each (language model, temperature, instance) combination), we can get fairly tight estimates of players' utility. See Appendix C.2 for details on how we use samples optimally and bound error.

6.1 Competition under a single tool

We begin by analyzing the setting in which all players share the same GAIT and can only vary the temperature τ. Figure 2 shows utility as a function of τ in the n-player game with the score function $s_1(x) = x$ when all players use llama3.1. Observe that utility is approximately quasiconcave as a function of τ: it increases with τ up to a certain point, and then decreases monotonically. Moreover, the peak occurs at larger values of τ as competition gets stronger with larger values of n. Lower temperatures are effective with mild competition

[12]For context, nemotron has a valid vocabulary size $|\mathcal{V}| = 169{,}159$, and we allow responses up to 6 tokens.

but yield low utilities when competition increases. We provide analogous figures for the other five GAITs in Figure 7.

Diversity and utility of equilibria. Because Assumption 1 is not strictly true (we analyze extent to which it is violated later in this subsection), pure symmetric equilibria do not necessarily exist here. Nevertheless, let $\tau_{\text{EQ}}(n, \gamma)$ be the symmetric equilibrium temperature if it exists for a single GAIT given n players and score function s_γ.[13] From our results in Section 4, we expect τ_{EQ} to increase in both n and γ (Theorems 4.5 and 4.6). We'd also expect the socially optimal temperature $\tau_{\text{OPT}}(n, \gamma)$ to be larger than the corresponding equilibrium temperature $\tau_{\text{EQ}}(n, \gamma)$ (Theorem 4.7).

Figure 3 supports our theoretical results in Theorem 4.5, that stronger competition induces more diversity (larger values of τ_{EQ}). Analogous results hold as we vary γ (Theorem 4.6) in Figure 8. Equilibria reflect a balance between the benefits of diversity and the costs of invalid answers that occur at high temperatures. Symmetric pure equilibria do not always exist (e.g., phi3.5 with $n = 4$). When they do exist, Figures 3 and 8 show that $\tau_{\text{EQ}}(n, \gamma) \leq \tau_{\text{OPT}}(n, \gamma)$, as Theorem 4.7 suggests.

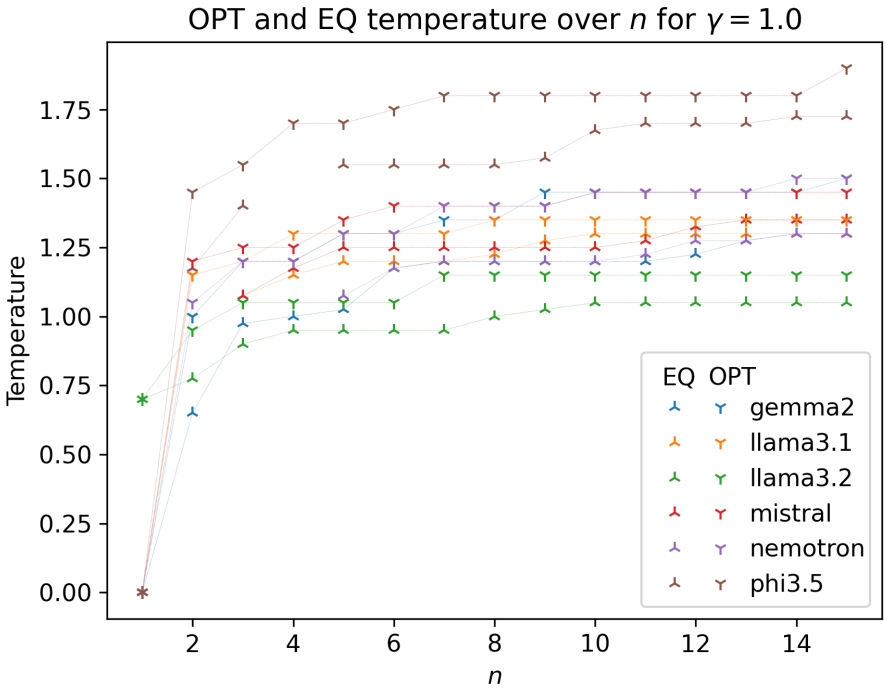


Figure 3: Socially optimal and equilibrium temperatures for each language model as a function of n. Observe that (1) all curves are increasing in n, and (2) equilibrium temperatures are lower than their socially optimal counterparts. $\gamma = 1.0$

Next, we consider players' utilities in Figure 4. Equilibrium utilities are, of course, lower than would be socially optimal, but the gap between them is not substantial; it is well within the factor of two predicted by our price of anarchy bound in Theorem 4.11.

[13] In principle, there could be multiple pure strategy equilibria. Empirically, this is not the case here.

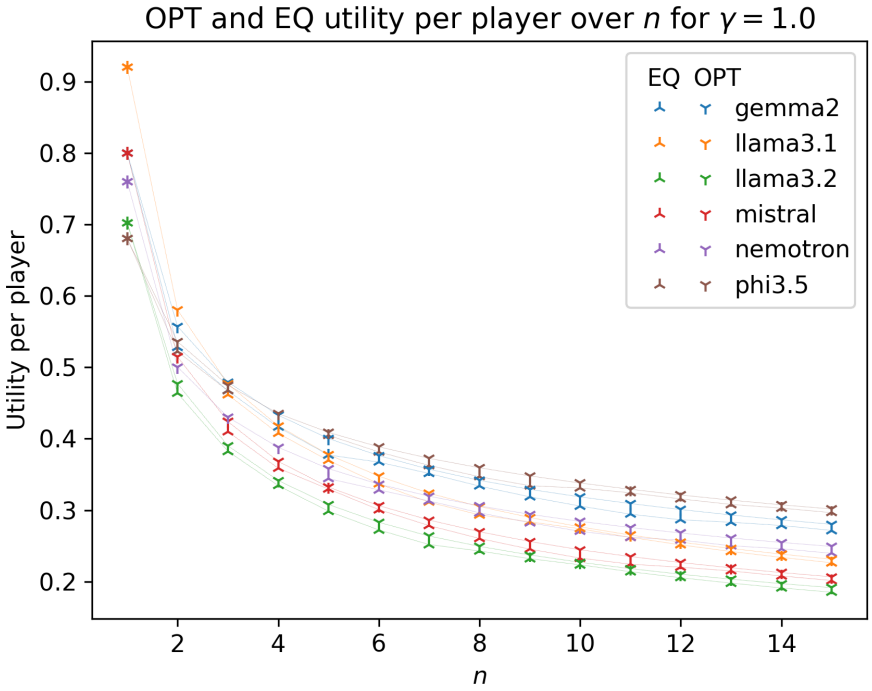
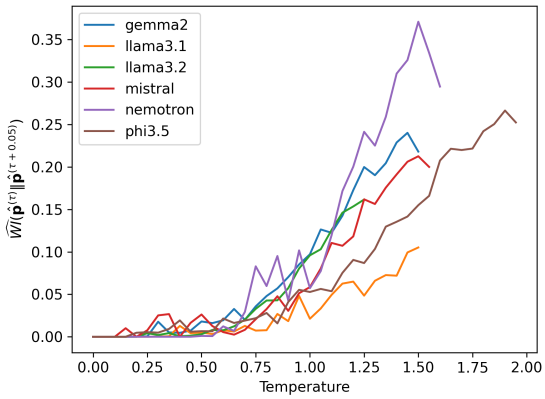


Figure 4: Socially optimal and equilibrium welfare (per player) for each language model as a function of n. Observe that (1) welfare is decreasing in n, and (2) equilibrium welfare is not too far from optimal. phi3.5 and gemma2 perform better in the presence of stronger competition. $\gamma = 1.0$.

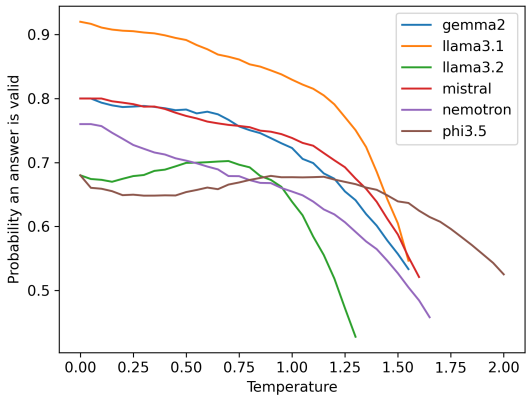
Perhaps more interestingly, we find strong evidence that the relative quality of GAITs depends on the strength of competition. With $n = 1$, llama3.1 yields the best performance out of all of our GAITs, providing a valid answer 92% of the time at a temperature of $\tau_{\text{OPT}}(1, \cdot) = \tau_{\text{EQ}}(1, \cdot) = 0.0$.[14] In contrast phi3.5 only yields a valid answer 68% of the time, also at an optimal temperature of 0.0. According to this naive benchmark, it might seem that llama3.1 is the "better" of the two GAITs in the Scattergories context. However, this comparison masks a surprising fact: under competition, the relative performance between these GAITs flips. phi3.5 outperforms llama3.1 when there are either more players (Figure 4) or stronger penalties for congestion (Figure 10). See Appendix C.1 for analogous plots varying γ.

Intuitively, llama3.1 is a "one-trick pony" in this setting. Its best answer is usually valid, but diversifying its distribution fails to place much probability mass on distinct, valid answers. phi3.5, gemma2, and nemotron, on the other hand, make more mistakes when producing a single answer, but are able to draw from a large set of valid answers at higher temperatures. We caution that our findings on these particular GAITs quality are specific to our setting and should not be used to draw more general conclusions about them without more comprehensive testing.

[14] When $n = 1$, the score function s does not matter.



(a) $\widehat{\mathsf{WI}}(\hat{\mathbf{p}}^{(\tau)} \,\|\, \hat{\mathbf{p}}^{(\tau+0.05)})$ as a function of τ. Larger values correspond to larger violations of Assumption 1. $\widehat{\mathsf{WI}}$ cannot exceed 1.



(b) $\mathbb{E}_{k\sim\mathbf{p}^{(\tau)}}[\mathbf{d}_k]$ for different language models and temperatures. Smaller values correspond to larger violations of Assumption 2.

Figure 5: Measures of the extents to which Assumptions 1 and 2 are violated.

Analyzing Assumptions 1 and 2. Before we move on, we evaluate whether either Assumption 1 or Assumption 2 holds. Strictly speaking, our theoretical results require that players can choose arbitrary distributions. When using temperature scaling, they only have one degree of freedom, so our theoretical results cannot directly apply here regardless. First, we measure the extent to which the order-preserving constraint of Assumption 1 is violated. Let $\pi(\mathbf{p})$ be the ranking over $[K]$ given by a distribution $\mathbf{p}$ in terms of decreasing probability mass, and let $\mathbf{p}^{(\tau)}$ be the GAIT's output distribution at temperature τ. Our measure of weighted inversions is

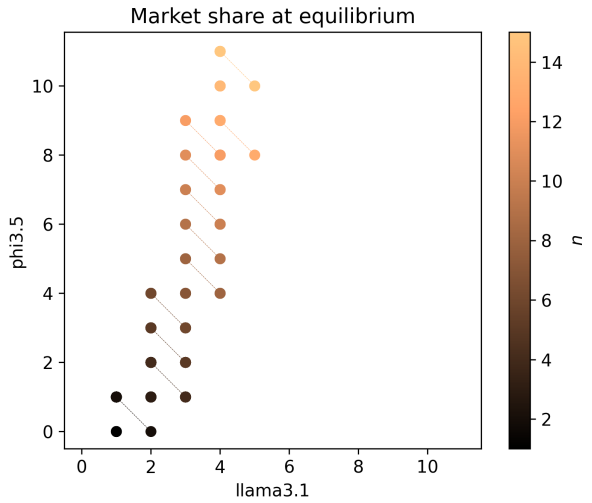
$$\mathsf{WI}(\mathbf{p}^{(\tau)} \,\|\, \mathbf{p}^{(\tau')}) \triangleq \frac{\min_{\mathbf{p}\in\Delta(\pi(\mathbf{p}^{(\tau)}))} d_{\mathrm{TV}}(\mathbf{p}, \mathbf{p}^{(\tau')})}{d_{\mathrm{TV}}(\mathbf{p}^{(\tau)}, \mathbf{p}^{(\tau')})},$$

where d_{TV} is total variation distance. In words, we find the minimal change that we would need to make to $\mathbf{p}^{(\tau')}$ to make it respect the ranking $\pi(\mathbf{p}^{(\tau)})$. We normalize by the distance between $\mathbf{p}^{(\tau)}$ and $\mathbf{p}^{(\tau')}$ so we measure the fraction of the change between $\mathbf{p}^{(\tau)}$ and $\mathbf{p}^{(\tau')}$ that is order-violating. For intuition, if $\mathbf{q}$ is the exact reverse of $\mathbf{p}$, then $\mathsf{WI}(\mathbf{p} \,\|\, \mathbf{q}) = 0.5$. Because we only have samples from each $\mathbf{p}^{(\tau)}$, we cannot directly compute WI. In Appendix C.3, we describe how we compute an approximate version $\widehat{\mathsf{WI}}$. Figure 5a shows $\widehat{\mathsf{WI}}(\hat{\mathbf{p}}^{(\tau)} \,\|\, \hat{\mathbf{p}}^{(\tau+0.05)})$ as a function of τ. For small values of τ, temperature scaling is nearly order-preserving, and it starts to break down as τ increases.

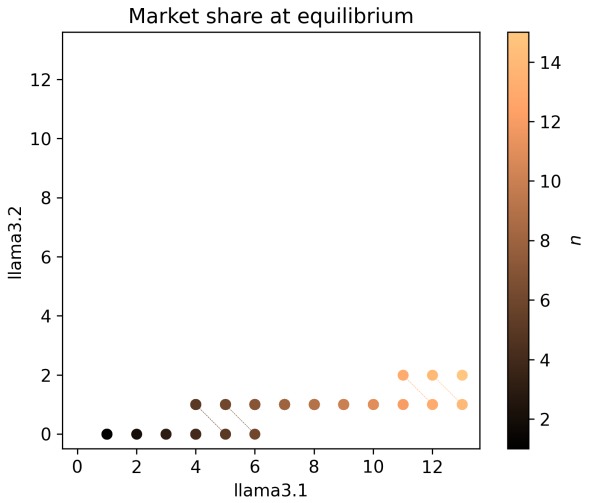
In Theorem 4.3, we found that if Assumption 2 holds, then Assumption 1 is not necessary for our results. In Figure 5b, we plot the (empirical) probability that an answer from a language model is valid. The smaller this probability, the greater the extent to which Assumption 2 is violated.[15]

[15]This relationship relies on the fact that there exist valid answers assigned low or 0 probability by a language model, which is in general true in our setting. See Appendix C.3 for more details.

6.2 Pairwise competition across tools



(a) Market shares in pairwise equilibria between llama3.1 and phi3.5.



(b) Market shares in pairwise equilibria between llama3.1 and llama3.2.

Figure 6: Each dot represents a possible pair (m_1, m_2) of equilibrium market shares for each language model. Dashed lines connect points with the same total number of players n, with a maximum of $n = 15$. $\gamma = 1.0$.

Next, we consider the setting in Section 5, where players strategically choose both a GAIT and a temperature τ. For simplicity, we limit ourselves to pairwise comparisons, and as before, our focus will be on market share at partially symmetric pure strategy equilibria.

Based on Figure 4, we might expect that the market share of each GAIT will shift dramatically as competition grows. For example, given the choice between llama3.1 and phi3.5, a single player ($n = 1$) would clearly prefer llama3.1. But with more players, we could see a mix where some choose llama3.1 and some choose phi3.5, and we might anticipate that for sufficiently large n, a majority of players would choose phi3.5. This is precisely what Figure 6a shows. Each point (m_1, m_2) represents market shares for llama3.1 and phi3.5 respectively at an equilibrium with $n = m_1 + m_2$ players. With $n = 1$ player, llama3.1 is optimal. But for larger n, the clear majority of players prefer phi3.5. As in Lemma 5.2, the "better" language model in isolation cedes market power under competition.

Next, we ask whether one language model "dominates" another in the sense of Lemma 5.3. In our sample, for any pair of language models, there exists sufficiently large n such that both get a nonzero share of the market. For example, as we can see from Figures 4 and 10, llama3.1 yields higher equilibrium utility than llama3.2 at any value of competition, either via n or γ. Despite this, we find that llama3.2 can still get nonzero market share (Figure 6b). This is perhaps even more surprising given that llama3.1 and llama3.2 are from the same family of language models, meaning their output distributions are especially likely to be homogeneous. Our results show that despite this relationship, llama3.2 can still prove useful to a (relatively small) fraction of the market. The same is true across all pairs of language models in our sample (Figure 11). Even without any sort of price differentiation, the market naturally

supports, and indeed encourages, competition. Of course, had we included a sufficiently low-quality language model in our sample, it would likely fail to capture any market share. We also do not find evidence that any of our language models is close enough to the perfect GAIT for Lemma 5.4 to hold.

7 Discussion

Our work studies the interplay between competition and creation through a game-theoretic lens. Using a simple model and empirical setting, we have described a rich set of phenomena that shed light on the diversity and quality of production under generative AI. And despite our model's simplicity, its key predictions—on diversity, quality, and market shares under competition—are borne out in our experiments.

Our findings have important implications for the development, evaluation, and use of generative AI tools. In highly competitive settings, GAITs are particularly useful if they can create high-quality outputs that are distinct from those of other tools on the market. Our model suggests that the value of data to a developer depends on both its quality in an absolute sense (i.e., the improvement it will enable on a benchmark) as well as its value in a competitive setting: data may be especially valuable if it captures signal others lack, increasing the value of exclusive rights to proprietary data. Further, our model provides additional motivation for objectives like alignment [88], which seek to ensure that tools sample from a diverse or representative distribution of outputs.

Existing GAIT evaluation methods typically consist of benchmarking their performance in isolation. Our results indicate that we learn something different when we evaluate tools jointly. Efforts like the ChatBot Arena [20], which allows users to choose between a pair of responses generated by anonymous language models, begin to get at this. It would be interesting to see whether a similar platform with k-way choices instead of pairwise choices would yield a different ranking over language models. We might expect high-quality but similar chatbots to split their demand, allowing a lower-quality but distinctive chatbot to be highly ranked.

Finally, users of GAITs must be increasingly sensitive to their competitors' behavior, particularly as AI adoption grows. Recruiters warn have already begun to warn job applicants that AI-written resumes "look exactly the same" [61]; applicants who want to use AI successfully need to explicitly diversify their use of tools [48]. Possible strategies include altering sampling parameters, varying a prompt, or including personal information to increase the likelihood of getting a unique response.

We see a wide range of potential future directions to build upon our work. From a modeling perspective, future work could incorporate more general notions of similarity and competition than our type-based model. Our model of producer-GAIT collaboration is somewhat limited; suppose, instead, producers sample multiple outputs from a GAIT and strategically choose one. Could a small number of choices [6, 64] dramatically increase welfare and mitigate monoculture? We might also question whether our conception of social welfare is broad enough to reveal the full costs of homogenization. In our model, consumer welfare only appears insofar as it directly impacts producer welfare. But perhaps some of the value of content diversity indirectly accrues to consumers who do not contribute to producer welfare

at all. Somewhat paradoxically, AI developers who scrape publicly available data benefit immensely from its diversity, but because the original creators of that data do not share in those benefits (and indeed, they may not even be aware that their data has been scraped), their behavior will be unaffected by developers' preferences for diversity. Rational economic behavior on the part of data creators [e.g., 90] may result in far more homogenization than would be socially optimal. Accounting for social welfare beyond its direct impact on producer utilities is a promising direction for future work.

On the empirical side, Scattergories is a natural competitive benchmark for language models. It could prove useful to researchers in working towards alignment—better-aligned models should have a competitive advantage here. More broadly, future work could build competitive benchmarks that go beyond the language modality, providing a comprehensive testbed for *competitive alignment*. Our hope is that both our theoretical model and empirical setting provide a starting point for future work to study the rich interaction between competition and diversity in generative AI.

Acknowledgements. We thank David Autor, Kate Donahue, Maryam Farboodi, Nikhil Garg, Amritha Jayanti, Jon Kleinberg, Simon Johnson, Sendhil Mullainathan, Ashesh Rambachan, Dave Rand, and participants of the Fall 2024 Schwarzman College of Computing Special Topics in Computing Workshop and Tuck Operations and Management Science seminar for helpful discussions and feedback.

References

[1] Marah Abdin, Sam Ade Jacobs, Ammar Ahmad Awan, Jyoti Aneja, Ahmed Awadallah, Hany Awadalla, Nguyen Bach, Amit Bahree, Arash Bakhtiari, Harkirat Behl, et al. Phi-3 technical report: A highly capable language model locally on your phone. *arXiv preprint arXiv:2404.14219*, 2024.

[2] Krishna Acharya, Varun Vangala, Jingyan Wang, and Juba Ziani. Producers equilibria and dynamics in engagement-driven recommender systems. *arXiv preprint arXiv:2401.16641*, 2024.

[3] Ravindra K Ahuja and James B Orlin. A fast scaling algorithm for minimizing separable convex functions subject to chain constraints. *Operations Research*, 49(5):784–789, 2001.

[4] Barrett R. Anderson, Jash Hemant Shah, and Max Kreminski. Homogenization effects of large language models on human creative ideation. In *Proceedings of the 16th Conference on Creativity & Cognition*, pages 413–425. ACM, 2024.

[5] Miriam Ayer, H Daniel Brunk, George M Ewing, William T Reid, and Edward Silverman. An empirical distribution function for sampling with incomplete information. *The annals of mathematical statistics*, pages 641–647, 1955.

[6] Yossi Azar, Andrei Z Broder, Anna R Karlin, and Eli Upfal. Balanced allocations. In *Proceedings of the twenty-sixth annual ACM symposium on theory of computing*, pages 593–602, 1994.

[7] Richard E Barlow, D J Bartholomew, J M Bremner, and H Daniel Brunk. *Statistical inference under order restrictions*. John Wiley & Sons, 1972.

[8] Ran Ben Basat, Moshe Tennenholtz, and Oren Kurland. A game theoretic analysis of the adversarial retrieval setting. *Journal of Artificial Intelligence Research*, 60:1127–1164, 2017.

[9] Johannes Gerd Becker and Damian S Damianov. On the existence of symmetric mixed strategy equilibria. *Economics Letters*, 90(1):84–87, 2006.

[10] Omer Ben-Porat and Moshe Tennenholtz. A game-theoretic approach to recommendation systems with strategic content providers. *Advances in Neural Information Processing Systems*, 31, 2018.

[11] Omer Ben-Porat, Gregory Goren, Itay Rosenberg, and Moshe Tennenholtz. From recommendation systems to facility location games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 1772–1779, 2019.

[12] David Blackwell. Conditional expectation and unbiased sequential estimation. *The Annals of Mathematical Statistics*, 18(1):105–110, 1947.

[13] Rishi Bommasani, Kathleen A Creel, Ananya Kumar, Dan Jurafsky, and Percy S Liang. Picking on the same person: Does algorithmic monoculture lead to outcome homogenization? *Advances in Neural Information Processing Systems*, 35:3663–3678, 2022.

[14] Giulia Brancaccio, Myrto Kalouptsidi, Theodore Papageorgiou, and Nicola Rosaia. Search frictions and efficiency in decentralized transport markets. *The Quarterly Journal of Economics*, 138(4):2451–2503, 2023.

[15] Hugh D Brunk. Maximum likelihood estimates of monotone parameters. *The Annals of Mathematical Statistics*, pages 607–616, 1955.

[16] Nicholas Buchholz. Spatial equilibrium, search frictions, and dynamic efficiency in the taxi industry. *The Review of Economic Studies*, 89(2):556–591, 2022.

[17] Daniel Buschek, Lukas Mecke, Florian Lehmann, and Hai Dang. Nine potential pitfalls when designing human-AI co-creative systems. *arXiv preprint arXiv:2104.00358*, 2021.

[18] Alice Cai, Steven R Rick, Jennifer L Heyman, Yanxia Zhang, Alexandre Filipowicz, Matthew Hong, Matt Klenk, and Thomas Malone. Designaid: Using generative AI and semantic diversity for design inspiration. In *Proceedings of The ACM Collective Intelligence Conference*, pages 1–11, 2023.

[19] Cheng-Han Chiang and Hung-yi Lee. Can large language models be an alternative to human evaluations? In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL*, pages 15607–15631. Association for Computational Linguistics, 2023.

[20] Wei-Lin Chiang, Lianmin Zheng, Ying Sheng, Anastasios Nikolas Angelopoulos, Tianle Li, Dacheng Li, Banghua Zhu, Hao Zhang, Michael I. Jordan, Joseph E. Gonzalez, and Ion Stoica. Chatbot Arena: An open platform for evaluating LLMs by human preference. In *Forty-first International Conference on Machine Learning, ICML*, 2024.

[21] Alex Chin and Zhiwei Qin. A unified representation framework for rideshare marketplace equilibrium and efficiency. In *Proceedings of the 31st ACM International Conference on Advances in Geographic Information Systems*, pages 1–11, 2023.

[22] Kathleen Creel and Deborah Hellman. The algorithmic leviathan: Arbitrariness, fairness, and opportunity in algorithmic decision-making systems. *Canadian Journal of Philosophy*, 52(1):26–43, 2022.

[23] Anil R Doshi and Oliver P Hauser. Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances*, 10(28):eadn5290, 2024.

[24] Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Amy Yang, Angela Fan, et al. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*, 2024.

[25] Ziv Epstein, Aaron Hertzmann, Memo Akten, Hany Farid, Jessica Fjeld, Morgan R Frank, Matthew Groh, Laura Herman, Neil Leach, Robert Mahari, Alex Pentland, Olga Russakovsky, Hope Schroeder, and Amy Smith. Art and the science of generative AI. *Science*, 380(6650):1110–1111, 2023.

[26] Seyed A Esmaeili, Kshipra Bhawalkar, Zhe Feng, Di Wang, and Haifeng Xu. How to strategize human content creation in the era of GenAI? *arXiv preprint arXiv:2406.05187*, 2024.

[27] William Feller. *An introduction to probability theory and its applications*, volume 1. New York: John Wiley, 3 edition, 1968.

[28] Nic Fishman and Leif Hancox-Li. Should attention be all we need? The epistemic and ethical implications of unification in machine learning. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, pages 1516–1527, 2022.

[29] Gaëtan Fournier and Marco Scarsini. Location games on networks: Existence and efficiency of equilibria. *Mathematics of Operations Research*, 44(1):212–235, 2019.

[30] Soheil Ghili and Vineet Kumar. Spatial distribution of supply and the role of market thickness: Theory and evidence from ridesharing. In *Proceedings of the 21st ACM Conference on Economics and Computation*, pages 505–506. ACM, 2020.

[31] Arpita Ghosh and Preston McAfee. Incentivizing high-quality user-generated content. In *Proceedings of the 20th international conference on World wide web*, pages 137–146, 2011.

[32] Corrado Gini. *Variabilità e mutabilità: contributo allo studio delle distribuzioni e delle relazioni statistiche.[Fasc. I.]*. Tipogr. di P. Cuppini, 1912.

[33] Irving L Glicksberg. A further generalization of the Kakutani fixed point theorem, with application to Nash equilibrium points. *Proceedings of the American Mathematical Society*, 3(1):170–174, 1952.

[34] Chuan Guo, Geoff Pleiss, Yu Sun, and Kilian Q Weinberger. On calibration of modern neural networks. In *International conference on machine learning*, pages 1321–1330. PMLR, 2017.

[35] Jeffrey T Hancock, Mor Naaman, and Karen Levy. AI-mediated communication: Definition, research agenda, and ethical considerations. *Journal of Computer-Mediated Communication*, 25(1):89–100, 2020.

[36] Donald B Hausch, William T Ziemba, and Mark Rubinstein. Efficiency of the market for racetrack betting. *Management science*, 27(12):1435–1452, 1981.

[37] Wassily Hoeffding. A class of statistics with asymptotically normal distribution. *The Annals of Mathematical Statistics*, 19(3):293–325, 1948.

[38] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[39] Ari Holtzman, Jan Buys, Li Du, Maxwell Forbes, and Yejin Choi. The curious case of neural text degeneration. In *8th International Conference on Learning Representations, ICLR*, 2020.

[40] Harold Hotelling. Stability in competition. *The economic journal*, 39(153):41–57, 1929.

[41] Jiri Hron, Karl Krauth, Michael I. Jordan, Niki Kilbertus, and Sarah Dean. Modeling content creator incentives on algorithm-curated platforms. In *The Eleventh International Conference on Learning Representations, ICLR*, 2023.

[42] Xinyan Hu, Meena Jagadeesan, Michael I Jordan, and Jacob Steinhardt. Incentivizing high-quality content in online recommender systems. *arXiv preprint arXiv:2306.07479*, 2023.

[43] Nicole Immorlica, Adam Tauman Kalai, Brendan Lucier, Ankur Moitra, Andrew Postlewaite, and Moshe Tennenholtz. Dueling algorithms. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 215–224, 2011.

[44] Meena Jagadeesan, Nikhil Garg, and Jacob Steinhardt. Supply-side equilibria in recommender systems. In *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems*, 2023.

[45] Meena Jagadeesan, Michael I. Jordan, Jacob Steinhardt, and Nika Haghtalab. Improved bayes risk can yield reduced social welfare under competition. In *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems*, 2023.

[46] Shomik Jain, Vinith Suriyakumar, Kathleen Creel, and Ashia Wilson. Algorithmic pluralism: A structural approach to equal opportunity. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency*, pages 197–206, 2024.

[47] Albert Q Jiang, Alexandre Sablayrolles, Arthur Mensch, Chris Bamford, Devendra Singh Chaplot, Diego de las Casas, Florian Bressand, Gianna Lengyel, Guillaume Lample, Lucile Saulnier, et al. Mistral 7b. *arXiv preprint arXiv:2310.06825*, 2023.

[48] Justin Kaashoek, Manish Raghavan, and John J Horton. The impact of generative AI on labor market matching. *An MIT exploration of generative AI*, 2024.

[49] Saurav Kadavath, Tom Conerly, Amanda Askell, Tom Henighan, Dawn Drain, Ethan Perez, Nicholas Schiefer, Zac Hatfield-Dodds, Nova DasSarma, Eli Tran-Johnson, Scott Johnston, Sheer El-Showk, Andy Jones, Nelson Elhage, Tristan Hume, Anna Chen, Yuntao Bai, Sam Bowman, Stanislav Fort, Deep Ganguli, Danny Hernandez, Josh Jacobson, Jackson Kernion, Shauna Kravec, Liane Lovitt, Kamal Ndousse, Catherine Olsson, Sam Ringer, Dario Amodei, Tom Brown, Jack Clark, Nicholas Joseph, Ben Mann, Sam McCandlish, Chris Olah, and Jared Kaplan. Language models (mostly) know what they know. *arXiv preprint arXiv:2207.05221*, 2022.

[50] Mark Thomas Kennedy and Nelson Phillips. The participation game. *arXiv preprint arXiv:2304.12700*, 2023.

[51] Robert Kirk, Ishita Mediratta, Christoforos Nalmpantis, Jelena Luketina, Eric Hambro, Edward Grefenstette, and Roberta Raileanu. Understanding the effects of RLHF on LLM generalisation and diversity. In *The Twelfth International Conference on Learning Representations, ICLR*, 2024.

[52] Jon Kleinberg and Manish Raghavan. Algorithmic monoculture and social welfare. *Proceedings of the National Academy of Sciences*, 118(22):e2018340118, 2021.

[53] Vicki Knoblauch. Generalizing location games to a graph. *The journal of industrial economics*, pages 683–688, 1991.

[54] Ryan Koo, Minhwa Lee, Vipul Raheja, Jong Inn Park, Zae Myung Kim, and Dongyeop Kang. Benchmarking cognitive biases in large language models as evaluators. In *Findings of the Association for Computational Linguistics, ACL*, pages 517–545. Association for Computational Linguistics, 2024.

[55] Elias Koutsoupias and Christos Papadimitriou. Worst-case equilibria. In *Annual symposium on theoretical aspects of computer science*, pages 404–413. Springer, 1999.

[56] A J Lee. *U-statistics: Theory and Practice*. Routledge, 2019.

[57] Isabelle Levent and Lila Shroff. The model is the message. In *The Second Workshop on Intelligent and Interactive Writing Assistants*, 2023.

[58] Weixin Liang, Zachary Izzo, Yaohui Zhang, Haley Lepp, Hancheng Cao, Xuandong Zhao, Lingjiao Chen, Haotian Ye, Sheng Liu, Zhi Huang, Daniel A. McFarland, and James Y. Zou. Monitoring AI-modified content at scale: A case study on the impact of ChatGPT on AI conference peer reviews. In *Forty-first International Conference on Machine Learning, ICML*, 2024.

[59] Yiqi Liu, Nafise Sadat Moosavi, and Chenghua Lin. LLMs as narcissistic evaluators: When ego inflates evaluation scores. In *Findings of the Association for Computational Linguistics, ACL*, pages 12688–12701. Association for Computational Linguistics, 2024.

[60] Charles Lovering, Michael Krumdick, Viet Dac Lai, Nilesh Kumar, Varshini Reddy, Rik Koncel-Kedziorski, and Chris Tanner. Are language model logits calibrated? *arXiv preprint arXiv:2410.16007*, 2024.

[61] Gili Malinksy. Using ChatGPT can make recruiters think you're sending them a fake resume–here's why. *CNBC*, 2024.

[62] Lev Manovich. *AI aesthetics*. Strelka press Moscow, 2018.

[63] Lisa Messeri and MJ Crockett. Artificial intelligence and illusions of understanding in scientific research. *Nature*, 627(8002):49–58, 2024.

[64] Michael Mitzenmacher. The power of two choices in randomized load balancing. *IEEE Transactions on Parallel and Distributed Systems*, 12(10):1094–1104, 2001.

[65] Martin Mladenov, Elliot Creager, Omer Ben-Porat, Kevin Swersky, Richard Zemel, and Craig Boutilier. Optimizing long-term social welfare in recommender systems: A constrained matching approach. In *International Conference on Machine Learning*, pages 6987–6998. PMLR, 2020.

[66] Dov Monderer and Lloyd S Shapley. Potential games. *Games and economic behavior*, 14(1):124–143, 1996.

[67] Saurav Muralidharan, Sharath Turuvekere Sreenivas, Raviraj Joshi, Marcin Chochowski, Mostofa Patwary, Mohammad Shoeybi, Bryan Catanzaro, Jan Kautz, and Pavlo Molchanov. Compact language models via pruning and knowledge distillation. *arXiv preprint arXiv:2407.14679*, 2024. URL `https://arxiv.org/abs/2407.14679`.

[68] Ingram Olkin and Albert W Marshall. *Inequalities: Theory of Majorization and Its Applications*. Academic Press, 1979.

[69] Vishakh Padmakumar and He He. Does writing with language models reduce content diversity? In *The Twelfth International Conference on Learning Representations, ICLR*, 2024.

[70] Arjun Panickssery, Samuel R Bowman, and Shi Feng. LLM evaluators recognize and favor their own generations. *arXiv preprint arXiv:2404.13076*, 2024.

[71] Peter S Park, Philipp Schoenegger, and Chongyang Zhu. Diminished diversity-of-thought in a standard large language model. *Behavior Research Methods*, pages 1–17, 2024.

[72] Jupinder Parmar, Shrimai Prabhumoye, Joseph Jennings, Mostofa Patwary, Sandeep Subramanian, Dan Su, Chen Zhu, Deepak Narayanan, Aastha Jhunjhunwala, Ayush Dattagupta, et al. Nemotron-4 15B technical report. *arXiv preprint arXiv:2402.16819*, 2024.

[73] Kenny Peng and Nikhil Garg. Monoculture in matching markets. In *Advances in Neural Information Processing Systems*, 2024.

[74] Kenny Peng and Nikhil Garg. Wisdom and foolishness of noisy matching markets. In *Proceedings of the 25th ACM Conference on Economics and Computation*, 2024.

[75] Kenny Peng, Manish Raghavan, Emma Pierson, Jon Kleinberg, and Nikhil Garg. Reconciling the accuracy-diversity trade-off in recommendations. In *Proceedings of the ACM on Web Conference 2024*, pages 1318–1329, 2024.

[76] C. R. Rao. Information and the accuracy attainable in the estimation of statistical parameters. *Bulletin of the Calcutta Mathematical Society*, 37(3):81–91, 1945.

[77] Alfréd Rényi. On measures of entropy and information. In *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability, volume 1: contributions to the theory of statistics*, volume 4, pages 547–562. University of California Press, 1961.

[78] Robert W Rosenthal. A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory*, 2:65–67, 1973.

[79] Günter Rote. Isotonic regression by dynamic programming. In *2nd Symposium on Simplicity in Algorithms (SOSA 2019)*. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2019.

[80] Tim Roughgarden and Éva Tardos. How bad is selfish routing? *Journal of the ACM (JACM)*, 49(2):236–259, 2002.

[81] Walter Rudin. *Principles of mathematical analysis*. McGraw-hill New York, 3 edition, 1976.

[82] Raaj K Sah. The effects of child mortality changes on fertility choice and parental welfare. *Journal of Political Economy*, 99(3):582–606, 1991.

[83] Raaj Kumar Sah. Comparative properties of sums of independent binomials with different parameters. *Economics Letters*, 31(1):27–30, 1989.

[84] Shibani Santurkar, Esin Durmus, Faisal Ladhak, Cinoo Lee, Percy Liang, and Tatsunori Hashimoto. Whose opinions do language models reflect? In *International Conference on Machine Learning*, pages 29971–30004. PMLR, 2023.

[85] Claude Elwood Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423, 1948.

[86] Michal Shur-Ofry, Bar Horowitz-Amsalem, Adir Rahamim, and Yonatan Belinkov. Growing a tail: Increasing output diversity in large language models. *arXiv preprint arXiv:2411.02989*, 2024.

[87] Anand Siththaranjan, Cassidy Laidlaw, and Dylan Hadfield-Menell. Distributional preference learning: Understanding and accounting for hidden context in RLHF. In *The Twelfth International Conference on Learning Representations, ICLR*, 2024.

[88] Taylor Sorensen, Jared Moore, Jillian Fisher, Mitchell L. Gordon, Niloofar Mireshghallah, Christopher Michael Rytting, Andre Ye, Liwei Jiang, Ximing Lu, Nouha Dziri, Tim Althoff, and Yejin Choi. Position: A roadmap to pluralistic alignment. In *Forty-first International Conference on Machine Learning, ICML*, 2024.

[89] Quentin F Stout. Unimodal regression via prefix isotonic regression. *Computational Statistics & Data Analysis*, 53(2):289–297, 2008.

[90] Boaz Taitler and Omer Ben-Porat. Braess's paradox of generative AI. *arXiv preprint arXiv:2409.05506*, 2024.

[91] Gemma Team. Gemma. 2024. doi: 10.34740/KAGGLE/M/3301. URL `https://www.kaggle.com/m/3301`.

[92] Gemma Team, Morgane Riviere, Shreya Pathak, Pier Giuseppe Sessa, Cassidy Hardin, Surya Bhupatiraju, Léonard Hussenot, Thomas Mesnard, Bobak Shahriari, Alexandre Ramé, et al. Gemma 2: Improving open language models at a practical size. *arXiv preprint arXiv:2408.00118*, 2024.

[93] Qwen Team. Qwen2.5: A party of foundation models, September 2024. URL `https://qwenlm.github.io/blog/qwen2.5/`.

[94] Richard H Thaler and William T Ziemba. Anomalies: Parimutuel betting markets: Racetracks and lotteries. *Journal of Economic perspectives*, 2(2):161–174, 1988.

[95] Katherine Tian, Eric Mitchell, Allan Zhou, Archit Sharma, Rafael Rafailov, Huaxiu Yao, Chelsea Finn, and Christopher D. Manning. Just ask for calibration: Strategies for eliciting calibrated confidence scores from language models fine-tuned with human feedback. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, EMNLP*, pages 5433–5442. Association for Computational Linguistics, 2023.

[96] Connor Toups, Rishi Bommasani, Kathleen Creel, Sarah Bana, Dan Jurafsky, and Percy S Liang. Ecosystem-level analysis of deployed machine learning reveals homogeneous outcomes. *Advances in Neural Information Processing Systems*, 36, 2024.

[97] Constantino Tsallis. Possible generalization of Boltzmann-Gibbs statistics. *Journal of statistical physics*, 52:479–487, 1988.

[98] Johan Ugander and Ziv Epstein. The Art of Randomness: Sampling and Chance in the Age of Algorithmic Reproduction. *Harvard Data Science Review*, 6(4), October 2024.

[99] Constance Van Eeden. Maximum likelihood estimation of ordered probabilities. *Indagationes Mathematicae*, 18(4), 1956.

[100] Adrian Vetta. Nash equilibria in competitive societies, with applications to facility location, traffic routing and auctions. In *The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings.*, pages 416–425. IEEE, 2002.

[101] Chaoqi Wang, Yibo Jiang, Chenghao Yang, Han Liu, and Yuxin Chen. Beyond reverse KL: Generalizing direct preference optimization with diverse divergence constraints. *arXiv preprint arXiv:2309.16240*, 2023.

[102] Peiyi Wang, Lei Li, Liang Chen, Zefan Cai, Dawei Zhu, Binghuai Lin, Yunbo Cao, Lingpeng Kong, Qi Liu, Tianyu Liu, and Zhifang Sui. Large language models are not fair evaluators. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL*, pages 9440–9450. Association for Computational Linguistics, 2024.

[103] J Weed. Can ChatGPT plan your vacation? here's what to know about AI and travel. *The New York Times*, 16, 2023.

[104] Fan Wu, Emily Black, and Varun Chandrasekaran. Generative monoculture in large language models. *arXiv preprint arXiv:2407.02209*, 2024.

[105] Jiancong Xiao, Ziniu Li, Xingyu Xie, Emily Getzen, Cong Fang, Qi Long, and Weijie J Su. On the algorithmic bias of aligning large language models with RLHF: Preference collapse and matching regularization. *arXiv preprint arXiv:2405.16455*, 2024.

[106] Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang, and Haifeng Xu. How bad is top-k recommendation under competing content creators? In *International Conference on Machine Learning*, pages 39674–39701. PMLR, 2023.

[107] Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang, and Haifeng Xu. Human vs. generative AI in content creation competition: Symbiosis or conflict? In *Forty-first International Conference on Machine Learning, ICML*, 2024.

[108] Fan Yao, Yiming Liao, Mingzhe Wu, Chuanhao Li, Yan Zhu, James Yang, Jingzhou Liu, Qifan Wang, Haifeng Xu, and Hongning Wang. User welfare optimization in recommender systems with competing content creators. In *Proceedings of the 30th*

ACM SIGKDD Conference on Knowledge Discovery and Data Mining, KDD 2024, Barcelona, Spain, August 25-29, 2024, pages 3874–3885. ACM, 2024.

[109] Simone Zhang, Janet Xu, and AJ Alvero. Generative AI meets open-ended survey responses: Participant use of AI and homogenization, Sep 2024. URL `osf.io/preprints/socarxiv/4esdp`.

[110] Lianmin Zheng, Wei-Lin Chiang, Ying Sheng, Siyuan Zhuang, Zhanghao Wu, Yonghao Zhuang, Zi Lin, Zhuohan Li, Dacheng Li, Eric P. Xing, Hao Zhang, Joseph E. Gonzalez, and Ion Stoica. Judging llm-as-a-judge with MT-Bench and Chatbot Arena. In *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems 2023, NeurIPS 2023, New Orleans, LA, USA, December 10 - 16, 2023*, 2023.

[111] Eric Zhou and Dokyun Lee. Generative artificial intelligence, human creativity, and art. *PNAS nexus*, 3(3):pgae052, 2024.

Organization of the appendix. All deferred proofs for Section 4, along with auxiliary supporting results, can be found in Appendix A. Appendix B contains deferred proofs for Section 5. We restate each result before proving it. In Appendix C, we provide details about our experiments and additional figures.

A Omitted Proofs and Auxiliary Results for Section 4

Here, we provide omitted proofs and auxiliary claims needed to support them. We restate each result before proving it. We begin with a few definitions.

$$b(k,n,p) \triangleq \binom{n}{k} p^k (1-p)^{n-k}$$

$$B(k,n,p) \triangleq \sum_{\ell=0}^{k} b(k,n,p)$$

For a sequence $\{\mathbf{a}_k\}$, and $k_2 \geq k_1$, define

$$\text{avg}(\mathbf{a}; k_1, k_2)_k \triangleq \begin{cases} \frac{\sum_{\ell=k_1}^{k_2} \mathbf{a}_\ell}{k_2 - k_1 + 1} & k_1 \leq k \leq k_2 \\ \mathbf{a}_k & \text{otherwise} \end{cases}.$$

Intuitively, this operation averages all entries between k_1 and k_2. For a function $f(x)$, define $\bar{f}(x)$ to be the discrete derivative $\bar{f}(x) \triangleq f(x+1) - f(x)$. For a function that takes two arguments (e.g., $g(n,p)$), we will take the discrete derivative of the first dimension by convention: $\bar{g}(n,p) \triangleq g(n+1,p) - g(n,p)$.

A.1 Omitted proofs for Section 4.1

Lemma 4.1. *Under Assumption 2, $\mathscr{G}(n, \mathbf{d}, s)$ has a unique symmetric equilibrium $\mathbf{p} = \mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s)$, which is a pure strategy equilibrium. Let $k^* = \max\{k : \mathbf{p}_k > 0\}$. $\mathbf{p}$ satisfies the following properties:*

1. $\mathbf{p} \in \Delta(\pi)$.
2. *For all $k \leq k^*$,*

$$\mathbf{d}_k u(n-1, \mathbf{p}_k) = U(\mathbf{p}).$$

3. *For all $k > k^*$, $U(\mathbf{p}) \geq \mathbf{d}_k$.*

This holds even if we remove Assumption 1 (i.e., if players can choose $\mathbf{p} \notin \Delta(\pi)$).

Proof. By Lemmas A.1 and A.2, $\mathscr{G}(n, \mathbf{d}, s)$ has a symmetric pure strategy equilibrium that yields utility $U(\mathbf{p})$. For now, we remove the restriction that $\mathbf{p} \in \Delta(\pi)$ and instead allow for any distribution $\mathbf{p}$.

Player i's utility for strategy $\mathbf{p}$ is

$$U(\mathbf{p}) = \sum_{k\in[K]} \mathbf{d}_k \mathbf{p}_k \mathbb{E}\left[\frac{1}{s(1+X(n-1,\mathbf{p}_k))}\right] = \sum_{k\in[K]} \mathbf{d}_k \mathbf{p}_k u(n-1,\mathbf{p}_k).$$

If we remove the restriction that $\mathbf{p} \in \Delta(\pi)$, then player i can deviate from $\mathbf{p}$ to any arbitrary distribution $\mathbf{p}'$. This means that for all k,

$$\mathbf{d}_k u(n-1,\mathbf{p}_k) \le U(\mathbf{p}),$$

because otherwise, a player could strictly increase their utility by placing all of their probability mass on type k. Similarly, for all k such that $\mathbf{p}_k > 0$,

$$\mathbf{d}_k u(n-1,\mathbf{p}_k) \ge U(\mathbf{p}),$$

because otherwise, a player could strictly increase their utility by removing mass from k and distributing it to any k' such that $\mathbf{d}_{k'} u(n-1,\mathbf{p}_{k'}) \ge U(\mathbf{p})$. (Such a k' must exist because these quantities average to $U(\mathbf{p})$.)

Finally, note that $u(n-1,\cdot)$ is strictly decreasing by Lemma A.10 and $u(n-1,0) = 1$. Therefore, if $\mathbf{d}_k > U(\mathbf{p})$,

$$\mathbf{d}_k u(n-1,\mathbf{p}_k) \le U(\mathbf{p}) < \mathbf{d}_k,$$

so $\mathbf{p}_k > 0$. If $\mathbf{d}_k \le U(\mathbf{p})$, then

$$\mathbf{d}_k u(n-1,\mathbf{p}_k) \le \mathbf{d}_k \le U(\mathbf{p}).$$

In this case, $\mathbf{p}_k = 0$ because otherwise the first inequality is strict. Let k^* be the largest k such that $\mathbf{d}_{k^*} > U(\mathbf{p})$. Putting these together, under Assumption 2,

$$\begin{aligned} \mathbf{p}_k > 0 &\iff k \le k^* \\ \mathbf{d}_k u(n-1,\mathbf{p}_k) &= U(\mathbf{p}) \qquad (k \le k^*) \end{aligned}$$

For $k \le k^*$, we have

$$u(n-1,\mathbf{p}_k) = \frac{U(\mathbf{p})}{\mathbf{d}_k}.$$

Again using the fact that $u(n-1,\cdot)$ is strictly decreasing, this means that $\mathbf{p}_k$ must also be decreasing, for $k \le k^*$. And since $\mathbf{p}_k = 0$ for $k > k^*$, we have shown $\mathbf{p} \in \Delta(\pi)$.

Finally, we show that $\mathbf{p}$ is the unique symmetric equilibrium. Assume towards contradiction that there is some other symmetric equilibrium $\mathbf{p}' \ne \mathbf{p}$. There must be some k such that $\mathbf{p}'_k > 0$, $\mathbf{p}_k > 0$, and $\mathbf{p}'_k \ne \mathbf{p}_k$. Because $u(n-1,\cdot)$ is strictly decreasing, $U(\mathbf{p}') = \mathbf{d}_k u(n-1,\mathbf{p}'_k) \ne \mathbf{d}_k u(n-1,\mathbf{p}_k) = U(\mathbf{p})$.

Let $u^{-1}(n,p)$ be the inverse of u, such that $u^{-1}(n,u(n,p)) = p$. We only require that u^{-1} is defined on the range $[u(n,1),u(n,0)]$; by convention, define it to be 0 otherwise. It is

well-defined because $u(n,\cdot)$ is strictly decreasing in that range. Observe that $u^{-1}(n-1,\cdot)$ is strictly decreasing on the range where it is nonzero. Then, let

$$S(c) \triangleq \sum_{k\in[K]} u^{-1}\left(n-1, \frac{c}{\mathbf{d}_k}\right).$$

Because at least one u^{-1} term is strictly decreasing in c (i.e., all terms corresponding to nonzero $\mathbf{p}_k$'s), S is strictly decreasing in c. Observe that

$$\begin{aligned} S(U(\mathbf{p})) &= \sum_{k\in[K]} u^{-1}\left(n-1, \frac{U(\mathbf{p})}{\mathbf{d}_k}\right) \\ &= \sum_{k\in[K]} \mathbf{p}_k \\ &= 1, \end{aligned}$$

and similarly, $S(U(\mathbf{p}')) = 1$. But we know that S is strictly decreasing, so $S(U(\mathbf{p})) = S(U(\mathbf{p}'))$ implies $U(\mathbf{p}) = U(\mathbf{p}')$, which is a contradiction. Therefore, $\mathbf{p}$ is the unique symmetric equilibrium. □

Lemma 4.2. *Under Assumption 2, $\mathscr{G}(n,\mathbf{d},s)$ has a unique symmetric optimal strategy $\mathbf{p}^* = \mathbf{p}_{\text{OPT}}(n,\mathbf{d},s)$. Let $c = \frac{1}{n}\|\nabla W(\mathbf{p}^*)\|_\infty$, and let $k^* = \max\{k : \mathbf{p}^*_k > 0\}$. $\mathbf{p}^*$ satisfies the following properties:*

1. *$\mathbf{p}^* \in \Delta(\pi)$.*
2. *If $s \in \mathcal{S}_\nearrow$, for all $k \le k^*$,*

$$\frac{1}{n}\left(\nabla W(\mathbf{p}^*)\right)_k = \frac{\partial}{\partial p}\mathbf{d}_k p u_s(n-1,p)\Big|_{p=\mathbf{p}^*_k} = c.$$

3. *If $s \in \mathcal{S}_\searrow$, for all $k \le k^*$, $\mathbf{d}_k(1-\mathbf{p}^*_k)^{n-1} = c$.*
4. *For all $k > k^*$, $c \ge \mathbf{d}_k$.*

This holds even if we remove Assumption 1 (i.e., if players can choose $\mathbf{p} \notin \Delta(\pi)$).

Proof. We begin with the case where $s \in \mathcal{S}_\nearrow$. By Lemma A.2, it suffices to consider pure strategies. The socially optimal symmetric pure strategy is the solution to the optimization problem

$$\begin{aligned} \max_{\mathbf{p}} \ & W(\mathbf{p}) \\ \text{s.t.} \ & \sum_{k\in[K]} \mathbf{p}_k \le 1 \\ & \mathbf{p}_{k+1} \le \mathbf{p}_k && (\forall k \in [K-1]) \\ & \mathbf{p}_K \ge 0. \end{aligned}$$

Because $\mathscr{G}$ is symmetric and social welfare is the sum of utilities for $s \in \mathcal{S}_{\nearrow}$, this is equivalent to maximizing a single player's utility:

$$\begin{aligned} \max_{\mathbf{p}} \ & U(\mathbf{p}) \\ \text{s.t.} \ & \sum_{k\in[K]} \mathbf{p}_k \le 1 \\ & \mathbf{p}_{k+1} \le \mathbf{p}_k && (\forall k \in [K-1]) \\ & \mathbf{p}_K \ge 0. \end{aligned}$$

As in the proof of Lemma 4.1, we will remove the assumption that $\mathbf{p} \in \Delta(\pi)$ and show that even without this assumption, $\mathbf{p}_{\text{OPT}} \in \Delta(\pi)$. Without the assumption that $\mathbf{p} \in \Delta(\pi)$, we can rewrite utility-maximization problem as

$$\begin{aligned} \max_{\mathbf{p}} \ & U(\mathbf{p}) \\ \text{s.t.} \ & \sum_{k\in[K]} \mathbf{p}_k \le 1 \\ & \mathbf{p}_k \ge 0. && (\forall k \in [K]) \end{aligned}$$

Intuitively, we would expect $\frac{\partial}{\partial \mathbf{p}_k} U(\mathbf{p})$ to be equal for all k such that $\mathbf{p}_k > 0$, since the optimal solution should equalize the marginal gain from increasing any nonzero $\mathbf{p}_k$. We formalize this in what follows. The Lagrangian is

$$\begin{aligned} \mathcal{L}(\mathbf{p}; \mu, \lambda) &= -U(\mathbf{p}) + \mu \left(\sum_{k\in[K]} \mathbf{p}_k - 1 \right) - \sum_{k\in[K]} \lambda_k \mathbf{p}_k \\ &= -\sum_{k\in[K]} \mathbf{d}_k \mathbf{p}_k u(n-1, \mathbf{p}_k) + \mu \left(\sum_{k\in[K]} \mathbf{p}_k - 1 \right) - \sum_{k\in[K]} \lambda_k \mathbf{p}_k. \end{aligned}$$

By stationarity, for the socially optimal strategy $\mathbf{p}^*$, for all $k \in [K]$,

$$\begin{aligned} \frac{\partial}{\partial \mathbf{p}_k} \mathcal{L}(\mathbf{p}; \mu, \lambda)\Big|_{\mathbf{p}_k = \mathbf{p}_k^*} &= 0 \\ -\mathbf{d}_k \frac{\partial}{\partial p} p u(n-1, p)\Big|_{p=\mathbf{p}_k^*} + \mu - \lambda_k &= 0. \end{aligned}$$

Consider an optimal solution $\mathbf{p}^*$. By complementary slackness,

$$\mathbf{p}_k^* > 0 \implies \lambda_k = 0.$$

Therefore, for all k such that $\mathbf{p}_k^* > 0$,

$$\mathbf{d}_k \frac{\partial}{\partial p} p u(n-1, p)\Big|_{p=\mathbf{p}_k^*} = \mu. \tag{3}$$

For all k such that $\mathbf{p}_k^* = 0$,

$$\mathbf{d}_k \frac{\partial}{\partial p} pu(n-1,p)\Big|_{p=\mathbf{p}_k^*} = \mu - \lambda_k$$
$$\mathbf{d}_k \frac{\partial}{\partial p} pu(n-1,p)\Big|_{p=0} \leq \mu.$$

Note that

$$\frac{\partial}{\partial p} pu(n-1,p)\Big|_{p=0} = \left(p \cdot \frac{\partial}{\partial p} u(n-1,p)\right)\Big|_{p=0} + u(n-1,0) = 1.$$

Therefore, for k such that $\mathbf{p}_k^* = 0$,

$$\mathbf{d}_k \leq \mu.$$

Next, we will show that $\frac{\partial}{\partial p} pu(n-1,p)$ is strictly decreasing in p. Define

$$w_s(n,p) \triangleq \mathbb{E}\left[\frac{X(n,p)}{s(X(n,p))}\right] = npu(n-1,p).$$

As before, we will drop the subscript s when it is clear from context. Then,

$$\frac{\partial^2}{\partial p^2} pu(n-1,p) < 0 \iff \frac{\partial^2}{\partial p^2} w(n,p) < 0.$$

Because $x/s(x)$ is strictly concave by assumption (**D.3** and **D.4(a)**), w is the expectation of a strictly concave function of a binomial random variable, which is strictly concave by Lemma A.11.

Given μ, we therefore have a tight characterization of $\mathbf{p}^*$:

1. If $\mathbf{d}_k > \mu$, then we must have $\mathbf{p}_k^* > 0$, and in particular,

$$\mathbf{d}_k \frac{\partial}{\partial p} pu(n-1,p)\Big|_{p=\mathbf{p}_k^*} = \mu.$$

2. If $\mathbf{d}_k \leq \mu$, then we must have

$$\mathbf{d}_k \frac{\partial}{\partial p} pu(n-1,p)\Big|_{p=\mathbf{p}_k^*} \leq \mathbf{d}_k \leq \mu.$$

   If $\mathbf{p}_k^* > 0$, the first of these inequalities is strict, which contradicts (3). Therefore, $\mathbf{d}_k \leq \mu$ implies $\mathbf{p}_k^* = 0$.

To show uniqueness, it suffices to note that the objective of our optimization problem is strictly concave, meaning it has a unique optimum. Because $\mathbf{p}_1^* > 0$, $\mu = \frac{\partial}{\partial \mathbf{p}_1} U(\mathbf{p})\big|_{\mathbf{p}=\mathbf{p}^*}$. Thus, $\mu = \|\nabla U(\mathbf{p}^*)\|_\infty = \frac{1}{n}\|\nabla W(\mathbf{p}^*)\|_\infty$.

Observe that $\mathbf{p}^* \in \Delta(\pi)$: for any k such that $\mathbf{p}^*_k > 0$ and $\mathbf{p}^*_{k+1} > 0$,

$$\begin{aligned}
\mathbf{d}_k \frac{\partial}{\partial p} pu(n-1,p)\Big|_{p=\mathbf{p}^*_k} &= \mu = \mathbf{d}_{k+1} \frac{\partial}{\partial p} pu(n-1,p)\Big|_{p=\mathbf{p}^*_{k+1}} \\
\frac{\partial}{\partial p} pu(n-1,p)\Big|_{p=\mathbf{p}^*_k} &\leq \frac{\partial}{\partial p} pu(n-1,p)\Big|_{p=\mathbf{p}^*_{k+1}} && (\mathbf{d}_k \geq \mathbf{d}_{k+1}) \\
\mathbf{p}^*_k &\geq \mathbf{p}^*_{k+1}. && (\tfrac{\partial}{\partial p} pu(n-1,p) \text{ strictly decreasing})
\end{aligned}$$

Next, we handle the case where $s \in \mathcal{S}_\searrow$. By definition,

$$\begin{aligned}
W(\mathbf{p}) &= \sum_{k \in [K]} \mathbf{d}_k (1 - (1 - \mathbf{p}^*_k)^n) \\
&= \sum_{k \in [K]} \mathbb{E}\left[\frac{X(n,p)}{X(n,p)} \cdot \mathbb{1}\left[X(n, p > 0)\right] \right] \\
&= \sum_{k \in [K]} \mathbb{E}\left[\frac{X(n,p)}{s_1(X(n,p))} \right].
\end{aligned}$$

Thus, welfare under any $s \in \mathcal{S}_\searrow$ is equivalent to welfare under the identity score function $s_1(x) = x$. Because $s_1 \in \mathcal{S}_\nearrow$, our results above hold. Let $c = \frac{1}{n}\|\nabla W(\mathbf{p}^*)\|_\infty$. For all k such that $\mathbf{p}^*_k > 0$,

$$c = \frac{\mathbf{d}_k}{n} \frac{\partial}{\partial p} w_{s_1}(n,p)\Big|_{p=\mathbf{p}^*_k}.$$

For $k > k^*$, $c \geq \mathbf{d}_k$. To conclude, we must show that for k such that $\mathbf{p}^*_k > 0$,

$$\begin{aligned}
c &= \frac{\mathbf{d}_k}{n} \frac{\partial}{\partial p} w_{s_1}(n,p)\Big|_{p=\mathbf{p}^*_k} \\
&= \frac{\mathbf{d}_k}{n} \frac{\partial}{\partial p} \mathbb{E}\left[\mathbb{1}\left[X(n,p) > 0\right]\right]\Big|_{p=\mathbf{p}^*_k} \\
&= \frac{\mathbf{d}_k}{n} \frac{\partial}{\partial p} (1 - (1-p)^n)\Big|_{p=\mathbf{p}^*_k} \\
&= \frac{\mathbf{d}_k}{n} n(1-p)^{n-1}\Big|_{p=\mathbf{p}^*_k} \\
&= \mathbf{d}_k (1 - \mathbf{p}^*_k)^{n-1}.
\end{aligned}$$

□

Theorem 4.3. *For every $\mathbf{d} \in \mathbb{R}^K_{\geq 0}$, there is some $\tilde{\mathbf{d}} \in \mathbb{R}^K_{\geq 0}$ such that $\tilde{\mathbf{d}}$ is decreasing in k, and for all $n \geq 2$, $s \in \mathcal{S}$,*

1. *$\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s) = \mathbf{p}_{\text{EQ}}(n, \tilde{\mathbf{d}}, s)$, and per-player utilities are equal.*
2. *$\mathbf{p}_{\text{OPT}}(n, \mathbf{d}, s) = \mathbf{p}_{\text{OPT}}(n, \tilde{\mathbf{d}}, s)$, and per-player utilities are equal.*

By Lemmas 4.1 and 4.2, $\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s)$ and $\mathbf{p}_{\text{OPT}}(n, \mathbf{d}, s)$ are both unique.

Proof. First, Lemmas A.1 and A.2 imply that $\mathscr{G}(n, \mathbf{d}, s)$ has a symmetric pure Nash equilibrium. By Lemma A.3, there exists decreasing $\tilde{\mathbf{d}}$ such that every symmetric equilibrium or socially optimal symmetric strategy in $\mathscr{G}(n, \mathbf{d}, s)$ is an equilibrium or socially optimal strategy respectively in $\mathscr{G}(n, \tilde{\mathbf{d}}, s)$ with the same utilities. By Lemmas 4.1 and 4.2, $\mathbf{p}_{\text{EQ}}(n, \tilde{\mathbf{d}}, s)$ and $\mathbf{p}_{\text{OPT}}(n, \tilde{\mathbf{d}}, s)$ are both unique for $\mathscr{G}(n, \tilde{\mathbf{d}}, s)$. This implies that $\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s)$ and $\mathbf{p}_{\text{OPT}}(n, \mathbf{d}, s)$ are both unique for $\mathscr{G}(n, \mathbf{d}, s)$ as well.

□

A.1.1 Additional results on equilibria

Lemma A.1. *$\mathscr{G}(n, \mathbf{d}, s)$ has a symmetric mixed Nash equilibrium.*

Proof. $\mathscr{G}$ is a symmetric game with a compact, convex strategy space $\Delta(\pi)$. Utilities are continuous in strategies. Using classical results by Glicksberg [33], Becker and Damianov [9, Theorem 1] show that this implies $\mathscr{G}$ has a symmetric mixed Nash equilibrium. □

Lemma A.2. *Every mixed strategy in the game $\mathscr{G}(n, \mathbf{d}, s)$ is equivalent to a pure strategy.*

Proof. Consider a mixed strategy q, which is a distribution over distributions $\mathbf{p}^{(1)}, \mathbf{p}^{(2)}, \ldots$. Let $\hat{\mathbf{p}}_k(q) = \mathbb{E}_{\mathbf{p} \sim q}[\mathbf{p}_k]$. A player samples type k with probability $\hat{\mathbf{p}}_k(q)$ under both q and $\hat{\mathbf{p}}(q)$, meaning no downstream game can distinguish between them. $\hat{\mathbf{p}} \in \Delta(\pi)$ because $\Delta(\pi)$ is convex: for any k and i, $\mathbf{p}_k^{(i)} \geq \mathbf{p}_{k+1}^{(i)}$. Therefore,

$$\hat{\mathbf{p}}_k(q) = \sum_i q_i \mathbf{p}_k^{(i)} \geq \sum_i q_i \mathbf{p}_{k+1}^{(i)} = \hat{\mathbf{p}}_{k+1}(q).$$

□

Lemma A.3. *For any $n \geq 2$, $s \in \mathcal{S}$, and $\mathbf{d}_k$, there exists a non-increasing $\tilde{\mathbf{d}}_k$ such that:*

1. *$\mathbf{p}$ is a symmetric equilibrium for $\mathscr{G}(n, \mathbf{d}, s)$ only if it is an equilibrium for $\mathscr{G}(n, \tilde{\mathbf{d}}, s)$. $U(\mathbf{p}; n, \mathbf{d}, s) = U(\mathbf{p}; n, \tilde{\mathbf{d}}, s)$, and $W(\mathbf{p}; n, \mathbf{d}, s) = W(\mathbf{p}; n, \tilde{\mathbf{d}}, s)$.*
2. *$\mathbf{p}^*$ is a socially optimal symmetric strategy for $\mathscr{G}(n, \mathbf{d}, s)$ only if it is a socially optimal symmetric strategy for $\mathscr{G}(n, \tilde{\mathbf{d}}, s)$. Moreover, $U(\mathbf{p}^*; n, \mathbf{d}, s) = U(\mathbf{p}^*; n, \tilde{\mathbf{d}}, s)$, and $W(\mathbf{p}^* n, \mathbf{d}, s*) = W(\mathbf{p}^*; n, \tilde{\mathbf{d}}, s)$.*

Proof. First, we will show that for any k_1, k_2 such that $\mathbf{d}_{k_1} \leq \mathbf{d}_{k_1+1} \leq \cdots \leq \mathbf{d}_{k_2}$, the two desired properties hold for $\tilde{\mathbf{d}} = \text{avg}(\mathbf{d}; k_1, k_2)$. Applying this iteratively will yield the result. At a high level, we will argue that if $\mathbf{d}_k \leq \mathbf{d}_{k+1}$, both equilibrium and optimal strategies will assign them equal probabilities, i.e., $\mathbf{p}_k = \mathbf{p}_{k+1}$. By linearity, outcomes will be identical to the case where $\mathbf{d}_k$ and $\mathbf{d}_{k+1}$ are replaced by their averages.

Equilibria. We begin by considering symmetric equilibria. Assume $\mathbf{d}_{k_1} \leq \mathbf{d}_{k_1+1} \leq \cdots \leq \mathbf{d}_{k_2}$. Let $\tilde{\mathbf{d}} = \text{avg}(\mathbf{d}; k_1, k_2)$. Let $\mathbf{p}$ be a symmetric equilibrium for $\mathscr{G}(n, \mathbf{d}, s)$. We will show that $\mathbf{p}$ is an equilibrium for $\mathscr{G}(n, \tilde{\mathbf{d}}, s)$.

By Lemma A.13, $\mathbf{p}$ satisfies $\mathbf{p} = \text{avg}(\mathbf{p}; k_1, k_2)$. Consider some other $\mathbf{p}' \neq \mathbf{p}$.

$$\begin{aligned}
U(\mathbf{p}', \mathbf{p}; n, \tilde{\mathbf{d}}, s) &\leq U(\text{avg}(\mathbf{p}'; k_1, k_2), \mathbf{p}; n, \tilde{\mathbf{d}}, s) && \text{(Lemma A.13)} \\
&= U(\text{avg}(\mathbf{p}'; k_1, k_2), \mathbf{p}; n, \mathbf{d}, s) && \text{(Lemma A.14)} \\
&\leq U(\mathbf{p}, \mathbf{p}; n, \mathbf{d}, s) && (\mathbf{p} \text{ is an equilibrium for } \mathbf{d}) \\
&= U(\mathbf{p}, \mathbf{p}; n, \tilde{\mathbf{d}}, s). && \text{(Lemma A.14)}
\end{aligned}$$

Further, by Lemma A.14, $U(\mathbf{p}; n, \mathbf{d}, s) = U(\mathbf{p}; n, \tilde{\mathbf{d}}, s)$ and $W(\mathbf{p}; n, \mathbf{d}, s) = W(\mathbf{p}; n, \tilde{\mathbf{d}}, s)$.

Optimal strategies. Next, let $\mathbf{p}^*$ be a symmetric optimal strategy for $\mathscr{G}(n, \mathbf{d}, s)$. Again using Lemma A.13, $\mathbf{p}^* = \text{avg}(\mathbf{p}^*; k_1, k_2)$. Therefore, for any other $\mathbf{p}'$,

$$\begin{aligned}
W(\mathbf{p}^*; n, \tilde{\mathbf{d}}, s) &= W(\mathbf{p}^*; n, \mathbf{d}, s) && \text{(Lemma A.14)} \\
&\geq W(\text{avg}(\mathbf{p}'; k_1, k_2); n, \mathbf{d}, s) && (\mathbf{p}^* \text{ socially optimal for } \mathbf{d}) \\
&= W(\text{avg}(\mathbf{p}'; k_1, k_2); n, \tilde{\mathbf{d}}, s) && \text{(Lemma A.14)} \\
&\geq W(\mathbf{p}'; n, \tilde{\mathbf{d}}, s). && \text{(Lemma A.13)}
\end{aligned}$$

Further, by Lemma A.14, $U(\mathbf{p}^*; n, \mathbf{d}, s) = U(\mathbf{p}^*; n, \tilde{\mathbf{d}}, s)$ and $W(\mathbf{p}^*; n, \mathbf{d}, s) = W(\mathbf{p}^*; n, \tilde{\mathbf{d}}, s)$.

Finally, observe that we can transform any $\mathbf{d}$ into a non-increasing $\hat{\mathbf{d}}$ by successively applying averaging operations via the Pool Adjacent Violators algorithm [7]. Let k^+ be the largest k such that $\mathbf{d}_k > \mathbf{d}_{k-1}$. Let k_1, k_2 be the minimal and maximal values such that $k_1 < k^+ \leq k_2$ and $\mathbf{d}_{k_1} \leq \mathbf{d}_{k_1+1} \leq \cdots \leq \mathbf{d}_{k_2}$. By the above argument, the desired properties hold for $\tilde{\mathbf{d}} = \text{avg}(\mathbf{d}; k_1, k_2)$. We iteratively apply this argument until $\tilde{\mathbf{d}}$ is non-increasing. Note that this takes at most K iterative applications because each time we do this, we are guaranteed that $\tilde{\mathbf{d}}_{k_1} = \tilde{\mathbf{d}}_{k_1+1} = \cdots = \tilde{\mathbf{d}}_{k_2}$ in all future iterations, reducing the number of distinct values in $\tilde{\mathbf{d}}$ by at least 1. □

A.2 Omitted proofs for Section 4.2

Theorem 4.5 (More players $\Longrightarrow$ more diverse equilibria)**.** *For all $n \geq 2$, $\mathbf{d} \in \mathbb{R}^K_{\geq 0}$, $s \in \mathcal{S}$,*

$$\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s) \succ \mathbf{p}_{\text{EQ}}(n+1, \mathbf{d}, s).$$

Proof. Assume without loss of generality that $\mathbf{d}$ is decreasing. Let $\mathbf{p}^{(n)} = \mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s)$ and $\mathbf{p}^{(n+1)} = \mathbf{p}_{\text{EQ}}(n+1, \mathbf{d}, s)$. Our plan will be to show that there is some k^* such that $\mathbf{p}^{(n)}_k > \mathbf{p}^{(n+1)}_k$ if and only if $k \leq k^*$, and then apply Lemma A.5 to conclude that $\mathbf{p}^{(n)} \succ \mathbf{p}^{(n+1)}$.

Our goal will be to show that for $k < k'$, if $\mathbf{p}^{(n)}_{k'} > 0$, then

$$\mathbf{d}_k u(n, \mathbf{p}^{(n)}_k) \leq \mathbf{d}_{k'} u(n, \mathbf{p}^{(n)}_{k'}). \tag{4}$$

(Note that $\mathbf{p}^{(n)}_{k'} > 0$ implies $\mathbf{p}^{(n)}_k > 0$ because $\mathbf{p}^{(n)} \in \Delta(\pi)$.) Let $c^{(n)} = U(\mathbf{p}^{(n)}; n, \mathbf{d}, s)$. By Lemma 4.1, for all k such that $\mathbf{p}^{(n)}_k > 0$, $\mathbf{d}_k u(n, \mathbf{p}^{(n)}_k) = c^{(n)}$. Therefore,

$$\mathbf{d}_k u(n-1, \mathbf{p}^{(n)}_k) = c^{(n)} = \mathbf{d}_{k'} u(n-1, \mathbf{p}^{(n)}_{k'}). \tag{5}$$

Dividing each side of (4) by each respective side of (5), our goal is to show that

$$\frac{\mathbf{d}_k u(n, \mathbf{p}_k^{(n)})}{\mathbf{d}_k u(n-1, \mathbf{p}_k^{(n)})} \leq \frac{\mathbf{d}_{k'} u(n, \mathbf{p}_{k'}^{(n)})}{\mathbf{d}_{k'} u(n-1, \mathbf{p}_{k'}^{(n)})}$$
$$\frac{u(n, \mathbf{p}_k^{(n)})}{u(n-1, \mathbf{p}_k^{(n)})} \leq \frac{u(n, \mathbf{p}_{k'}^{(n)})}{u(n-1, \mathbf{p}_{k'}^{(n)})}.$$

Because $\mathbf{p}_{k'}^{(n)} \leq \mathbf{p}_k^{(n)}$, it suffices to show that

$$\frac{\partial}{\partial p} \frac{u(n,p)}{u(n-1,p)} \leq 0.$$

This holds by Lemma A.9. We can now use Lemma A.4 with $\mathbf{p} = \mathbf{p}^{(n+1)}$, $\mathbf{p}' = \mathbf{p}^{(n)}$, and $Q_k = \mathbf{d}_k u(n, \mathbf{p}^{(n)})$. By (4), $\{Q_k\}$ is increasing as long as $\mathbf{p}_k^{(n)} > 0$, so $\mathbf{p}^{(n)} \succ \mathbf{p}^{(n+1)}$. □

Theorem 4.6 (Greater externalities $\Longrightarrow$ more diverse equilibria)**.** *For all $n \geq 2$, $\mathbf{d} \in \mathbb{R}^K_{\geq 0}$, $\gamma_0 \leq \gamma_1$,*

$$\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s_{\gamma_0}) \succ \mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s_{\gamma_1}).$$

Proof. Assume without loss of generality that $\mathbf{d}$ is decreasing. Let $\mathbf{p}^{(\gamma_0)} = \mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s_{\gamma_0})$ and $\mathbf{p}^{(\gamma_1)} = \mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s_{\gamma_1})$. We will use a similar approach to the proof of Theorem 4.5 to show that $Q_k = \{\mathbf{d}_k u_{s_{\gamma_1}}(n-1, \mathbf{p}_k^{(\gamma_0)})\}$ is increasing as long as $\mathbf{p}_k^{(\gamma_0)} > 0$. We will then apply Lemma A.4 to conclude that $\mathbf{p}^{(\gamma_0)} \succ \mathbf{p}^{(\gamma_1)}$. Abusing notation slightly, define

$$u_\gamma(n,p) \triangleq \mathbb{E}\left[\frac{1}{s_\gamma(1+X(n,p))}\right] = \mathbb{E}\left[\frac{1}{(1+X(n,p))^\gamma}\right].$$

As we did with Theorem 4.5, we begin by showing that for $k < k'$, if $\mathbf{p}_{k'}^{(\gamma_0)} > 0$, then

$$\mathbf{d}_k u_{\gamma_1}(n-1, \mathbf{p}_k^{(\gamma_0)}) \leq \mathbf{d}_{k'} u_{\gamma_1}(n-1, \mathbf{p}_{k'}^{(\gamma_0)}). \tag{6}$$

Observe that

$$u_{\gamma_1}(n-1,p) = u_{\gamma_0}(n-1,p) + \int_{\gamma_0}^{\gamma_1} \frac{\partial}{\partial \gamma} u_\gamma(n-1,p)\, d\gamma. \tag{7}$$

Further, by Lemma 4.1, if $k < k'$ and $\mathbf{p}_k^{(\gamma_0)} \geq \mathbf{p}_{k'}^{(\gamma_0)} > 0$, then

$$\mathbf{d}_k u_{\gamma_0}(n-1, \mathbf{p}_k^{(\gamma_0)}) = U(\mathbf{p}^{(\gamma_0)}; n, \mathbf{d}, s_{\gamma_0}) = \mathbf{d}_{k'} u_{\gamma_0}(n-1, \mathbf{p}_k^{(\gamma_0)}). \tag{8}$$

Substituting (7) and dividing both sides by the left and right of (8) respectively, (6) is true if and only if

$$\frac{\mathbf{d}_k \left(u_{\gamma_0}(n-1, \mathbf{p}_k^{(\gamma_0)}) + \int_{\gamma_0}^{\gamma_1} \frac{\partial}{\partial \gamma} u_\gamma(n-1, \mathbf{p}_k^{(\gamma_0)})\, d\gamma\right)}{\mathbf{d}_k u_{\gamma_0}(n-1, \mathbf{p}_k^{(\gamma_0)})}$$
$$\leq \frac{\mathbf{d}_{k'} \left(u_{\gamma_0}(n-1, \mathbf{p}_{k'}^{(\gamma_0)}) + \int_{\gamma_0}^{\gamma_1} \frac{\partial}{\partial \gamma} u_\gamma(n-1, \mathbf{p}_{k'}^{(\gamma_0)})\, d\gamma\right)}{\mathbf{d}_{k'} u_{\gamma_0}(n-1, \mathbf{p}_{k'}^{(\gamma_0)})}.$$

Simplifying, this is

$$\frac{\int_{\gamma_0}^{\gamma_1} \frac{\partial}{\partial\gamma} u_\gamma(n-1, \mathbf{p}_k^{(\gamma_0)})\, d\gamma}{u_{\gamma_0}(n-1, \mathbf{p}_k^{(\gamma_0)})} - \frac{\int_{\gamma_0}^{\gamma_1} \frac{\partial}{\partial\gamma} u_\gamma(n-1, \mathbf{p}_{k'}^{(\gamma_0)})\, d\gamma}{u_{\gamma_0}(n-1, \mathbf{p}_{k'}^{(\gamma_0)})} \leq 0$$
$$\int_{\gamma_0}^{\gamma_1} \frac{\partial}{\partial\gamma} \left(\frac{u_\gamma(n-1, \mathbf{p}_k^{(\gamma_0)})}{u_{\gamma_0}(n-1, \mathbf{p}_k^{(\gamma_0)})} - \frac{u_\gamma(n-1, \mathbf{p}_{k'}^{(\gamma_0)})}{u_{\gamma_0}(n-1, \mathbf{p}_{k'}^{(\gamma_0)})} \right) d\gamma \leq 0.$$

It suffices to show that the term inside the integral is nonpositive for all $\gamma \in [\gamma_0, \gamma_1]$. For now, assume that $\gamma_1 \leq \gamma_0 + \frac{1}{\log n}$. We will show that under this assumption, (6) holds, and we will use Lemma A.4 to conclude that $\mathbf{p}^{(\gamma_0)} \succ \mathbf{p}^{(\gamma_1)}$ for all such γ_1. Iteratively applying this result will allow us to remove the assumption. Our goal is to prove that

$$\frac{\partial}{\partial\gamma} \left(\frac{u_\gamma(n-1, \mathbf{p}_k^{(\gamma_0)})}{u_{\gamma_0}(n-1, \mathbf{p}_k^{(\gamma_0)})} - \frac{u_\gamma(n-1, \mathbf{p}_{k'}^{(\gamma_0)})}{u_{\gamma_0}(n-1, \mathbf{p}_{k'}^{(\gamma_0)})} \right) \leq 0$$
$$\frac{\frac{\partial}{\partial\gamma} u_\gamma(n-1, \mathbf{p}_k^{(\gamma_0)})}{u_{\gamma_0}(n-1, \mathbf{p}_k^{(\gamma_0)})} \leq \frac{\frac{\partial}{\partial\gamma} u_\gamma(n-1, \mathbf{p}_{k'}^{(\gamma_0)})}{u_{\gamma_0}(n-1, \mathbf{p}_{k'}^{(\gamma_0)})}.$$

Because $\mathbf{p}_k^{(\gamma_0)} \geq \mathbf{p}_{k'}^{(\gamma_0)}$, it suffices to show that this ratio is decreasing in p:

$$\frac{\partial}{\partial p} \frac{\frac{\partial}{\partial\gamma} u_\gamma(n-1, p)}{u_{\gamma_0}(n-1, p)} \leq 0. \tag{9}$$

First, we take the partial with respect to γ:

$$\begin{aligned}
\frac{\partial}{\partial\gamma} u_\gamma(n-1, p) &= \frac{\partial}{\partial\gamma} \mathbb{E}\left[(1 + X(n-1, p))^{-\gamma}\right] \\
&= \mathbb{E}\left[\frac{\partial}{\partial\gamma}(1 + X(n-1, p))^{-\gamma}\right] \\
&= \mathbb{E}\left[-\frac{\log(1 + X(n-1, p))}{(1 + X(n-1, p))^{-\gamma}}\right].
\end{aligned}$$

Define

$$z_\gamma(n, p) \triangleq \mathbb{E}\left[\frac{\log(1 + X(n, p))}{(1 + X(n, p))^\gamma}\right] = -\frac{\partial}{\partial\gamma} u_\gamma(n, p).$$

Plugging this into (9), we want to show that

$$\frac{\partial}{\partial p} \frac{z_\gamma(n-1, p)}{u_{\gamma_0}(n-1, p)} \geq 0. \tag{10}$$

This holds if and only if

$$\begin{aligned}
\frac{\partial}{\partial p}\log\left(\frac{z_\gamma(n-1,p)}{u_{\gamma_0}(n-1,p)}\right) &\geq 0 \\
\frac{\partial}{\partial p}\log\left(z_\gamma(n-1,p)\right) &\geq \frac{\partial}{\partial p}\log\left(u_{\gamma_0}(n-1,p)\right) \\
\frac{\frac{\partial}{\partial p}z_\gamma(n-1,p)}{z_\gamma(n-1,p)} &\geq \frac{\frac{\partial}{\partial p}u_{\gamma_0}(n-1,p)}{u_{\gamma_0}(n-1,p)} \\
\frac{\frac{n-1}{p}\bar{z}_\gamma(n-2,p)}{z_\gamma(n-1,p)} &\geq \frac{\frac{n-1}{p}\bar{u}_{\gamma_0}(n-2,p)}{u_{\gamma_0}(n-1,p)} \qquad \text{(by Lemma A.7)} \\
\frac{z_\gamma(n-1,p)-z_\gamma(n-2,p)}{z_\gamma(n-1,p)} &\geq \frac{u_{\gamma_0}(n-1,p)-u_{\gamma_0}(n-2,p)}{u_{\gamma_0}(n-1,p)} \\
\frac{z_\gamma(n-2,p)}{z_\gamma(n-1,p)} &\leq \frac{u_{\gamma_0}(n-2,p)}{u_{\gamma_0}(n-1,p)}
\end{aligned}$$

Here, we need one more pair of definitions:

$$\begin{aligned}
u_\gamma^+(n,p) &\triangleq \mathbb{E}\left[(2+X(n,p))^{-\gamma}\right] \\
z_\gamma^+(n,p) &\triangleq \mathbb{E}\left[\frac{\log(2+X(n,p))}{(2+X(n,p))^\gamma}\right].
\end{aligned}$$

With these, we must show that

$$\begin{aligned}
\frac{z_\gamma(n-2,p)}{z_\gamma(n-1,p)} &\leq \frac{u_{\gamma_0}(n-2,p)}{u_{\gamma_0}(n-1,p)} \\
\frac{z_\gamma(n-1,p)}{z_\gamma(n-2,p)} &\geq \frac{u_{\gamma_0}(n-1,p)}{u_{\gamma_0}(n-2,p)} \\
\frac{(1-p)z_\gamma(n-2,p)+pz_\gamma^+(n-2,p)}{z_\gamma(n-2,p)} &\geq \frac{(1-p)u_{\gamma_0}(n-2,p)+pu_{\gamma_0}^+(n-2,p)}{u_{\gamma_0}(n-2,p)} \\
1-p+\frac{pz_\gamma^+(n-2,p)}{z_\gamma(n-2,p)} &\geq 1-p+\frac{pu_{\gamma_0}^+(n-2,p)}{u_{\gamma_0}(n-2,p)} \\
\frac{z_\gamma^+(n-2,p)}{z_\gamma(n-2,p)} &\geq \frac{u_{\gamma_0}^+(n-2,p)}{u_{\gamma_0}(n-2,p)} \\
z_\gamma^+(n-2,p)u_{\gamma_0}(n-2,p) &\geq u_{\gamma_0}^+(n-2,p)z_\gamma(n-2,p).
\end{aligned}$$

This holds if and only if

$$\begin{aligned}
&\left(\sum_{\ell=0}^{n-2}b(\ell,n-2,p)\frac{\log(2+\ell)}{(2+\ell)^\gamma}\right)\left(\sum_{\ell=0}^{n-2}b(\ell,n-2,p)\frac{1}{(1+\ell)^{\gamma_0}}\right) \\
-&\left(\sum_{\ell=0}^{n-2}b(\ell,n-2,p)\frac{\log(1+\ell)}{(1+\ell)^\gamma}\right)\left(\sum_{\ell=0}^{n-2}b(\ell,n-2,p)\frac{1}{(2+\ell)^{\gamma_0}}\right)\geq 0.
\end{aligned}$$

Expanding terms, this is

$$\sum_{\ell_1=0}^{n-3}\sum_{\ell_2=\ell_1+1}^{n-2} b(\ell_1,n-2,p)b(\ell_2,n-2,p)\left[\frac{\log(2+\ell_1)}{(2+\ell_1)^{\gamma}}\cdot\frac{1}{(1+\ell_2)^{\gamma_0}}+\frac{\log(2+\ell_2)}{(2+\ell_2)^{\gamma}}\cdot\frac{1}{(1+\ell_1)^{\gamma_0}}\right.$$
$$\left.-\frac{\log(1+\ell_1)}{(1+\ell_1)^{\gamma}}\cdot\frac{1}{(2+\ell_2)^{\gamma_0}}-\frac{\log(1+\ell_2)}{(1+\ell_2)^{\gamma}}\cdot\frac{1}{(2+\ell_1)^{\gamma_0}}\right] \tag{11}$$

$$+\sum_{\ell=0}^{n-2} b(\ell,n-2,p)^2\left[\frac{\log(2+\ell)}{(2+\ell)^{\gamma}}\cdot\frac{1}{(1+\ell)^{\gamma_0}}+\frac{\log(2+\ell)}{(2+\ell)^{\gamma}}\cdot\frac{1}{(1+\ell)^{\gamma_0}}\right.$$
$$\left.-\frac{\log(1+\ell)}{(1+\ell)^{\gamma}}\cdot\frac{1}{(2+\ell)^{\gamma_0}}-\frac{\log(1+\ell)}{(1+\ell)^{\gamma}}\cdot\frac{1}{(2+\ell)^{\gamma_0}}\right]\geq 0. \tag{12}$$

We will show that the terms in brackets in (11) and (12) are both nonnegative. We begin with (11).

$$\begin{aligned}
&\frac{\log(2+\ell_1)}{(2+\ell_1)^{\gamma}}\cdot\frac{1}{(1+\ell_2)^{\gamma_0}}+\frac{\log(2+\ell_2)}{(2+\ell_2)^{\gamma}}\cdot\frac{1}{(1+\ell_1)^{\gamma_0}}\\
&-\frac{\log(1+\ell_1)}{(1+\ell_1)^{\gamma}}\cdot\frac{1}{(2+\ell_2)^{\gamma_0}}-\frac{\log(1+\ell_2)}{(1+\ell_2)^{\gamma}}\cdot\frac{1}{(2+\ell_1)^{\gamma_0}}\\
=&\frac{1}{(2+\ell_1)^{\gamma_0}(1+\ell_2)^{\gamma_0}}\left(\frac{\log(2+\ell_1)}{(2+\ell_1)^{\gamma-\gamma_0}}-\frac{\log(1+\ell_2)}{(1+\ell_2)^{\gamma-\gamma_0}}\right)\\
&+\frac{1}{(1+\ell_1)^{\gamma_0}(2+\ell_2)^{\gamma_0}}\left(\frac{\log(2+\ell_2)}{(2+\ell_2)^{\gamma-\gamma_0}}-\frac{\log(1+\ell_1)}{(1+\ell_1)^{\gamma-\gamma_0}}\right)
\end{aligned}$$

Let $t(\ell)\triangleq\frac{\log(\ell)}{\ell^{\gamma-\gamma_0}}$. Our goal is to show that the above is nonnegative, or

$$\begin{aligned}
\frac{1}{(1+\ell_1)^{\gamma_0}(2+\ell_2)^{\gamma_0}}\left(t(2+\ell_2)-t(1+\ell_1)\right)&\geq\frac{1}{(2+\ell_1)^{\gamma_0}(1+\ell_2)^{\gamma_0}}\left(t(1+\ell_2)-t(2+\ell_1)\right)\\
t(2+\ell_2)-t(1+\ell_1)&\geq\frac{((1+\ell_1)(2+\ell_2))^{\gamma_0}}{((2+\ell_1)(1+\ell_2))^{\gamma_0}}\left(t(1+\ell_2)-t(2+\ell_1)\right)
\end{aligned}$$

Because $\ell_1<\ell_2$,

$$\frac{((1+\ell_1)(2+\ell_2))^{\gamma_0}}{((2+\ell_1)(1+\ell_2))^{\gamma_0}}=\frac{(2+2\ell_1+\ell_2+\ell_1\ell_2)^{\gamma_0}}{(2+\ell_1+2\ell_2+\ell_1\ell_2)^{\gamma_0}}\leq 1.$$

It therefore suffices to show that for $\ell_2>\ell_1$,

$$\begin{aligned}
t(2+\ell_2)-t(1+\ell_1)&\geq t(1+\ell_2)-t(2+\ell_1)\\
t(2+\ell_1)+t(2+\ell_2)&\geq t(1+\ell_1)+t(1+\ell_2)
\end{aligned}\tag{13}$$

If t is weakly increasing, this is trivially true: each term on the left hand side is weakly larger than the corresponding term on the right hand side. While t is not increasing in general, it

is as long as its arguments are sufficiently small relative to $\gamma - \gamma_0$:

$$\begin{aligned}\frac{d}{d\ell}t(\ell) &\geq 0\\ \frac{\ell^{\gamma-\gamma_0}\ell^{-1} - \log(\ell)(\gamma-\gamma_0)\ell^{\gamma-\gamma_0-1}}{\ell^{2(\gamma-\gamma_0)}} &\geq 0\\ 1 - (\gamma-\gamma_0)\log(\ell) &\geq 0\\ (\gamma-\gamma_0) &\leq \frac{1}{\log(\ell)}\end{aligned}$$

Here, we make use of our assumption $\gamma_1 - \gamma_0 \leq \frac{1}{\log n}$. Because $\gamma \in [\gamma_0, \gamma_1]$, t is increasing as long as we plug in values that are at most n. The largest value ℓ_2 can take is $n-2$, so the largest value we plug in is $2+n-2=n$. Thus, (13) holds, and the first term in brackets (11) is nonnegative. We use a similar, though simpler, argument to handle the second term (12).

$$\begin{aligned}&\frac{\log(2+\ell)}{(2+\ell)^{\gamma}}\cdot\frac{1}{(1+\ell)^{\gamma_0}} + \frac{\log(2+\ell)}{(2+\ell)^{\gamma}}\cdot\frac{1}{(1+\ell)^{\gamma_0}}\\ &- \frac{\log(1+\ell)}{(1+\ell)^{\gamma}}\cdot\frac{1}{(2+\ell)^{\gamma_0}} - \frac{\log(1+\ell)}{(1+\ell)^{\gamma}}\cdot\frac{1}{(2+\ell)^{\gamma_0}}\\ =&\frac{2}{(2+\ell)^{\gamma_0}(1+\ell)^{\gamma_0}}\left(\frac{\log(2+\ell)}{(2+\ell)^{\gamma-\gamma_0}} - \frac{\log(1+\ell)}{(1+\ell)^{\gamma-\gamma_0}}\right)\\ &\frac{2}{(2+\ell)^{\gamma_0}(1+\ell)^{\gamma_0}}\left(t(2+\ell) - t(1+\ell)\right)\end{aligned}$$

Again, this is nonnegative as long as t is weakly increasing, which it is for $\ell \leq n-2$. With this, we have shown (6) under the assumption that $\gamma_1 \leq \gamma_0 + \frac{1}{\log n}$. Still under this assumption, we use Lemma A.4 with $\mathbf{p}' = \mathbf{p}^{(\gamma_0)}$, $\mathbf{p} = \mathbf{p}^{(\gamma_1)}$, and $Q_k = \mathbf{d}_k u_{\gamma_1}(n-1, \mathbf{p}_k^{(\gamma_0)})$. By (6), $\{Q_k\}$ is increasing as long as $\mathbf{p}_k^{(\gamma_0)} > 0$, so $\mathbf{p}^{(\gamma_0)} \succ \mathbf{p}^{(\gamma_1)}$.

To finish the proof, we must remove the assumption that $\gamma_1 \leq \gamma_0 + \frac{1}{\log n}$. For any $\gamma_1 \geq \gamma_0$, we simply need to iteratively apply this argument $\lceil(\gamma_1 - \gamma_0)\log n\rceil$ times to get

$$\mathbf{p}^{(\gamma_0)} \succ \mathbf{p}^{(\gamma_0 + 1/\log(n))} \succ \mathbf{p}^{(\gamma_0 + 2/\log(n))} \succ \cdots \succ \mathbf{p}^{(\gamma_1)}.$$

This suffices for any finite γ_1. To handle the case where $\gamma_1 = \infty$, we invoke Lemma 4.10 which states that for any $s \in \mathcal{S}$ (including s_{γ_0}), $\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s) \succ \mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s_\infty)$. □

Theorem 4.7 (Optimal strategies are more diverse than equilibria)**.** *For all $n \geq 2$, $\mathbf{d} \in \mathbb{R}^K_{\geq 0}$, $s \in \mathcal{S}$,*

$$\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s) \succ \mathbf{p}_{\text{OPT}}(n, \mathbf{d}, s).$$

Proof. Assume without loss of generality that $\mathbf{d}$ is decreasing. Let $\mathbf{p}^* = \mathbf{p}_{\text{OPT}}(n, \mathbf{d}, s)$. We begin with the case where $s \in \mathcal{S}_{\nearrow}$. By Lemma 4.2, for all k such that $\mathbf{p}^*_k > 0$,

$$\frac{\partial}{\partial p}\mathbf{d}_k p u(n-1, p)\Big|_{p=\mathbf{p}^*_k} = c$$

for $c = \frac{1}{n}\|\nabla W(\mathbf{p})\|_\infty$. Expanding,

$$\begin{aligned}
\frac{\partial}{\partial p}\mathbf{d}_k p u(n-1,p)\Big|_{p=\mathbf{p}_k^*} &= c \\
\mathbf{d}_k u(n-1,\mathbf{p}_k^*) + \mathbf{d}_k \mathbf{p}_k^* \frac{\partial}{\partial p} u(n-1,p)\Big|_{p=\mathbf{p}_k^*} &= c \\
\mathbf{d}_k u(n-1,\mathbf{p}_k^*) + \mathbf{d}_k \mathbf{p}_k^* \frac{n-1}{\mathbf{p}_k^*}\bar{u}(n-2,\mathbf{p}_k^*) &= c && \text{(by Lemma A.7)} \\
\mathbf{d}_k u(n-1,\mathbf{p}_k^*) &= c - \mathbf{d}_k(n-1)\bar{u}(n-2,\mathbf{p}_k^*) && (14)
\end{aligned}$$

We will show that the left hand side is decreasing in k: For $k < k'$, our goal is to show that

$$\mathbf{d}_k u(n-1,\mathbf{p}_k^*) \geq \mathbf{d}_{k'} u(n-1,\mathbf{p}_{k'}^*). \tag{15}$$

If $\mathbf{p}_k^* = 0$, then $\mathbf{p}_{k'}^* = 0$ as well. In this case, because $u(n-1,0) = 1$, this is simply $\mathbf{d}_k \geq \mathbf{d}_{k'}$, which is true by assumption. We can therefore focus on the case where $\mathbf{p}_k^* > 0$.

$$\begin{aligned}
& \mathbf{d}_k u(n-1,\mathbf{p}_k^*) \geq \mathbf{d}_{k'} u(n-1,\mathbf{p}_{k'}^*) \\
\iff\ & c - \mathbf{d}_k(n-1)\bar{u}(n-2,\mathbf{p}_k^*) \geq c - \mathbf{d}_{k'}(n-1)\bar{u}(n-2,\mathbf{p}_{k'}^*) && \text{(by (14))} \\
\iff\ & \mathbf{d}_k \bar{u}(n-2,\mathbf{p}_k^*) \leq \mathbf{d}_{k'}\bar{u}(n-2,\mathbf{p}_{k'}^*) && (16)
\end{aligned}$$

Next, observe that by Lemma 4.2,

$$\mathbf{d}_k u(n-1,\mathbf{p}_k^*) + \mathbf{d}_k(n-1)\bar{u}(n-2,\mathbf{p}_k^*) \geq \mathbf{d}_{k'} u(n-1,\mathbf{p}_{k'}^*) + \mathbf{d}_{k'}(n-1)\bar{u}(n-2,\mathbf{p}_{k'}^*). \tag{17}$$

This holds because by assumption, $\mathbf{p}_k^* > 0$, so the left hand side is c, and the right hand side is at most c. Note that both sides of (16) are weakly negative, and both sides of (17) are positive. Therefore, it suffices to show that the inequality holds if we divide both sides

of (16) by the respective sides of (17) to get

$$\begin{aligned}
&\frac{\mathbf{d}_k \bar{u}(n-2, \mathbf{p}_k^*)}{\mathbf{d}_k u(n-1, \mathbf{p}_k^*) + \mathbf{d}_k (n-1)\bar{u}(n-2, \mathbf{p}_k^*)} \leq \frac{\mathbf{d}_{k'} \bar{u}(n-2, \mathbf{p}_{k'}^*)}{\mathbf{d}_{k'} u(n-1, \mathbf{p}_{k'}^*) + \mathbf{d}_{k'} (n-1)\bar{u}(n-2, \mathbf{p}_{k'}^*)} \\
\iff &\frac{u(n-1, \mathbf{p}_k^*) + (n-1)\bar{u}(n-2, \mathbf{p}_k^*)}{\bar{u}(n-2, \mathbf{p}_k^*)} \geq \frac{u(n-1, \mathbf{p}_{k'}^*) + (n-1)\bar{u}(n-2, \mathbf{p}_{k'}^*)}{\bar{u}(n-2, \mathbf{p}_{k'}^*)} \\
\iff &\frac{u(n-1, \mathbf{p}_k^*)}{\bar{u}(n-2, \mathbf{p}_k^*)} \geq \frac{u(n-1, \mathbf{p}_{k'}^*)}{\bar{u}(n-2, \mathbf{p}_{k'}^*)} \\
\iff &\frac{\bar{u}(n-2, \mathbf{p}_k^*)}{u(n-1, \mathbf{p}_k^*)} \leq \frac{\bar{u}(n-2, \mathbf{p}_{k'}^*)}{u(n-1, \mathbf{p}_{k'}^*)} \\
\iff &\frac{u(n-1, \mathbf{p}_k^*) - u(n-2, \mathbf{p}_k^*)}{u(n-1, \mathbf{p}_k^*)} \leq \frac{u(n-1, \mathbf{p}_{k'}^*) - u(n-2, \mathbf{p}_{k'}^*)}{u(n-1, \mathbf{p}_{k'}^*)} \\
\iff &-\frac{u(n-2, \mathbf{p}_k^*)}{u(n-1, \mathbf{p}_k^*)} \leq -\frac{u(n-2, \mathbf{p}_{k'}^*)}{u(n-1, \mathbf{p}_{k'}^*)} \\
\iff &\frac{u(n-2, \mathbf{p}_k^*)}{u(n-1, \mathbf{p}_k^*)} \geq \frac{u(n-2, \mathbf{p}_{k'}^*)}{u(n-1, \mathbf{p}_{k'}^*)} \\
\iff &\frac{u(n-1, \mathbf{p}_k^*)}{u(n-2, \mathbf{p}_k^*)} \leq \frac{u(n-1, \mathbf{p}_{k'}^*)}{u(n-2, \mathbf{p}_{k'}^*)} \\
\impliedby &\frac{\partial}{\partial p} \frac{u(n-1, p)}{u(n-2, p)} \leq 0.
\end{aligned}$$

This last inequality holds by Lemma A.9. With this, we can apply Lemma A.4 with $\mathbf{p}' = \mathbf{p}^*$, $\mathbf{p} = \mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s)$, and $Q_k = \mathbf{d}_k u(n-1, \mathbf{p}_k^*)$. As we have shown, $\{Q_k\}$ is decreasing, so $\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s) \succ \mathbf{p}^*$.

Next, we handle the case where $s \in \mathcal{S}_\searrow$.

$$\begin{aligned}
\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s) &\succ \mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s_\infty) && \text{(Lemma 4.10)} \\
&= \mathbf{p}_{\text{OPT}}(n, \mathbf{d}, s), && \text{(Lemma 4.9)}
\end{aligned}$$

which proves the claim. □

Lemma 4.8 (Strategies as $n \to \infty$)**.** *Assume without loss of generality that $\mathbf{d} \in \mathbb{R}^K_{\geq 0}$ is weakly decreasing. Let $\mathbf{p}^{(n)} = \mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s_\gamma)$, and let $\mathbf{q}^{(n)} = \mathbf{p}_{\text{OPT}}(n, \mathbf{d}, s_\gamma)$. Let $\mathbf{p}^\dagger(\gamma)_k \triangleq \mathbf{d}_k^{1/\gamma} / \sum_{k' \in [K]} \mathbf{d}_{k'}^{1/\gamma}$. Then,*

- $\left\{\mathbf{p}^{(n)}\right\}_{n\to\infty}$ *approaches* $\mathbf{p}^\dagger(\gamma)$.
- *If* $\gamma < 1$, $\left\{\mathbf{q}^{(n)}\right\}_{n\to\infty}$ *approaches* $\mathbf{p}^\dagger(\gamma)$.
- *If* $\gamma \geq 1$ *(or more generally, for any* $s \in \mathcal{S}_\searrow$*),* $\left\{\mathbf{q}^{(n)}\right\}_{n\to\infty}$ *approaches the uniform distribution.*

Proof. Intuitively, we will use the fact that for large n, $X(n, p)$ concentrates to its expectation np. We then use the equilibrium conditions in Lemma 4.1 to show that $\mathbf{d}_k / (n\mathbf{p}_k^{(n)})^\gamma$ must be equal for all k, which yields the desired results for $\mathbf{p}^\dagger(\gamma)$.

Assume $\mathbf{d}_K > 0$ (otherwise $\mathbf{p}_K^{(n)} = 0$ as well). Let $c^{(n)} = U(\mathbf{p}^{(n)}; n, \mathbf{d}, s)$. By Lemma 4.1, for all k such that $\mathbf{p}_k^{(n)} > 0$,

$$\mathbf{d}_k u(n-1, \mathbf{p}_k^{(n)}) = \mathbf{d}_1 u(n-1, \mathbf{p}_1^{(n)}) = c^{(n)}.$$

Moreover, because $\mathbf{p}_1^{(n)} \geq \frac{1}{K}$, $c^{(n)} \leq \mathbf{d}_k u(n-1, 1/K)$.

$$\begin{aligned}
\lim_{n\to\infty} u(n-1, 1/K) &= \lim_{n\to\infty} \mathbb{E}\left[\frac{1}{s_\gamma(1 + X(n-1, 1/K))}\right] \\
&= \lim_{n\to\infty} \mathbb{E}\left[\frac{1}{(1 + X(n-1, 1/K))^\gamma}\right] \\
&= 0
\end{aligned}$$

because for any constant x, $\lim_{n\to\infty} \Pr[X(n-1, 1/K) > x] = 1$. As a result, $\lim_{n\to\infty} c^{(n)} = 0$, so for any k, $\mathbf{d}_k > c^{(n)}$ for sufficiently large n. This means that for sufficiently large n, $\mathbf{p}_k^{(n)} > 0$ for all k. We will now assume that this is the case moving forwards. Next, we will show that there is some constant $a > 0$ such that for sufficiently large n, $\mathbf{p}_k^{(n)} \geq a$ for all k. In particular, let n_0 be the smallest n such that $\mathbf{p}_K^{(n)} > 0$. Let $a = \mathbf{p}_K^{(n_0)}$. By Theorem 4.5, $\mathbf{p}_k^{(n)} \succ \mathbf{p}_k^{(n+1)}$ for all n. This implies that $\mathbf{p}_K^{(n)}$ must be monotonically increasing in n, so for all $n \geq n_0$, $\mathbf{p}_K^{(n)} \geq a$. As a result, $\mathbf{p}_k^{(n)} \geq a$ for all $n \geq n_0$, since $\mathbf{p}^{(n)} \in \Delta(\pi)$.

We now use concentration to bound $u(n, p)$:

$$\begin{aligned}
u(n,p) &= \mathbb{E}\left[(1 + X(n,p))^{-\gamma}\right] \\
&= \mathbb{E}\left[(1 + X(n,p))^{-\gamma} \;\middle|\; |X(n,p) - np| \leq (1+\delta)np\right] && (18)\\
&\cdot \Pr\left[|X(n,p) - np| \leq (1+\delta)np\right] && (19)\\
&+ \mathbb{E}\left[(1 + X(n,p))^{-\gamma} \;\middle|\; |X(n,p) - np| > (1+\delta)np\right] && (20)\\
&\cdot \Pr\left[|X(n,p) - np| > (1+\delta)np\right] && (21)
\end{aligned}$$

We bound each of these individually. First, (18):

$$\mathbb{E}\left[(1 + X(n,p))^{-\gamma} \;\middle|\; |X(n,p) - np| \leq (1+\delta)np\right] = (1 + np + \alpha_1)^{-\gamma}$$

for some $\alpha_1 \in [-\delta np, \delta np]$ by the mean value theorem. Using Taylor's theorem,

$$\begin{aligned}
(1 + np + \alpha_1)^{-\gamma} &= (1+np)^{-\gamma} + \alpha_1 \frac{d}{dx} x^{-\gamma}\Big|_{x=1+np+\alpha_2} \\
&= (1+np)^{-\gamma} - \alpha_1 \gamma (1 + np + \alpha_2)^{-\gamma-1}
\end{aligned}$$

for some α_2 between 0 and α_1. We can therefore bound $\alpha_2 \geq \min(0, \alpha_1)$. Putting this together,

$$\begin{aligned}
(1+np)^{-\gamma} - \delta np\gamma(1 + np(1-\delta))^{-\gamma-1} &\leq \mathbb{E}\left[(1 + X(n,p))^{-\gamma} \;\middle|\; |X(n,p) - np| \leq (1+\delta)np\right] \\
&\leq (1+np)^{-\gamma} + \delta np\gamma(1+np)^{-\gamma-1}
\end{aligned}$$

Next, to bound (19) and (21), define

$$\xi(\delta, z) \triangleq 2\exp\left(-\frac{\delta^2 z}{3}\right).$$

Standard Chernoff bounds yield

$$\begin{aligned} 0 \leq \Pr[|X(n,p) - np| > (1+\delta)np] &\leq \xi(\delta, np) \\ 1 - \xi(\delta, np) \leq \Pr[|X(n,p) - np| \leq (1+\delta)np] &\leq 1. \end{aligned}$$

Finally, observe that (20) is bounded between 0 and 1. Putting this all together, we have that for any $\delta > 0$, for sufficiently large n,

$$\begin{aligned} &\left((1+np)^{-\gamma} - \delta np\gamma(1+np(1-\delta)^{-\gamma-1}\right)(1-\xi(\delta,np)) \\ &\leq u(n,p) \\ &\leq (1+np)^{-\gamma} + \delta np\gamma(1+np)^{-\gamma-1} + \xi(\delta,np). \end{aligned} \tag{22}$$

We will use the squeeze theorem to show that

$$\lim_{n\to\infty} u(n-1, \mathbf{p}_k^{(n)})(1+(n-1)\mathbf{p}_k^{(n)})^{\gamma} = 1. \tag{23}$$

Let $h_n = (n-1)\mathbf{p}_k^{(n)}$, and let $\delta_n = 1/\log(h_n)$. Recall that for sufficiently large n, $\mathbf{p}_k^{(n)} \geq a$, so $a(n-1) \leq h_n \leq n-1$. Taking the left hand side of (22),

$$\begin{aligned} &\lim_{n\to\infty} \left((1+h_n)^{-\gamma} - \delta_n h_n\gamma(1+h_n(1-\delta_n))^{-\gamma-1}\right)(1-\xi(\delta_n,h_n))(1+h_n)^{\gamma} \\ &= \lim_{n\to\infty} \left(1 - \frac{\gamma}{\log h_n}\left(\frac{1+h_n}{1+h_n\left(1-\frac{1}{\log h_n}\right)}\right)^{\gamma}\left(\frac{h_n}{1+h_n\left(1-\frac{1}{\log h_n}\right)}\right)\right)\left(1-2\exp\left(-\frac{\frac{h_n}{(\log h_n)^2}}{3}\right)\right) \\ &= 1 + \lim_{n\to\infty} -\frac{\gamma}{\log h_n}\left(1-\frac{1}{(\log h_n)(1+h_n)-1}\right)^{\gamma}\left(\frac{1}{\frac{1}{h_n}+1-\frac{1}{\log h_n}}\right) \\ &= 1. \end{aligned}$$

In the last step, we implicitly applied the squeeze theorem with $a(n-1) \leq h_n \leq n-1$. The right hand side of (22) becomes

$$\begin{aligned} &\lim_{n\to\infty} \left((1+h_n)^{-\gamma} + \delta_n h_n\gamma(1+h_n)^{-\gamma-1} + \xi(\delta_n,h_n)\right)(1+h_n)^{\gamma} \\ &= \lim_{n\to\infty} 1 + \frac{\gamma}{\log h_n}\left(\frac{h_n}{1+h_n}\right) + 2\exp\left(-\frac{\frac{h_n}{(\log h_n)^2}}{3} + \gamma\log(1+h_n)\right) \\ &= 1. \end{aligned}$$

Therefore, (23) holds.

Consider any k, k'. At equilibrium,

$$\begin{aligned}
\mathbf{d}_k u(n-1, \mathbf{p}_k^{(n)}) &= \mathbf{d}_{k'} u(n-1, \mathbf{p}_{k'}^{(n)}) \\
\frac{\mathbf{d}_k}{\mathbf{d}_{k'}} &= \frac{u(n-1, \mathbf{p}_{k'}^{(n)})}{u(n-1, \mathbf{p}_k^{(n)})} \\
\frac{\mathbf{d}_k s_\gamma((n-1)\mathbf{p}_{k'}^{(n)})}{\mathbf{d}_{k'} s_\gamma((n-1)\mathbf{p}_k^{(n)})} &= \frac{u(n-1, \mathbf{p}_{k'}^{(n)}) s_\gamma((n-1)\mathbf{p}_{k'}^{(n)})}{u(n-1, \mathbf{p}_k^{(n)}) s_\gamma((n-1)\mathbf{p}_k^{(n)})}.
\end{aligned}$$

Therefore, by (23),

$$\begin{aligned}
\lim_{n\to\infty} \frac{\mathbf{d}_k s_\gamma((n-1)\mathbf{p}_{k'}^{(n)})}{\mathbf{d}_{k'} s_\gamma((n-1)\mathbf{p}_k^{(n)})} &= 1 \\
\lim_{n\to\infty} \frac{\mathbf{d}_k ((n-1)\mathbf{p}_{k'}^{(n)})^\gamma}{\mathbf{d}_{k'} ((n-1)\mathbf{p}_k^{(n)})^\gamma} &= 1 \\
\lim_{n\to\infty} \left(\frac{\mathbf{p}_{k'}^{(n)}}{\mathbf{p}_k^{(n)}}\right)^\gamma &= \frac{\mathbf{d}'_k}{\mathbf{d}_k} \\
\lim_{n\to\infty} \frac{\mathbf{p}_{k'}^{(n)}}{\mathbf{p}_k^{(n)}} &= \left(\frac{\mathbf{d}'_k}{\mathbf{d}_k}\right)^{1/\gamma}.
\end{aligned}$$

Since this holds for all k, k' and $\sum_k \mathbf{p}_k^{(n)} = 1$,

$$\lim_{n\to\infty} \mathbf{p}_k^{(n)} = \frac{\mathbf{d}_k^{1/\gamma}}{\sum_{k' \in [K]} \mathbf{d}_{k'}^{1/\gamma}}.$$

Technically, we have only proved the claim for finite γ; we will return to $\gamma = \infty$ later.

Next, we use a similar argument for $\mathbf{q}_k^{(n)} = \mathbf{p}_{\mathrm{OPT}}(n, \mathbf{d}, s_\gamma)$. We begin with the case where $\gamma < 1$, so $s_\gamma \in \mathcal{S}_\nearrow$. By (4.2), for all k such that $\mathbf{q}_k^{(n)} > 0$,

$$\begin{aligned}
\frac{\mathbf{d}_k}{n} \frac{\partial}{\partial p} w(n,p)\Big|_{p=\mathbf{q}_k^{(n)}} &= \frac{\mathbf{d}_1}{n} \frac{\partial}{\partial p} w(n,p)\Big|_{p=\mathbf{q}_1^{(n)}} \\
\frac{\mathbf{d}_k}{\mathbf{q}_k^{(n)}} \bar{w}(n-1, \mathbf{q}_k^{(n)}) &= \frac{\mathbf{d}_1}{\mathbf{q}_1^{(n)}} \bar{w}(n-1, \mathbf{q}_1^{(n)}). && \text{(by Lemma A.7)}
\end{aligned}$$

Observe that

$$\begin{aligned}
\frac{1}{p}\bar{w}(n-1,p) &= \frac{1}{p}(w(n,p) - w(n-1,p)) \\
&= \frac{1}{p}\left(\mathbb{E}\left[\frac{X(n,p)}{s_\gamma(X(n,p))}\right] - \mathbb{E}\left[\frac{X(n-1,p)}{s_\gamma(X(n-1,p))}\right]\right) \\
&= \frac{1}{p}\left(\mathbb{E}\left[X(n,p)^{1-\gamma}\right] - \mathbb{E}\left[X(n-1,p)^{1-\gamma}\right]\right) \\
&= \frac{1}{p}\left(p\mathbb{E}\left[(1+X(n-1,p))^{1-\gamma}\right] + (1-p)\mathbb{E}\left[X(n-1,p)^{1-\gamma}\right] - \mathbb{E}\left[X(n-1,p)^{1-\gamma}\right]\right) \\
&= \mathbb{E}\left[(1+X(n-1,p))^{1-\gamma} - X(n-1,p)^{1-\gamma}\right].
\end{aligned}$$

We again use the fact that $X(n,p)$ concentrates.

$$\begin{aligned}
&\mathbb{E}\left[(1+X(n,p))^{1-\gamma} - X(n,p)^{1-\gamma}\right] \\
&= \mathbb{E}\left[(1+X(n,p))^{1-\gamma} - X(n,p)^{1-\gamma} \mid |X(n,p)-np| \le (1+\delta)np\right] \\
&\cdot \Pr[|X(n,p)-np| \le (1+\delta)np] \\
&+ \mathbb{E}\left[(1+X(n,p))^{1-\gamma} - X(n,p)^{1-\gamma} \mid |X(n,p)-np| > (1+\delta)np\right] \\
&\cdot \Pr[|X(n,p)-np| > (1+\delta)np]
\end{aligned}$$

The first term is

$$\begin{aligned}
&\mathbb{E}\left[(1+X(n,p))^{1-\gamma} - X(n,p)^{1-\gamma} \mid |X(n,p)-np| \le (1+\delta)np\right] \\
&= (1+np+\alpha_3)^{1-\gamma} - (np+\alpha_3)^{1-\gamma} \qquad \text{(Mean value theorem, } \alpha_3 \in [-\delta np, \delta np]) \\
&= (1+np)^{1-\gamma} - (np)^{1-\gamma} + \alpha_3 \frac{d}{dx}(1+x)^{1-\gamma} - x^{1-\gamma}\Big|_{x=np+\alpha_4} \\
&\qquad\qquad \text{(Taylor's theorem, } \alpha_4 \text{ between 0 and } \alpha_3) \\
&= (1+np)^{1-\gamma} - (np)^{1-\gamma} + \alpha_3(1-\gamma)((1+np+\alpha_4)^{-\gamma} - (np+\alpha_4)^{-\gamma}).
\end{aligned}$$

Next, observe that again using Taylor's theorem,

$$\begin{aligned}
(1+y)^z - y^z &= \frac{d}{dx}x^z\Big|_{y+\beta} \\
&= z(y+\beta)^{z-1}
\end{aligned}$$

for some $\beta \in [0,1]$. Therefore,

$$\begin{aligned}
&\mathbb{E}\left[(1+X(n,p))^{1-\gamma} - X(n,p)^{1-\gamma} \mid |X(n,p)-np| \le (1+\delta)np\right] \\
&= (1-\gamma)(np+\beta_1)^{-\gamma} + \alpha_3(1-\gamma)(-\gamma)(np+\alpha_4+\beta_2)^{-\gamma-1} \\
&= (1-\gamma)\left((np+\beta_1)^{-\gamma} - \alpha_3\gamma(np+\alpha_4+\beta_2)^{-\gamma-1}\right).
\end{aligned}$$

for $\beta_1, \beta_2 \in [0,1]$. Using the fact that $\alpha_4 \ge \min(0, \alpha_3)$, this yields

$$\begin{aligned}
&(1-\gamma)\left((np+1)^{-\gamma} - \delta np\gamma(np(1-\delta))^{-\gamma-1}\right) \\
&\le \mathbb{E}\left[(1+X(n,p))^{1-\gamma} - X(n,p)^{1-\gamma} \mid |X(n,p)-np| \le (1+\delta)np\right] \\
&\le (1-\gamma)\left((np)^{-\gamma} + \delta np\gamma(np+1)^{-\gamma-1}\right).
\end{aligned}$$

Combining with Chernoff bounds, we have

$$\begin{aligned}
&(1-\gamma)\left((np+1)^{-\gamma} - \delta np\gamma(np(1-\delta))^{-\gamma-1}\right)(1-\xi(\delta,np)) \\
&\le \frac{1}{p}\bar{w}(n,p) \\
&\le (1-\gamma)\left((np)^{-\gamma} + \delta np\gamma(np+1)^{-\gamma-1}\right) + \xi(\delta,np). \qquad (24)
\end{aligned}$$

Again, we will use the squeeze theorem to prove that

$$\lim_{n\to\infty} \frac{1}{\mathbf{q}_k^{(n)}} \bar{w}(n-1, \mathbf{q}_k^{(n)}) \cdot \frac{((n-1)\mathbf{q}_k^{(n)})^\gamma}{1-\gamma} = 1. \qquad (25)$$

By Theorem 4.5, $\mathbf{p}^{(n)} \succ \mathbf{q}^{(n)}$ for all n. As noted above, there exists some $a > 0$ such that for sufficiently large n, $\mathbf{p}_k^{(n)} \geq a$ for all k. Therefore, for sufficiently large n and any k, $\mathbf{q}_k^{(n)} \geq \mathbf{q}_K^{(n)} \geq \mathbf{p}_K^{(n)} \geq a$. Again, let $h_n = (n-1)\mathbf{q}_k^{(n)}$ and $\delta_n = 1/\log h_n$, so $a(n-1) \leq h_n \leq n-1$. We begin with the left hand side of (24), which we multiply by $h_n^\gamma/(1-\gamma)$.

$$\begin{aligned}
&\lim_{n\to\infty} (1-\gamma)\left((h_n+1)^{-\gamma} - \delta_n h_n \gamma (h_n(1-\delta_n))^{-\gamma-1}\right)(1-\xi(\delta_n,h_n))\frac{h_n^\gamma}{1-\gamma} \\
&= \lim_{n\to\infty} \left(\left(\frac{h_n}{1+h_n}\right)^\gamma - \delta_n h_n^{\gamma+1}\gamma(h_n(1-\delta_n))^{-\gamma-1}\right)\left(1-2\exp\left(-\frac{\frac{h_n}{(\log h_n)^2}}{3}\right)\right) \\
&= \lim_{n\to\infty} \left(\left(1-\frac{1}{1+h_n}\right)^\gamma - \frac{\gamma}{\log h_n}\left(1-\frac{1}{\log h_n}\right)^{-\gamma-1}\right)\left(1-2\exp\left(-\frac{\frac{h_n}{(\log h_n)^2}}{3}\right)\right) \\
&= 1.
\end{aligned}$$

The right hand side of (24) multiplied by $h_n^\gamma/(1-\gamma)$ is

$$\begin{aligned}
&\lim_{n\to\infty} \left((1-\gamma)\left((h_n)^{-\gamma} + \delta_n h_n \gamma (h_n+1)^{-\gamma-1}\right) + \xi(\delta_n,h_n)\right)\frac{h_n^\gamma}{1-\gamma} \\
&= \lim_{n\to\infty} 1 + \frac{\gamma h_n^{\gamma+1}}{\log h_n}(h_n+1)^{-\gamma-1} + 2\exp\left(-\frac{\frac{h_n}{(\log h_n)^2}}{3}\right)\frac{h_n^\gamma}{1-\gamma} \\
&= 1 + \lim_{n\to\infty} \frac{\gamma}{\log h_n}\left(\frac{h_n}{1+h_n}\right)^{-\gamma-1} + \frac{2}{1-\gamma}\exp\left(-\frac{\frac{h_n}{(\log h_n)^2}}{3} + \gamma\log h_n\right) \\
&= 1.
\end{aligned}$$

This proves (25). We now use a similar argument to the one we used for $\mathbf{p}_k^{(n)}$. For any k, k',

$$\begin{aligned}
\frac{\mathbf{d}_k}{\mathbf{q}_k^{(n)}}\bar{w}(n-1,\mathbf{q}_k^{(n)}) &= \frac{\mathbf{d}_{k'}}{\mathbf{q}_{k'}^{(n)}}\bar{w}(n-1,\mathbf{q}_{k'}^{(n)}) \\
\frac{\mathbf{d}_k((n-1)\mathbf{q}_{k'}^{(n)})^\gamma}{\mathbf{d}'_k((n-1)\mathbf{q}_k^{(n)})^\gamma} &= \frac{\frac{1}{\mathbf{q}_{k'}^{(n)}}\bar{w}(n-1,\mathbf{q}_{k'}^{(n)})((n-1)\mathbf{q}_{k'}^{(n)})^\gamma}{\frac{1}{\mathbf{q}_k^{(n)}}\bar{w}(n-1,\mathbf{q}_k^{(n)})((n-1)\mathbf{q}_k^{(n)})^\gamma} \\
\lim_{n\to\infty}\frac{\mathbf{d}_k((n-1)\mathbf{q}_{k'}^{(n)})^\gamma}{\mathbf{d}'_k((n-1)\mathbf{q}_k^{(n)})^\gamma} &= \lim_{n\to\infty}\frac{\frac{1}{\mathbf{q}_{k'}^{(n)}}\bar{w}(n-1,\mathbf{q}_{k'}^{(n)})((n-1)\mathbf{q}_{k'}^{(n)})^\gamma}{\frac{1}{\mathbf{q}_k^{(n)}}\bar{w}(n-1,\mathbf{q}_k^{(n)})((n-1)\mathbf{q}_k^{(n)})^\gamma} \\
\frac{\mathbf{d}_k}{\mathbf{d}_{k'}}\lim_{n\to\infty}\left(\frac{\mathbf{q}_{k'}^{(n)}}{\mathbf{q}_k^{(n)}}\right)^\gamma &= 1 \\
\mathbf{q}_k^{(\infty)} &\propto \mathbf{d}_k^{1/\gamma}.
\end{aligned}$$

Finally, we handle the case where $\gamma \geq 1$ (and more generally, any $s \in \mathcal{S}_\searrow$). By Lemma 4.2,

for any $s \in \mathcal{S}_\searrow$,

$$\begin{aligned}\frac{1}{n}\frac{\partial}{\partial p}w(n,p) &= \frac{1}{n}\frac{\partial}{\partial p}\Pr[X(n,p) > 0] \\ &= \frac{1}{n}\frac{\partial}{\partial p}1-(1-p)^n \\ &= (1-p)^{n-1}.\end{aligned}$$

Then, for any k, k',

$$\begin{aligned}\mathbf{d}_k(1-\mathbf{q}_k^{(n)})^{n-1} &= \mathbf{d}_{k'}(1-\mathbf{q}_{k'}^{(n)})^{n-1} \\ \frac{1-\mathbf{q}_k^{(n)}}{1-\mathbf{q}_{k'}^{(n)}} &= \left(\frac{\mathbf{d}_{k'}}{\mathbf{d}_k}\right)^{1/(n-1)} \\ \lim_{n\to\infty}\frac{1-\mathbf{q}_k^{(n)}}{1-\mathbf{q}_{k'}^{(n)}} &= \lim_{n\to\infty}\left(\frac{\mathbf{d}_{k'}}{\mathbf{d}_k}\right)^{1/(n-1)} \\ \frac{1-\mathbf{q}_k^{(\infty)}}{1-\mathbf{q}_{k'}^{(\infty)}} &= 1 \\ \mathbf{q}_k^{(\infty)} &= \frac{1}{K}. \qquad \text{(because this must hold for all } k, k')\end{aligned}$$

Finally, we can use this to return to the case where $\gamma = \infty$ for $\mathbf{p}_k^{(\infty)}$. By Lemma 4.9, when $\gamma = \infty$, $\mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s_\infty) = \mathbf{p}_{\text{OPT}}(n, \mathbf{d}, s_\infty)$. Thus, in this case, $\mathbf{p}_k^{(\infty)} = 1/K \propto \mathbf{d}^{1/\infty}$ as well. □

Lemma 4.9 (Equivalence when $\gamma = \infty$)**.** *For all $n \geq 2$, $\mathbf{d} \in \mathbb{R}^K_{\geq 0}$, $s \in \mathcal{S}_\searrow$,*

$$\mathbf{p}_{\text{OPT}}(n, \mathbf{d}, s) = \mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s_\infty).$$

Proof. Assume without loss of generality that $\mathbf{d}$ is decreasing. Let $\mathbf{p}^* = \mathbf{p}_{\text{OPT}}(n, \mathbf{d}, s)$ and $\mathbf{p} = \mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s_\infty)$. By Lemma 4.2, for $s \in \mathcal{S}_\searrow$,

$$\mathbf{p}_k^* > 0 \iff \mathbf{d}_k(1-\mathbf{p}_k^*)^{n-1} = c$$

for $c = \frac{1}{n}\|\nabla W(\mathbf{p}^*)\|_\infty$. In addition,

$$\mathbf{p}_k^* = 0 \iff \mathbf{d}_k \leq c.$$

By Lemma 4.1,

$$\mathbf{p}_k > 0 \iff \mathbf{d}_k u_{s_\infty}(n-1, \mathbf{p}_k) = c' \tag{26}$$

for $c' = U(\mathbf{p}; n, \mathbf{d}, s_\infty)$. In addition,

$$\mathbf{p}_k = 0 \iff \mathbf{d}_k \leq c'.$$

Observe that

$$\begin{aligned} u_{s_\infty}(n-1,\mathbf{p}_k) &= \mathbb{E}\left[\frac{1}{(1+X(n-1,\mathbf{p}_k))^\infty}\right] \\ &= \Pr[X(n-1,\mathbf{p}_k)=0] \\ &= (1-\mathbf{p}_k)^{n-1}. \end{aligned}$$

Substituting into (26),

$$\mathbf{p}_k > 0 \iff \mathbf{d}_k(1-\mathbf{p}_k)^{n-1} = c'.$$

Because $(1-p)^{n-1}$ is strictly decreasing in p and $\sum_{k\in[K]}\mathbf{p}^*_k = \sum_{k\in[K]}\mathbf{p}_k = 1$, by a water-filling argument, it must be the case that $c' = c$. Therefore, $\mathbf{p}^* = \mathbf{p}$. □

Lemma 4.10 ($\gamma = \infty$ yields the most diverse equilibria)**.** *For all $n \geq 2$, $\mathbf{d} \in \mathbb{R}^K_{\geq 0}$, $s \in \mathcal{S}$,*

$$\mathbf{p}_{\text{EQ}}(n,\mathbf{d},s) \succ \mathbf{p}_{\text{EQ}}(n,\mathbf{d},s_\infty).$$

Proof. Assume without loss of generality that $\mathbf{d}$ is decreasing. Let $\mathbf{p} = \mathbf{p}_{\text{EQ}}(n,\mathbf{d},s)$ and let $\mathbf{q} = \mathbf{p}_{\text{EQ}}(n,\mathbf{d},s_\infty)$. Our goal will be to show that for $k' > k$, if $\mathbf{p}_{k'} > 0$,

$$\mathbf{d}_k u_{s_\infty}(n-1,\mathbf{p}_k) \leq \mathbf{d}_{k'} u_{s_\infty}(n-1,\mathbf{p}_{k'}). \tag{27}$$

We will then apply Lemma A.4. To do so, observe that under these conditions,

$$\mathbf{d}_k u_s(n-1,\mathbf{p}_k) = \mathbf{d}_{k'} u_s(n-1,\mathbf{p}_{k'}) \tag{28}$$

by Lemma 4.1. Dividing both sides of (27) by the respective sides of (28), it suffices to show that

$$\begin{aligned} \frac{u_{s_\infty}(n-1,\mathbf{p}_k)}{u_s(n-1,\mathbf{p}_k)} \leq \frac{u_{s_\infty}(n-1,\mathbf{p}_{k'})}{u_s(n-1,\mathbf{p}_{k'})} &\iff \frac{\Pr[X(n-1,\mathbf{p}_k)=0]}{u_s(n-1,\mathbf{p}_k)} \leq \frac{\Pr[X(n-1,\mathbf{p}_{k'})=0]}{u_s(n-1,\mathbf{p}_{k'})} \\ &\impliedby \frac{\partial}{\partial p}\frac{\Pr[X(n-1,p)=0]}{u_s(n-1,p)} \leq 0 \\ &\iff \frac{\partial}{\partial p}\log\left(\frac{\Pr[X(n-1,p)=0]}{u_s(n-1,p)}\right) \leq 0 \\ &\iff \frac{\frac{\partial}{\partial p}\Pr[X(n-1,p)=0]}{\Pr[X(n-1,p)=0]} \leq \frac{\frac{\partial}{\partial p}u_s(n-1,p)}{u_s(n-1,p)} \\ &\iff \frac{-(n-1)(1-p)^{n-2}}{(1-p)^{n-1}} \leq \frac{\frac{n-1}{p}\bar{u}_s(n-2,p)}{u_s(n-1,p)} \\ &\iff -\frac{1}{1-p} \leq \frac{1}{p}\frac{u_s(n-1,p)-u_s(n-2,p)}{u_s(n-1,p)} \\ &\iff -\frac{p}{1-p} \leq 1 - \frac{u_s(n-2,p)}{u_s(n-1,p)} \\ &\iff \frac{u_s(n-2,p)}{u_s(n-1,p)} \leq \frac{1}{1-p} \\ &\iff (1-p)u_s(n-2,p) \leq u_s(n-1,p). \end{aligned}$$

Observe that

$$\begin{aligned} u_s(n-1,p) &= \mathbb{E}\left[\frac{1}{s(1+X(n-1,p))}\right] \\ &= (1-p)\mathbb{E}\left[\frac{1}{s(1+x(n-2,p))}\right] + p\mathbb{E}\left[\frac{1}{s(2+X(n-2,p))}\right] \\ &\geq (1-p)\mathbb{E}\left[\frac{1}{s(1+x(n-2,p))}\right] \\ &= (1-p)u_s(n-2,p) \end{aligned}$$

as desired. Therefore, if $k' > k$ and $p_{k'} > 0$, (27) holds. Applying Lemma A.4 with $Q_k = \mathbf{d}_k u_{s_\infty}(n-1, \mathbf{p}_k)$, because $\{Q_k\}$ is increasing when $\mathbf{p}_k > 0$, $\mathbf{p} \succ \mathbf{q}$. □

A.2.1 Additional results on diversity.

Lemma A.4. *Let $\mathbf{p} = \mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s)$. Let $\mathbf{p}'$ be some other distribution. Consider the sequence $\{Q_k\}_{k\in[K]}$ where*

$$Q_k \triangleq \mathbf{d}_k u_s(n-1, \mathbf{p}'_k).$$

If $\{Q_k\}$ is decreasing, then $\mathbf{p} \succ \mathbf{p}'$. If $\{Q_k\}$ is increasing for k such that $\mathbf{p}'_k > 0$, then $\mathbf{p}' \succ \mathbf{p}$.

Proof. Assume without loss of generality that $\mathbf{d}$ is decreasing. Because $u(n-1,\cdot)$ is strictly decreasing,

$$\mathbf{p}'_k > \mathbf{p}_k \iff \mathbf{d}_k u(n-1, \mathbf{p}'_k) < \mathbf{d}_k u(n-1, \mathbf{p}_k). \tag{29}$$

Let $c = U(\mathbf{p}; n, \mathbf{d}, s)$.

Case 1: $\{Q_k\}$ is increasing. Consider any $k, \hat{k}$ such that $k < \hat{k}$. If $\mathbf{p}_k > 0$, then by Lemma 4.1,

$$\mathbf{d}_k u(n-1, \mathbf{p}_k) = c \geq \mathbf{d}_{\hat{k}} u(n-1, \mathbf{p}_{\hat{k}}). \tag{30}$$

If $\mathbf{p}_k = 0$, then $\mathbf{p}_{\hat{k}} = 0$ as well, so

$$\mathbf{d}_k u(n-1, \mathbf{p}_k) = \mathbf{d}_k \geq \mathbf{d}_{\hat{k}} \geq \mathbf{d}_{\hat{k}} u(n-1, \mathbf{p}_{\hat{k}}).$$

In either case,

$$\mathbf{d}_k u(n-1, \mathbf{p}_k) \geq \mathbf{d}_{\hat{k}} u(n-1, \mathbf{p}_{\hat{k}}).$$

Therefore, if $\mathbf{p}'_{\hat{k}} > \mathbf{p}_{\hat{k}}$, then

$$\begin{aligned} Q_k &\leq Q_{\hat{k}} && (\{Q_k\} \text{ increasing for } \mathbf{p}'_{\hat{k}} > 0) \\ &= \mathbf{d}_{\hat{k}} u(n-1, \mathbf{p}'_{\hat{k}}) \\ &< \mathbf{d}_{\hat{k}} u(n-1, \mathbf{p}_{\hat{k}}) && (\text{by (29)}) \\ &\leq \mathbf{d}_k u(n-1, \mathbf{p}_k). && (\text{by (30)}) \end{aligned}$$

Again using (29), $Q_k < \mathbf{d}_k u(n-1, \mathbf{p}_k)$ implies $\mathbf{p}'_k > \mathbf{p}_k$. As a result, we have

$$\mathbf{p}'_{\hat{k}} > \mathbf{p}_{\hat{k}} \Longrightarrow \mathbf{p}'_k > \mathbf{p}_k$$

for any $k < \hat{k}$. Let k^* be the largest k such that $\mathbf{p}'_k > \mathbf{p}_k$. We have shown that for $k \leq k^*$, $\mathbf{p}'_k > \mathbf{p}_k$. On the other hand, for $k > k^*$, $\mathbf{p}'_k = 0$, so $\mathbf{p}'_k \leq \mathbf{p}_k$. Therefore, $\mathbf{p}'_k > \mathbf{p}_k \Longleftrightarrow k \leq k^*$. This implies $\mathbf{p}' \succ \mathbf{p}$ by Lemma A.5.

Case 2: $\{Q_k\}$ is decreasing. Similarly, we will show that for any $k < \hat{k}$,

$$\mathbf{p}_{\hat{k}} > \mathbf{p}'_{\hat{k}} \Longrightarrow \mathbf{p}_k > \mathbf{p}'_k.$$

The premise implies that $\mathbf{p}_{\hat{k}} > 0$ and therefore $\mathbf{p}_k > 0$. By Lemma 4.1, this means

$$\mathbf{d}_k u(n-1, \mathbf{p}_k) = \mathbf{d}_{\hat{k}} u(n-1, \mathbf{p}_{\hat{k}}) = c$$

for some c. This means that

$$\begin{aligned}
\mathbf{p}_{\hat{k}} > \mathbf{p}'_{\hat{k}} &\Longleftrightarrow \mathbf{d}_{\hat{k}} u(n-1, \mathbf{p}_{\hat{k}}) < \mathbf{d}_{\hat{k}} u(n-1, \mathbf{p}'_{\hat{k}}) && \text{(by (29))} \\
&\Longleftrightarrow c < Q_{\hat{k}} \\
&\Longrightarrow c < Q_k && (\{Q_k\} \text{ is decreasing}) \\
&\Longleftrightarrow \mathbf{d}_k u(n-1, \mathbf{p}_k) < \mathbf{d}_k u(n-1, \mathbf{p}'_k) \\
&\Longleftrightarrow \mathbf{p}_k > \mathbf{p}'_k. && \text{(by (29))}
\end{aligned}$$

Again, this means that there is some k^* such that for all $k \leq k^*$, $\mathbf{p}_k > \mathbf{p}'_k$. Lemma A.5 yields $\mathbf{p} \succ \mathbf{p}'$. □

Lemma A.5. *Suppose distributions* $\mathbf{p}$ *and* $\mathbf{q}$ *satisfy*

$$\mathbf{p}_k < \mathbf{q}_k \Longleftrightarrow k \leq k^*$$

for some k^**. Then,* $\mathbf{q} \succ \mathbf{p}$*.*

Proof. By definition [e.g., 68], $\mathbf{q} \succ \mathbf{p}$ if and only if for all k,

$$\sum_{\ell=0}^{k} \mathbf{q}_\ell \geq \sum_{\ell=0}^{k} \mathbf{p}_\ell. \tag{31}$$

By assumption, there is some k^* such that $\mathbf{q}_k > \mathbf{p}_k$ if and only if $k \leq k^*$. For $k \leq k^*$,

$$\sum_{\ell=0}^{k} \mathbf{q}_\ell \geq \sum_{\ell=0}^{k} \mathbf{p}_\ell$$

because $\mathbf{q}_\ell > \mathbf{p}_\ell$ for each ℓ. For $k > k^*$,

$$\begin{aligned}\sum_{\ell=0}^{k} \mathbf{q}_\ell &= 1 - \sum_{\ell=k+1}^{K} \mathbf{q}_\ell \\ &\geq 1 - \sum_{\ell=k+1}^{K} \mathbf{p}_\ell && (\mathbf{q}_\ell \leq \mathbf{p}_\ell \text{ for } k > k^*) \\ &= \sum_{\ell=0}^{k} \mathbf{p}.\end{aligned}$$

Therefore, $\mathbf{q} \succ \mathbf{p}$. □

A.3 Omitted proofs for Section 4.3

Vetta [100, Thm. 5] shows that the price of anarchy is at most 2 for any *valid utility game*. We will show that $\mathscr{G}(n, \mathbf{d}, s)$ is a valid utility game. Before we do, introduce notation allowing us to consider asymmetric strategies. As before, it suffices to consider pure strategies, since each mixed strategy is equivalent to a pure strategy. Define $\mathbf{P}$ to be the strategy matrix where column i is the strategy for player i. Because we will still use subscripts to denote $k \in [K]$, we will use $\mathbf{P}(\cdot)$ to index the columns of $\mathbf{P}$. We denote the ith player's strategy by $\mathbf{P}(i)$, the strategy matrix for the remaining $n-1$ players by $\mathbf{P}(-i)$, and for any $S \subseteq [n]$, $\mathbf{P}(S)$ denotes the strategy matrix given by restricting to only players in S. Again slightly abusing notation, define

$$\begin{aligned}X_k(\mathbf{P}) &\triangleq \sum_{i \in [n]} X(1, \mathbf{P}(i)_k) \\ U_i(\mathbf{P}) &\triangleq \sum_{k \in [K]} \mathbf{d}_k \mathbf{P}(i)_k \mathbb{E}\left[\frac{1}{s(1 + X_k(\mathbf{P}(-i)))}\right] \\ W(\mathbf{P}) &\triangleq \begin{cases} \sum_{k \in [K]} \mathbf{d}_k \mathbb{E}\left[\frac{X_k(\mathbf{P})}{s(X_k(\mathbf{P}))}\right] & s \in \mathcal{S}_\nearrow \\ \sum_{k \in [K]} \mathbf{d}_k \Pr[X_k(\mathbf{P}) > 0] & s \in \mathcal{S}_\searrow \end{cases}\end{aligned}$$

A valid utility game [100] satisfies the following three conditions:

1. **Submodularity:** The social welfare function is submodular in the set of players participating. For $A \subseteq B \subseteq [n]$,

$$W(\mathbf{P}(A \cup \{i\})) - W(\mathbf{P}(A)) \geq W(\mathbf{P}(B \cup \{i\})) - W(\mathbf{P}(B)).$$

2. **Players' marginal contributions are rewarded:** For each player, their utility is at least their marginal contribution to social welfare given all other players' behaviors:

$$U_i(\mathbf{P}) \geq W(\mathbf{P}) - W(\mathbf{P}(-i)).$$

3. **Social welfare is at least players' utilities:** The sum of all players' utilities is at most the social welfare:

$$\sum_{i\in[n]} U_i(\mathbf{P}) \le W(\mathbf{P}).$$

With this, we can prove Theorem 4.11.

Theorem 4.11 (Price of anarchy)**.** *The price of anarchy for $\mathscr{G}$ is 2. Formally, let $\mathbf{P}$ be any equilibrium of $\mathscr{G}(n, \mathbf{d}, s)$, and let $\mathbf{P}^*$ be any socially optimal strategy matrix. Then,*

$$W(\mathbf{P}^*) \le 2W(\mathbf{P}).$$

There exists a family of instances $\{\mathscr{G}(n, \mathbf{d}^{(n)}, s)\}_{n\to\infty}$ for which this is asymptotically tight.

Proof. Assume without loss of generality that $\mathbf{d}$ is decreasing. We will show that $\mathscr{G}(n, \mathbf{d}, s)$ is a valid utility game.

Submodularity. It suffices to consider the case where $s \in \mathcal{S}_{\nearrow}$, since when $s \in \mathcal{S}_{\searrow}$, social welfare is equivalent to the case $s = s_1 \in \mathcal{S}_{\nearrow}$. We want to show that for $A \subseteq B \subseteq [n]$,

$$W(\mathbf{P}(A\cup\{i\})) - W(\mathbf{P}(A)) \ge W(\mathbf{P}(B\cup\{i\})) - W(\mathbf{P}(B)). \tag{32}$$

To do so, it suffices to show that for each $k \in [K]$,

$$\begin{aligned}&\mathbf{d}_k\mathbb{E}\left[\frac{X_k(\mathbf{P}(A\cup\{i\}))}{s(X_k(\mathbf{P}(A\cup\{i\})))}\right] - \mathbb{E}\left[\frac{X_k(\mathbf{P}(A))}{s(X_k(\mathbf{P}(A)))}\right]\\ \ge\ &\mathbf{d}_k\mathbb{E}\left[\frac{X_k(\mathbf{P}(B\cup\{i\}))}{s(X_k(\mathbf{P}(B\cup\{i\})))}\right] - \mathbb{E}\left[\frac{X_k(\mathbf{P}(B))}{s(X_k(\mathbf{P}(B)))}\right].\end{aligned}$$

Observe that because $B \supseteq A$,

$$X_k(\mathbf{P}(B)) = X_k(\mathbf{P}(B\backslash A)) + X_k(\mathbf{P}(A)).$$

Because $X_k(\cdot) \ge 0$, this means

$$X_k(\mathbf{P}(B)) \ge X_k(\mathbf{P}(A)).$$

Because $s \in \mathcal{S}_{\nearrow}$, by assumption, $x/s(x)$ is increasing and concave. For any $a \le b$ and $c \ge 0$,

$$\frac{a+c}{s(a+c)} - \frac{a}{s(a)} \ge \frac{b+c}{s(b+c)} - \frac{b}{s(b)}.$$

Thus,

$$\frac{X_k(\mathbf{P}(A)+X_k(\mathbf{P}(i))}{s(X_k(\mathbf{P}(A)+X_k(\mathbf{P}(i)))} - \frac{X_k(\mathbf{P}(A))}{s(X_k(\mathbf{P}(A)))} \ge \frac{X_k(\mathbf{P}(B))X_k(\mathbf{P}(i)}{s(X_k(\mathbf{P}(B))X_k(\mathbf{P}(i))} - \frac{X_k(\mathbf{P}(B))}{s(X_k(\mathbf{P}(B)))}.$$

Taking expectations of both sides, multiplying both sides by $\mathbf{d}_k$, and summing over $k \in [K]$ yields (32).

Players' marginal contributions are rewarded. Next, we must show

$$U_i(\mathbf{P}) \geq W(\mathbf{P}) - W(\mathbf{P}(-i)). \tag{33}$$

For $s \in \mathcal{S}_{\nearrow}$,

$$\begin{aligned}
W(\mathbf{P}) &= \sum_{k\in[K]} \mathbf{d}_k \mathbb{E}\left[\frac{X_k(\mathbf{P})}{s(X_k(\mathbf{P}))}\right] \\
&= \sum_{k\in[K]} \mathbf{d}_k \mathbb{E}\left[\frac{X_k(\mathbf{P}(-i)) + X_k(\mathbf{P}(i))}{s(X_k(\mathbf{P}(-i)) + X_k(\mathbf{P}(i)))}\right] \\
&= \sum_{k\in[K]} \mathbf{d}_k \left(\mathbb{E}\left[\mathbb{E}\left[\frac{X_k(\mathbf{P}(-i)) + X_k(\mathbf{P}(i))}{s(X_k(\mathbf{P}(-i)) + X_k(\mathbf{P}(i)))}\right] \;\middle|\; X_k(\mathbf{P}(i))\right]\right) \\
&= \sum_{k\in[K]} \mathbf{d}_k \left(\mathbf{P}(i)_k \mathbb{E}\left[\frac{1}{s(1 + X_k(\mathbf{P}(-i)))}\right] + (1 - \mathbf{P}(i)_k)\mathbb{E}\left[\frac{X_k(\mathbf{P}(-i))}{s(X_k(\mathbf{P}(-i)))}\right]\right) \\
&= U_i(\mathbf{P}) + \sum_{k\in[K]} \mathbf{d}_k (1 - \mathbf{P}(i)_k)\mathbb{E}\left[\frac{X_k(\mathbf{P}(-i))}{s(X_k(\mathbf{P}(-i)))}\right] \\
&\leq U_i(\mathbf{P}) + \sum_{k\in[K]} \mathbf{d}_k \mathbb{E}\left[\frac{X_k(\mathbf{P}(-i))}{s(X_k(\mathbf{P}(-i)))}\right] \\
&= U_i(\mathbf{P}) + W(\mathbf{P}(-i)).
\end{aligned}$$

For $s \in \mathcal{S}_{\searrow}$,

$$\begin{aligned}
W(\mathbf{P}) &= \sum_{k\in[K]} \mathbf{d}_k \Pr[X_k(\mathbf{P}) > 0] \\
&= \sum_{k\in[K]} \mathbf{d}_k \Pr[X_k(\mathbf{P}(-i)) > 0 \vee X_k(\mathbf{P}(i)) > 0] \\
&= \sum_{k\in[K]} \mathbf{d}_k \left(\Pr[X_k(\mathbf{P}(i)) > 0 \wedge X_k(\mathbf{P}(-i)) = 0] + \Pr[X_k(\mathbf{P}(-i)) > 0]\right) \\
&= W(\mathbf{P}(-i)) + \sum_{k\in[K]} \mathbf{d}_k \mathbf{P}(i)_k \Pr[X_k(\mathbf{P}(-i)) = 0] \\
&= W(\mathbf{P}(-i)) + \sum_{k\in[K]} \mathbf{d}_k \mathbf{P}(i)_k \mathbb{E}\left[\frac{1}{s(1 + X_k(\mathbf{P}(-i)))} \cdot \mathbb{1}\left[X_k(\mathbf{P}(-i)) = 0\right]\right] \\
&\leq W(\mathbf{P}(-i)) + \sum_{k\in[K]} \mathbf{d}_k \mathbf{P}(i)_k \mathbb{E}\left[\frac{1}{s(1 + X_k(\mathbf{P}(-i)))}\right] \\
&= U_i(\mathbf{P}) + W(\mathbf{P}(-i)).
\end{aligned}$$

Thus, (33) holds.

Social welfare is at least players' utilities. Finally, we must show

$$\sum_{i\in[n]} U_i(\mathbf{P}) \leq W(\mathbf{P}). \tag{34}$$

We proceed as follows.

$$\begin{aligned}
\sum_{i\in[n]} U_i(\mathbf{P}) &= \sum_{k\in[K]} \mathbf{d}_k \sum_{i\in[n]} \mathbf{P}(i)_k \mathbb{E}\left[\frac{1}{s(1+X_k(\mathbf{P}(-i)))}\right] \\
&= \sum_{k\in[K]} \mathbf{d}_k \sum_{i\in[n]} \mathbb{E}\left[\frac{X_k(\mathbf{P}(i))}{s(X_k(\mathbf{P}(i))+X_k(\mathbf{P}(-i)))}\right] \\
&= \sum_{k\in[K]} \mathbf{d}_k \sum_{i\in[n]} \mathbb{E}\left[\frac{X_k(\mathbf{P}(i))}{s(X_k(\mathbf{P}))}\right] \\
&= \sum_{k\in[K]} \mathbf{d}_k \mathbb{E}\left[\frac{X_k(\mathbf{P})}{s(X_k(\mathbf{P}))}\right].
\end{aligned}$$

For $s \in \mathcal{S}_{\nearrow}$, this last line is equal to $W(\mathbf{P})$, and (34) holds. For $s \in \mathcal{S}_{\searrow}$,

$$\begin{aligned}
\sum_{k\in[K]} \mathbf{d}_k \mathbb{E}\left[\frac{X_k(\mathbf{P})}{s(X_k(\mathbf{P}))}\right] &\le \sum_{k\in[K]} \mathbf{d}_k \mathbb{E}\left[\mathbb{1}\left[X_k(\mathbf{P})>0\right]\right] \\
&= \sum_{k\in[K]} \mathbf{d}_k \Pr[X_k(\mathbf{P})>0] \\
&= W(\mathbf{P}).
\end{aligned}$$

Thus, $\mathscr{G}(n,\mathbf{d},s)$ is a valid utility game.

To show that this is tight, for some n, let

$$\mathbf{d} = [1;\ 1/n;\ \ \ldots\ 1/n].$$

Choose $K > n^2$. Let the score function be the identity s_1. We will show that there exists a strategy profile with social welfare approaching 2 and a (symmetric) equilibrium with social welfare 1. Define $\mathbf{P}^+$ as follows:

$$\begin{aligned}
\mathbf{P}^+(1) &= [1;\ 0;\ \ \ldots\ 0] \\
\mathbf{P}^+(i) &= [1/K;\ \ \ldots\ 1/K]. \qquad (i>1)
\end{aligned}$$

Because $s_1 \in \mathcal{S}_\searrow$, social welfare under $\mathbf{P}^+$ is

$$\begin{aligned}
W(\mathbf{P}^+) &= \sum_{k\in[K]} \mathbf{d}_k(1 - \Pr[X_k(\mathbf{P}^+) = 0]) \\
&= \mathbf{d}_1 + \sum_{k=2}^{K} \mathbf{d}_k \left(1 - \Pr\left[X\left(n-1, \frac{1}{K}\right) = 0\right]\right) \\
&= 1 + \sum_{k=2}^{K} \frac{1}{n}\left(1 - \left(1 - \frac{1}{K}\right)^{n-1}\right) \\
&= 1 + \frac{K-1}{n}\left(1 - \left(1 - \frac{1}{K}\right)^{n-1}\right) \\
&\geq 1 + \frac{K-1}{n}\left(1 - \left(1 - \frac{n-1}{K} + \binom{n-1}{2}\frac{1}{K^2}\right)\right) \quad \text{(by Lemma A.15 since } K > n\text{)} \\
&= 1 + \frac{K-1}{n}\left(\frac{n-1}{K} - \binom{n-1}{2}\frac{1}{K^2}\right) \\
&= 1 + \frac{(K-1)(n-1)}{Kn}\left(1 - \frac{n-2}{2K}\right) \\
&\geq 1 + \frac{(n^2-1)(n-1)}{n^3}\left(1 - \frac{1}{2n}\right) \qquad (K > n^2) \\
&\geq 2 - \frac{n^2+n}{n^3}\left(1 - \frac{1}{2n}\right) - \frac{1}{2n} \\
&\geq 2 - \frac{3}{2n} - \frac{1}{n^2}.
\end{aligned}$$

For sufficiently large n, we can make this arbitrarily close to 2. In contrast, we will show that $\mathbf{P}$ is an equilibrium, where

$$\mathbf{P}(i) = [1;\ 0;\ \ldots\ 0]$$

for all i. To show that it is an equilibrium, observe that each player gets utility $U_i(\mathbf{P}) = \frac{1}{n}$, which weakly greater than $\mathbf{d}_k$ for any $k > 1$. Therefore, $\mathbf{P}$ is (weakly) an equilibrium. However, $W(\mathbf{P}) = 1$. The price of anarchy is therefore 2. □

A.4 Auxiliary results

A.4.1 Results from prior work

Lemma A.6 (Rudin [81, Thm 3.41])**.** *For any sequences $\{x_i\}, \{y_i\}$, let $X_i = \sum_{j=0}^{i} x_i$. Then,*

$$\sum_{i=0}^{n} x_i y_i = -\sum_{i=0}^{n-1} X_i(y_{i+1} - y_i) + X_n y_n. \tag{35}$$

Lemma A.7 (Sah [82])**.** *Define*

$$\begin{aligned} f(n,p) &\triangleq \sum_{\ell=0}^{n} b(\ell,n,p)g(\ell) \\ \bar{f}(n,p) &\triangleq f(n+1,p) - f(n,p) \\ \bar{\bar{f}}(n,p) &\triangleq \bar{f}(n+1,p) - \bar{f}(n,p) \\ \bar{g}(\ell) &\triangleq g(1+\ell) - g(\ell) \\ \bar{\bar{g}}(\ell) &\triangleq \bar{g}(1+\ell) - \bar{g}(\ell) \end{aligned}$$

Then,

$$\bar{f}(n,p) = p\sum_{\ell=0}^{n} b(\ell,n,p)\bar{g}(\ell) \tag{36}$$

$$\bar{\bar{f}}(n,p) = p^2\sum_{\ell=0}^{n} b(\ell,n,p)\bar{\bar{g}}(\ell) \tag{37}$$

$$\frac{\partial}{\partial p}f(n,p) = \frac{n}{p}\bar{f}(n-1,p) \tag{38}$$

$$\frac{\partial}{\partial p}\bar{f}(n,p) = \frac{1}{p}\left(\bar{f}(n,p) + n\bar{\bar{f}}(n-1,p)\right). \tag{39}$$

Proof. (36), (37), and (39) are directly from Sah [82, eq. (4–6)]. A similar argument, which we reproduce below for completeness, gives us (38). See Sah [83] for additional detail.

From Sah [82, eq. (A6)] (see also Feller [27, eq. (10.9) p. 173]),

$$\frac{\partial}{\partial p}B(\ell,n,p) = -nb(\ell,n-1,p). \tag{40}$$

Therefore,

$$\begin{aligned} \frac{\partial}{\partial p}f(n,p) &= \frac{\partial}{\partial p}\sum_{\ell=0}^{n} b(\ell,n,p)g(\ell) \\ &= \frac{\partial}{\partial p}\left[g(n) - \sum_{\ell=0}^{n-1} B(\ell,n,p)\bar{g}(\ell)\right] && \text{(by Lemma A.6)} \\ &= -\sum_{\ell=0}^{n-1}(-nb(\ell,n-1,p))\bar{g}(\ell) && \text{(by (40))} \\ &= n\sum_{\ell=0}^{n-1} b(\ell,n-1,p)\bar{g}(\ell) \\ &= \frac{n}{p}\bar{f}(n-1,p). && \text{(by Equation (36))} \end{aligned}$$

□

A.4.2 Additional results.

Lemma A.8. *$\mathscr{G}$ is a potential game.*

Proof. Define potentially asymmetric strategy profiles $\mathbf{P}$ as in Appendix A.3. $\mathscr{G}$ is a potential game if there exists some function Φ such that, for any $\mathbf{P}$ and $\mathbf{P}'$ differing only in player i's strategy,

$$\Phi(\mathbf{P}') - \Phi(\mathbf{P}) = U_i(\mathbf{P}') - U_i(\mathbf{P}). \tag{41}$$

We will show that the this holds for the following definition of Φ, which is fairly standard in the literature on congestion games [78].

$$\Phi(\mathbf{P}) \triangleq \mathbb{E}\left[\sum_{k\in[K]} \mathbf{d}_k \sum_{\ell=1}^{X_k(\mathbf{P})} \frac{1}{s(\ell)}\right].$$

Observe that

$$\begin{aligned}
\mathbb{E}\left[\sum_{k\in[K]} \mathbf{d}_k \sum_{\ell=1}^{X_k(\mathbf{P})} \frac{1}{s(\ell)}\right] &= \sum_{k\in[K]} \mathbf{d}_k \mathbb{E}\left[\frac{X_k(\mathbf{P}(i))}{s(1+X_k(\mathbf{P}(-i)))} + \sum_{\ell=1}^{X_k(\mathbf{P}(-i))} \frac{1}{s(\ell)}\right] \\
&= \sum_{k\in[K]} \mathbf{d}_k \mathbf{P}(i)_k \frac{1}{1+X_k(\mathbf{P}(-i))} + \sum_{k\in[K]} \mathbf{d}_k \mathbb{E}\left[\sum_{\ell=1}^{X_k(\mathbf{P}(-i))} \frac{1}{s(\ell)}\right] \\
&= U_i(\mathbf{P}) + \Phi(\mathbf{P}(-i)).
\end{aligned}$$

Therefore, if $\mathbf{P}'$ and $\mathbf{P}$ only differ in column i,

$$\Phi(\mathbf{P}') - \Phi(\mathbf{P}) = U_i(\mathbf{P}') - U_i(\mathbf{P}).$$

□

Lemma A.9.

$$\frac{\partial}{\partial p}\frac{u(n,p)}{u(n-1,p)} \leq 0.$$

Proof. We proceed as follows.

$$\begin{aligned}
\frac{\partial}{\partial p}\frac{u(n,p)}{u(n-1,p)} \leq 0 &\iff \frac{\partial}{\partial p}\log\left(\frac{u(n,p)}{u(n-1,p)}\right) \leq 0 \\
&\iff \frac{\frac{\partial}{\partial p}u(n,p)}{u(n,p)} - \frac{\frac{\partial}{\partial p}u(n-1,p)}{u(n-1,p)} \leq 0 \\
&\iff \frac{\frac{n}{p}\bar{u}(n-1,p)}{u(n,p)} \leq \frac{\frac{n-1}{p}\bar{u}(n-2,p)}{u(n-1,p)} \\
&\iff \frac{n(u(n,p)-u(n-1,p))}{u(n,p)} \leq \frac{(n-1)(u(n-1,p)-u(n-2,p))}{u(n-1,p)} \\
&\iff n - \frac{nu(n-1,p)}{u(n,p)} \leq (n-1) - \frac{(n-1)u(n-2,p)}{u(n-1,p)} \\
&\iff 1 - \frac{nu(n-1,p)}{u(n,p)} \leq -\frac{(n-1)u(n-2,p)}{u(n-1,p)} \\
&\iff u(n-1,p)u(n,p) + (n-1)u(n-2,p)u(n,p) \leq nu(n-1,p)^2.
\end{aligned}$$

Using the fact that $u(n,p) \leq u(n-1,p)$ (because $X(n,p)$ stochastically dominates $X(n-1,p)$), $u(n,p)u(n-1,p) \leq u(n-1,p)^2$. Therefore, it suffices to show that

$$\begin{aligned}
(n-1)u(n-2,p)u(n,p) &\leq (n-1)u(n-1,p)^2 \\
u(n-2,p)u(n,p) &\leq u(n-1,p)^2. \qquad (42)
\end{aligned}$$

From Lemma A.7, we have

$$\begin{aligned}
&u(n,p) - u(n-1,p) - (u(n-1,p) - u(n-2,p)) \\
=&\,\bar{\bar{u}}(n-2,p) \\
=&\,p^2 \sum_{\ell=0}^{n-2} b(\ell, n-2, p)\left(\frac{1}{s(3+\ell)} - \frac{1}{s(2+\ell)} - \left(\frac{1}{s(2+\ell)} - \frac{1}{s(1+\ell)}\right)\right).
\end{aligned}$$

Under the assumption that $1/s(x)$ is convex (**D.2**), each term in the sum is weakly negative. As a result, u is discrete convex in n:

$$\begin{aligned}
u(n,p) - u(n-1,p) - (u(n-1,p) - u(n-2,p)) \leq 0 \\
u(n-1,p) \geq \frac{u(n,p) + u(n-2,p)}{2}.
\end{aligned}$$

Putting this into (42), we have

$$\begin{aligned}
u(n-1,p)^2 - u(n,p)u(n-2,p) &\geq \left(\frac{u(n,p)+u(n-2,p)}{2}\right)^2 - u(n,p)u(n-2,p) \\
&= \frac{u(n,p)^2 + 2u(n,p)u(n-2,p) + u(n-2,p)^2}{4} \\
&- u(n,p)u(n-2,p) \\
&= \frac{u(n,p)^2 + u(n-2,p)^2}{4} - \frac{u(n,p)u(n-2,p)}{2} \\
&= \frac{1}{2}\left(u(n,p)^2 - 2u(n,p)u(n-2,p) + u(n-2,p)^2\right) \\
&= \frac{1}{2}\left(u(n,p) - u(n-2,p)\right)^2 \\
&\geq 0.
\end{aligned}$$

□

Lemma A.10. *Suppose g is weakly decreasing and $g(1) < g(0)$. Let $f(n,p) = \mathbb{E}\left[g(X(n,p))\right]$. Then,*

$$\frac{\partial}{\partial p} f(n,p) < 0.$$

Proof. If $n = 1$, the claim is true because $\frac{\partial}{\partial p} f(n,p) = g(0) - g(1)$. Otherwise,

$$\begin{aligned}
\frac{\partial}{\partial p} f(n,p) &= \frac{n}{p}\bar{f}(n-1,p) && \text{(Lemma A.7)} \\
&= n\sum_{\ell=0}^{n-1} b(\ell, n-1, p)\bar{g}(\ell) && \text{(Lemma A.7)} \\
&= n\sum_{\ell=0}^{n-1} b(\ell, n-1, p)(g(1+\ell) - g(\ell)) \\
&\leq nb(0, n-1, p)(g(1) - g(0)) \\
&= n(1-p)^{n-1}\left(g(1) - g(0)\right) \\
&< 0.
\end{aligned}$$

□

Lemma A.11. *If $g(x)$ is convex, then $\mathbb{E}\left[g(X(n,p))\right]$ is convex in p. If $g(x)$ is concave, then $\mathbb{E}\left[g(X(n,p))\right]$ is concave in p. Additionally, if $n > 1$ and*

$$\frac{g(2) + g(0)}{2} \neq g(1),$$

then these are strict.

Proof. Let $f(n,p) = \mathbb{E}\left[g(X(n,p))\right]$. If $n = 1$, $f(1,p) = pg(1) + (1-p)g(0)$, which is both convex and concave. Otherwise,

$$\begin{aligned}
\frac{\partial^2}{\partial p^2} f(n,p) &= \frac{\partial}{\partial p}\frac{n}{p}\bar{f}(n-1,p) && \text{(Lemma A.7)} \\
&= n\left(\frac{1}{p}\cdot\frac{1}{p}\left(\bar{f}(n-1,p) + (n-1)\bar{\bar{f}}(n-2,p)\right) + \bar{f}(n-1,p)(-p^{-2})\right) && \text{(Lemma A.7)} \\
&= \frac{n(n-1)}{p^2}\bar{\bar{f}}(n-2,p) \\
&= n(n-1)\sum_{\ell=0}^{n-2} b(\ell,n,p)\bar{\bar{g}}(\ell) && \text{(Lemma A.7)} \\
&= n(n-1)\sum_{\ell=0}^{n-2} b(\ell,n,p)(\bar{g}(1+\ell) - \bar{g}(\ell)) \\
&= n(n-1)\sum_{\ell=0}^{n-2} b(\ell,n,p)(g(2+\ell) - g(1+\ell) - (g(1+\ell) - g(\ell))) \\
&= n(n-1)\sum_{\ell=0}^{n-2} b(\ell,n,p)(g(2+\ell) + g(\ell) - 2g(1+\ell)).
\end{aligned}$$

If g is convex, then

$$\frac{g(2+\ell) + g(\ell)}{2} \geq g(1+\ell).$$

Similarly, if g is concave, then

$$\frac{g(2+\ell) + g(\ell)}{2} \leq g(1+\ell).$$

This proves convexity or concavity. To show strictness, observe that under the given assumption, the term

$$b(0,n,p)((g(2) - g(1)) - (g(1) - g(0)) \neq 0,$$

meaning f is strictly convex/concave in p. □

Lemma A.12. *Let $\{a_\ell\}_{\ell=1}^k$ be a positive, non-decreasing sequence. Let f be a concave, nonnegative, increasing function. Then, for any positive c,*

$$\begin{aligned}
&\max_b \sum_{\ell=1}^k a_\ell f(b_\ell) \\
&s.t. \sum_\ell b_\ell \leq c \\
&\quad \{b_\ell\}_{\ell=1}^k \text{ is non-increasing}
\end{aligned}$$

is attained by $b_\ell^* = c/k$ for all ℓ. If $a_k > a_1$ or f is strictly concave, then this is the unique maximizer.

Proof. In the case where $a_k = a_1$, then all the a_ℓ's are equal and the claim is trivially true. From here, we consider the case where $a_k > a_1$. The Lagrangian is

$$\mathcal{L}(b;\mu,\lambda) = -\sum_{\ell=1}^{k} a_\ell f(b_\ell) + \mu\left(\sum_{\ell=1}^{k} b_\ell - c\right) + \sum_{\ell=1}^{k-1} \lambda_\ell (b_{\ell+1} - b_\ell).$$

An optimal solution must satisfy stationarity: for all ℓ,

$$\frac{\partial}{\partial b_\ell}\mathcal{L}(b;\mu,\lambda) = 0.$$

This means

$$\begin{aligned} -a_1 f'(b_1) + \mu - \lambda_1 &= 0 \\ -a_\ell f'(b_\ell) + \mu + \lambda_{\ell-1} - \lambda_\ell &= 0 \qquad (2 \le \ell \le k-1) \\ -a_k f'(b_k) + \mu + \lambda_{k-1} &= 0. \end{aligned}$$

Summing all these equations, we have

$$\begin{aligned} -\sum_{\ell=1}^{k} a_\ell f'(b_\ell) + k\mu &= 0 \\ \mu &= \frac{1}{k}\sum_{\ell=1}^{k} a_\ell f'(b_\ell). \end{aligned}$$

We will show that $\lambda_j > 0$ for all $j \le k-1$. Sum the first j equations above to get

$$-\sum_{\ell=1}^{j} a_\ell f'(b_\ell) + j\mu - \lambda_j = 0$$

$$\begin{aligned} \lambda_j &= j\mu - \sum_{\ell=1}^{j} a_\ell f'(b_\ell) \\ &= \frac{j}{k}\sum_{\ell=1}^{k} a_\ell f'(b_\ell) - \sum_{\ell=1}^{j} a_\ell f'(b_\ell) \\ &= \frac{j}{k}\left(\sum_{\ell=1}^{j} a_\ell f'(b_\ell) + \sum_{j+1}^{k} a_\ell f'(b_\ell)\right) - \sum_{\ell=1}^{j} a_\ell f'(b_\ell) \\ &= \frac{j}{k}\sum_{j+1}^{k} a_\ell f'(b_\ell) - \frac{k-j}{k}\sum_{\ell=1}^{j} a_\ell f'(b_\ell) \\ &= \frac{j(k-j)}{k}\frac{1}{k-j}\sum_{j+1}^{k} a_\ell f'(b_\ell) - \frac{j(k-j)}{k}\frac{1}{j}\sum_{\ell=1}^{j} a_\ell f'(b_\ell) \\ &> 0. \end{aligned}$$

The last step follows because both a_ℓ and $f'(b_\ell)$ are weakly increasing, and as long as either $a_k > a_1$ or $f'(b_k) > f(b_1)$, the overall sequence must strictly increase somewhere. The average of the last $k - j$ terms of a non-decreasing sequence that strictly increases somewhere must be strictly larger than the average of the first j terms. Complementary slackness tells us that $\lambda_\ell > 0$ implies $b_\ell = b_{\ell+1}$ for all $\ell \leq k-1$, meaning that the unique optimizer is $b^*_\ell = c/k$ for all ℓ. □

Lemma A.13. *Suppose* $\mathbf{d}_{k_1} \leq \mathbf{d}_{k_1+1} \leq \cdots \leq \mathbf{d}_{k_2}$. *Given a strategy* $\mathbf{p}$, *let* $\hat{\mathbf{p}} = \text{avg}(\mathbf{p}; k_1, k_2)$. *If* $\mathbf{p} \neq \hat{\mathbf{p}}$, *then*

$$U(\hat{\mathbf{p}}, \mathbf{p}) > U(\mathbf{p}, \mathbf{p})$$
$$W(\hat{\mathbf{p}}) > W(\mathbf{p}).$$

Proof.

$$U(\hat{\mathbf{p}}, \mathbf{p}) - U(\mathbf{p}, \mathbf{p}) = \sum_{k=k_1}^{k_2} \mathbf{d}_k \left(\hat{\mathbf{p}}_k - \mathbf{p}_k\right) u(n-1, \mathbf{p}_k)$$

Observe that $\mathbf{d}$ is weakly increasing over this range by assumption. $u(n-1, \mathbf{p}_k)$ is weakly increasing as well by Lemma A.10 because $\mathbf{p}$ is weakly decreasing in k by construction. Moreover, it must strictly increase somewhere under the assumption that $\hat{\mathbf{p}} \neq \mathbf{p}$, because this implies that $\mathbf{p}$ strictly decreases somewhere between k_1 and k_2. By Lemma A.12 with $a_\ell = \mathbf{d}_\ell u(n-1, \mathbf{p}_\ell)$, utility is maximized by uniform weights given by $\hat{\mathbf{p}}$, and the inequality is strict:

$$U(\hat{\mathbf{p}}, \mathbf{p}) > U(\mathbf{p}, \mathbf{p}).$$

To show that social welfare improves, first we take the case where $s \in \mathcal{S}_\nearrow$. Then, consider expected social welfare considering only types k_1 through k_2:

$$\sum_{k=k_1}^{k_2} \mathbf{d}_k \mathbb{E}\left[\frac{X(n, \mathbf{p}_k)}{s(X(n, \mathbf{p}_k))}\right].$$

$\mathbb{E}\left[X(n,p)/s(X(n,p))\right]$ is strictly concave in p because by assumption, $x/s(x)$ is discrete-concave (**D.4(a)**) and $2/s(2) < 2$ (**D.2** and **D.3**). Lemma A.12 yields the desired claim, that this expression is strictly maximized at $\hat{\mathbf{p}}$.

In the case where $s \in \mathcal{S}_\searrow$, observe that this expression is

$$\sum_{k=k_1}^{k_2} \mathbf{d}_k \Pr[X(n, \mathbf{p}_k) > 0].$$

This coincides with social welfare under the identity score function $s_1 \in \mathcal{S}_\nearrow$, which we just showed is strictly maximized at $\hat{\mathbf{p}}$. Therefore, the claim holds for any $s \in \mathcal{S}_\searrow$ as well. □

Lemma A.14. *Assume* $\mathbf{d}_{k_1} \leq \mathbf{d}_{k_1+1} \leq \cdots \leq \mathbf{d}_{k_2}$. *Let* $\tilde{\mathbf{d}} = \text{avg}(\mathbf{d}; k_1, k_2)$. *For any* $\mathbf{p}, \mathbf{p}'$ *such that* $\mathbf{p} = \text{avg}(\mathbf{p}; k_1, k_2)$ *and* $\mathbf{p}' = \text{avg}(\mathbf{p}'; k_1, k_2)$, $U(\mathbf{p}, \mathbf{p}'; n, \mathbf{d}, s) = U(\mathbf{p}, \mathbf{p}'; n, \tilde{\mathbf{d}}, s)$ *and* $W(\mathbf{p}; n, \mathbf{d}, s) = W(\mathbf{p}; n, \tilde{\mathbf{d}}, s)$.

Proof. We begin with utility.

$$\begin{aligned}
U(\mathbf{p},\mathbf{p}';n,\tilde{\mathbf{d}},s) - U(\mathbf{p},\mathbf{p}';n,\mathbf{d},s) &= \sum_{k=k_1}^{k_2} (\tilde{\mathbf{d}}_k - \mathbf{d}_k)\mathbf{p}_k u(n-1,\mathbf{p}'_k) \\
&= \sum_{k=k_1}^{k_2} (\tilde{\mathbf{d}}_k - \mathbf{d}_k)\mathbf{p}_{k_1} u(n-1,\mathbf{p}'_{k_1}) \\
&\qquad\qquad (\mathbf{p} = \mathrm{avg}(\mathbf{p};k_1,k_2) \text{ and } \mathbf{p}' = \mathrm{avg}(\mathbf{p}';k_1,k_2)) \\
&= (k_2 - k_1 + 1)\mathbf{p}_{k_1} u(n-1,\mathbf{p}'_{k_1}) \sum_{k=k_1}^{k_2} (\tilde{\mathbf{d}}_k - \mathbf{d}_k) \\
&= 0. \qquad\qquad (\mathbf{d} = \mathrm{avg}(\mathbf{d};k_1,k_2))
\end{aligned}$$

Next, we handle social welfare in the case where $s \in \mathcal{S}_{\nearrow}$. We slightly abuse notation to let $W(\mathbf{p},\mathbf{p}';n,\mathbf{d},s)$ refer to the social welfare when one player plays $\mathbf{p}$ and the rest play $\mathbf{p}'$. Let $\tilde{X}_k = X(1,\mathbf{p}_k) + X(n-1,\mathbf{p}'_k)$.

$$\begin{aligned}
W(\mathbf{p},\mathbf{p}';n,\tilde{\mathbf{d}},s) - W(\mathbf{p},\mathbf{p}';n,\mathbf{d},s) &= \sum_{k=k_1}^{k_2} (\tilde{\mathbf{d}}_k - \mathbf{d}_k)\mathbb{E}\left[\frac{\tilde{X}_k}{s(\tilde{X}_k)}\right] \\
&= (k_2 - k_1 + 1)\mathbb{E}\left[\frac{\tilde{X}_{k_1}}{s(\tilde{X}_{k_1})}\right] \sum_{k=k_1}^{k_2} (\tilde{\mathbf{d}}_k - \mathbf{d}_k) \\
&= 0. \qquad\qquad (\mathbf{d} = \mathrm{avg}(\mathbf{d};k_1,k_2))
\end{aligned}$$

As in the proof of Lemma A.13, expected social welfare in the case where $s \in \mathcal{S}_{\searrow}$ coincides with the identity function $s_1 \in \mathcal{S}_{\nearrow}$, so the claim holds for all $s \in \mathcal{S}_{\searrow}$ as well. □

Lemma A.15. *For $n > 0$ and $0 < x < 1/n$,*

$$(1-x)^n \le 1 - nx + \binom{n}{2}x^2.$$

Proof.

$$(1-x)^n = \sum_{i=0}^{\infty} \binom{n}{i}(-x)^i = 1 - nx + \binom{n}{2}x^2 + \sum_{i=3}^{\infty} \binom{n}{i}(-x)^i.$$

We need to show

$$\sum_{i=3}^{\infty} \binom{n}{i}(-x)^i \le 0.$$

Observe that the first term is negative, and for $x < 1/n$, the absolute value of each term is decreasing. Therefore, the sum is bounded from above by 0. □

B Omitted proofs for Section 5

Lemma 5.2 (The dominant GAIT depends on the strength of competition)**.** *There exists* $(n, \Pi = \{\pi^{(1)}, \pi^{(2)}\}, \mathbf{d})$ *such that* $\underline{\mathcal{M}}_1(1, \Pi, \mathbf{d}) = 1$ *and* $\overline{\mathcal{M}}_1(n, \Pi, \mathbf{d}) < \underline{\mathcal{M}}_2(n, \Pi, \mathbf{d})$.

Proof. We will provide an example where $|\Pi| = 2$. Let $\mathbf{d} = [1;\ 1-\varepsilon;\ 1-\varepsilon;\ 0;\ \ldots\ 0]$. Let $\mathbf{d}(\pi)$ denote $\mathbf{d}$ ordered according to π. We consider two GAITs $\pi^{(1)}$ and $\pi^{(2)}$ given by

$$\begin{aligned} \pi^{(1)} &= [1;\ 4;\ 5;\ \ldots\ K;\ 2;\ 3] \qquad \Longrightarrow \qquad \mathbf{d}(\pi^{(1)}) = [1;\ 0;\ \ldots\ 0;\ 1-\varepsilon;\ 1-\varepsilon] \\ \pi^{(2)} &= [2;\ 3;\ 4;\ \ldots\ K;\ 1] \qquad \Longrightarrow \qquad \mathbf{d}(\pi^{(2)}) = [1-\varepsilon;\ 1-\varepsilon;\ 0;\ \ldots\ 0;\ 1] \end{aligned}$$

First, observe that $\underline{\mathcal{M}}_1(1, \pi, \mathbf{d}) = 1$. This is because the distribution $\mathbf{p}_1 = [1;\ 0;\ \ldots\ 0] \in \Delta(\pi^{(1)})$ yields utility 1. In contrast, a single player's optimal strategy in $\Delta(\pi^{(2)})$ is $[0;\ 1;\ 0;\ \ldots\ 0]$, which yields utility $1-\varepsilon$.

Next, we will show that there exists n such that $\overline{\mathcal{M}}_1(n, \Pi, \mathbf{d}) < \underline{\mathcal{M}}_2(n, \Pi, \mathbf{d})$. In particular, choose $n = 3$. Assume towards contradiction that there is an equilibrium in which $m_1 \geq 2$. For sufficiently large K, these m_1 players must choose the equilibrium strategy $\mathbf{p} = [1;\ 0;\ \ldots\ 0]$, yielding utility at most $1/m_1$. This is because the constraint $\mathbf{p} \in \Delta(\pi^{(1)})$ means that $\mathbf{p}_1 + \mathbf{p}_2(K-2) \leq 1$ and $\mathbf{p}_1 + \mathbf{p}_3(K-1) \leq 1$, so for large K, it is optimal to choose $\mathbf{p}_2 = \mathbf{p}_3 = 0$.

Since $m_2 \leq 1$, if a player chooses $\pi^{(2)}$, their equilibrium strategy $\mathbf{q}$ satisfies $\mathbf{q}_2 + \mathbf{q}_3 = 1$ for a similar reason: for $\mathbf{q} \in \Delta(\pi^{(2)})$, $\mathbf{q}_2 + \mathbf{q}_3 + \mathbf{q}_1(K-2) \leq 1$. Again, for large K, it is optimal to choose $\mathbf{q}_1 = 0$. This means that all players who choose $\pi^{(1)}$ get utility $1/m_1$. They could deviate to choosing $\pi^{(2)}$ with $\mathbf{q}' = [0;\ 1/2;\ 1/2;\ 0;\ \ldots\ 0]$, which would yield utility at least

$$\begin{aligned} (1-\varepsilon)\frac{1}{2}\left(\mathbb{E}\left[\frac{1}{1+X(1,\mathbf{q}_2)}\right] + \mathbb{E}\left[\frac{1}{1+X(1,\mathbf{q}_3)}\right]\right) &= \frac{1-\varepsilon}{2}\left(1 - \mathbf{q}_2 + \frac{\mathbf{q}_2}{2} + 1 - \mathbf{q}_3 + \frac{\mathbf{q}_3}{2}\right) \\ &= \frac{1-\varepsilon}{2}\left(2 - \frac{\mathbf{q}_2}{2} - \frac{\mathbf{q}_3}{2}\right) \\ &= \frac{1-\varepsilon}{2}\left(2 - \frac{1}{2}\right) \qquad (\mathbf{q}_2 + \mathbf{q}_3 = 1) \\ &= \frac{3(1-\varepsilon)}{4} \\ &> \frac{1}{m_1} \end{aligned}$$

for sufficiently small ε. By contradiction, there is no pure partially symmetric equilibrium in which $m_1 \geq 2$, meaning $\overline{\mathcal{M}}_1(3, \Pi, \mathbf{d}) \leq 1$. Finally, we will show that there is an pure partially symmetric equilibrium in which $m_1 = 1$ and $m_2 = 2$. Let

$$\mathbf{p}^+ = [1;\ 0;\ \ldots\ 0] \qquad\qquad \mathbf{q}^+ = [0;\ 1/2;\ 1/2;\ 0;\ \ldots\ 0].$$

With $m_1 = 1$ and $m_2 = 2$, the player who chooses $\pi^{(1)}$ gets utility 1, and players who choose $\pi^{(2)}$ each get utility $3(1-\varepsilon)/4$. No player can strictly improve their utility by either switching to the other tool or changing their distribution. Therefore, $\overline{\mathcal{M}}_1(3, \Pi, \mathbf{d}) < \underline{\mathcal{M}}_2(3, \Pi, \mathbf{d})$. □

Lemma 5.3 (Strict domination is possible)**.** *Suppose $s(x) < \infty$ for $x \in \mathbb{N}_{>0}$, and suppose $\exists\ \hat{k}$ such that $\mathbf{d}_{\hat{k}} = 0$. For any $\pi^{(1)}$ such that $\hat{k}$ is not ranked first by $\pi^{(1)}$, there exists $\pi^{(2)}$ such that for all $\Pi \supseteq \{\pi^{(1)}, \pi^{(2)}\}$, $\overline{\mathcal{M}}_2(n, \Pi, \mathbf{d}) = 0$ for all n.*

Proof. Without loss of generality, let $\pi^{(1)} = [1;\ 2;\ \dots\ K]$ and let $\mathbf{d}_{\hat{k}} = 0$. Let $\pi^{(2)} = [\hat{k};\ 1;\ \dots\ \hat{k}{-}1;\ \hat{k}{+}1;\ K]$. Our goal is to show that $\overline{\mathcal{M}}_2(n, \Pi, \mathbf{d}) = 0$ for all n, $\Pi \subseteq \{\pi^{(1)}, \pi^{(2)}\}$. To do so, assume towards contradiction that at equilibrium, some player i chooses $\pi^{(2)}$ with distribution $\mathbf{p}$. Then, player i could strictly improve their utility by:

- switching to $\pi^{(1)}$, and
- transferring all probability mass from $\mathbf{p}_{\hat{k}}$ to $\mathbf{p}_1$.

This is because for all $k \neq \hat{k}$, player i still puts at least as much probability mass on type k. Unless $\mathbf{p}$ is the uniform distribution, at least one of these increases must be strict, and because $s(x) < \infty$, this strictly increases expected utility. If $\mathbf{p}$ is the uniform distribution, then transferring probability mass from $\hat{k}$ to 1 strictly increases utility. Therefore, there can be no equilibrium in which a player chooses $\pi^{(2)}$. □

Lemma 5.4 (Only the perfect tool weakly dominates)**.** *If $\pi^{(1)}$ is a perfect ordering according to $\mathbf{d}$, i.e., $\mathbf{d}_{\pi_1^{(1)}} \geq \mathbf{d}_{\pi_2^{(1)}} \geq \cdots \geq \mathbf{d}_{\pi_K^{(1)}}$, then $\overline{\mathcal{M}}_1(n, \Pi, \mathbf{d}) = n$ for all $\Pi \ni \pi^{(1)}$. Otherwise, there exist $\pi^{(2)}$ and n_0 such that for all $n \geq n_0$, $\overline{\mathcal{M}}_1(n, \{\pi^{(1)}, \pi^{(2)}\}, \mathbf{d}) < n$.*

Proof. Without loss of generality, let $\pi^{(1)} = [1;\ 2;\ \dots\ K]$. First, we consider the case where $\pi^{(1)}$ is a perfect ordering. For a given n, let $\mathbf{p}$ be the equilibrium distribution for $\mathscr{G}(n, \mathbf{d}, s)$ when players can only use $\pi^{(1)}$. We will show that this remains an equilibrium even when they are given the option to switch to other GAITs.

To do so, recall that by Lemma 4.1, $\mathbf{p}$ satisfies

$$U(\mathbf{p}; n, \mathbf{d}, s) \geq \mathbf{d}_k u(n-1, \mathbf{p}_k)$$

for all k. Suppose a player deviates to some distribution $\mathbf{p}'$ via some other GAIT π'. Given that the remaining $n-1$ players are all playing $\mathbf{p}$, the deviating player's utility is

$$\sum_{k \in [K]} \mathbf{p}'_k \mathbf{d}_k u(n-1, \mathbf{p}_k) \leq \sum_{k \in [K]} \mathbf{p}'_k U(\mathbf{p}; n, \mathbf{d}, s) = U(\mathbf{p}; n, \mathbf{d}, s).$$

Thus, no player can strictly improve their utility by deviating, meaning that $\mathbf{p}$ is an equilibrium.

Next, if $\pi^{(1)}$ is not a perfect ordering, then let $\hat{k}$ be such that $\mathbf{d}_{\hat{k}-1} < \mathbf{d}_{\hat{k}}$. Assume towards contradiction that there exists an equilibrium in which all players use $\pi^{(1)}$. This means they must play the strategy $\mathbf{p} = \mathbf{p}_{\text{EQ}}(n, \mathbf{d}, s) = \mathbf{p}_{\text{EQ}}(n, \tilde{\mathbf{d}}, s)$ where $\tilde{\mathbf{d}}$ is the result of the application of the Pool Adjacent Violators algorithm to $\mathbf{d}$ (Theorem 4.3). Let n_0 be large enough that $\mathbf{p}_{\text{EQ}}(n_0, \tilde{\mathbf{d}}, s)_{\hat{k}} > 0$. (Such an n_0 must exist by a similar argument to the one used in the proof of Lemma 4.8.) Using Lemma 4.1, as long as $n \geq n_0$,

$$\mathbf{d}_{\hat{k}} u(n-1, \mathbf{p}_{\hat{k}}) > \tilde{\mathbf{d}}_{\hat{k}} u(n-1, \mathbf{p}_{\hat{k}}) = U(\mathbf{p}; n, \tilde{\mathbf{d}}, s) = U(\mathbf{p}; n, \mathbf{d}, s).$$

Let $\pi^{(2)}$ rank $\hat{k}$ first. Then, a player could deviate to choosing $\pi^{(2)}$ and place all of their probability mass on $\hat{k}$ to get utility $\mathbf{d}_{\hat{k}} u(n-1, \mathbf{p}_{\hat{k}}) > U(\mathbf{p}; n, \mathbf{d}, s)$. Therefore, there is no equilibrium in which all players use $\pi^{(1)}$. □

C Additional Information on Experiments

C.1 Additional figures

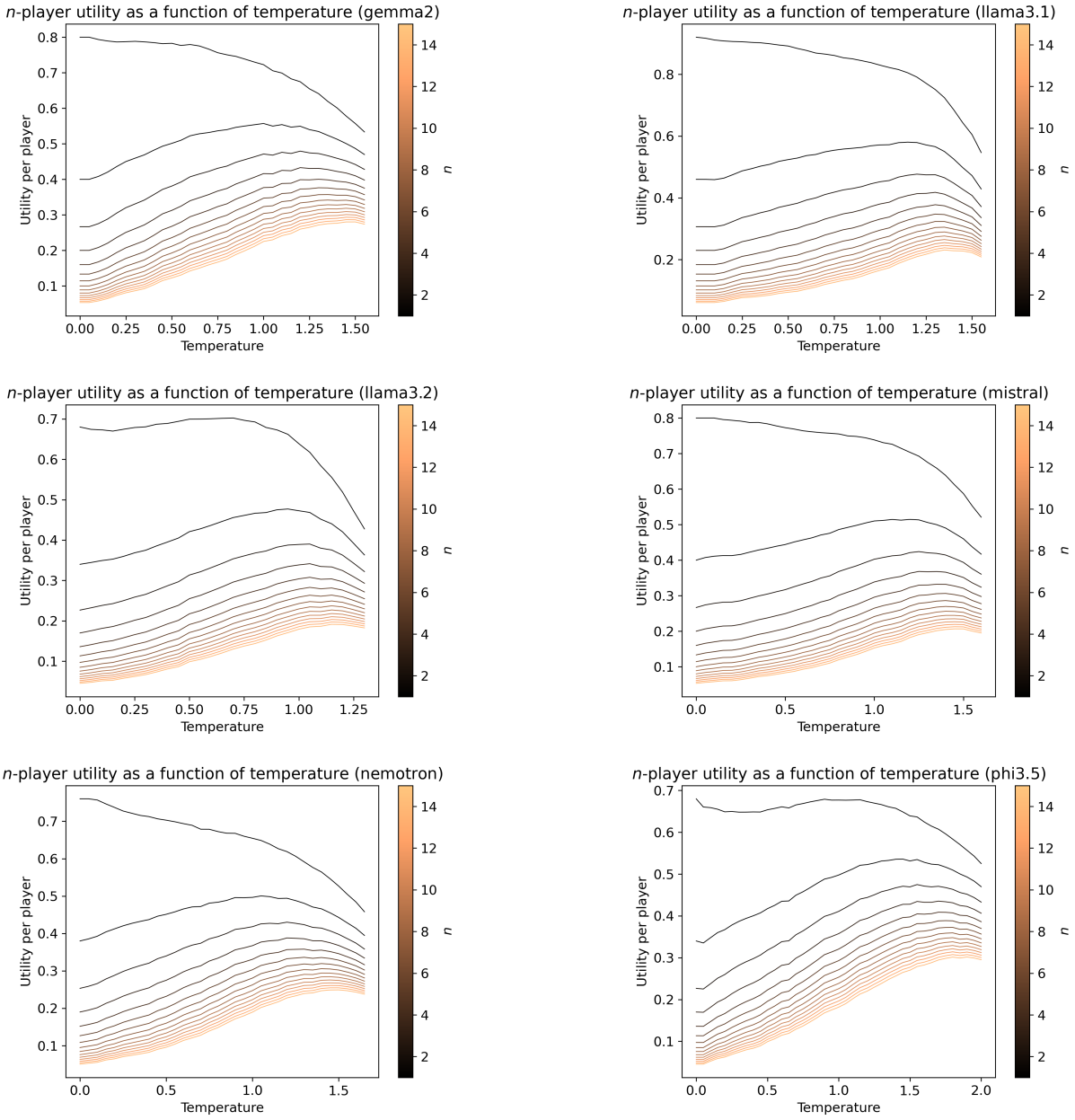


Figure 7: Social welfare as a function of τ for different values of n. $\gamma = 1.0$.

Figure 7 replicates Figure 2 for all the language models we use. Figures 8 to 10, provide analogous results to Figures 3 and 4 respectively using γ instead of n. Because social welfare is different from the sum of players' utilities when $\gamma > 1.0$, we plot them separately. Figure 11 shows pairwise equilibria for all pairs of language models in a style similar to Figure 6. We

provide the prompts we use in Figure 13. For gemma2, we omit the "System" role because gemma2 does not support it.

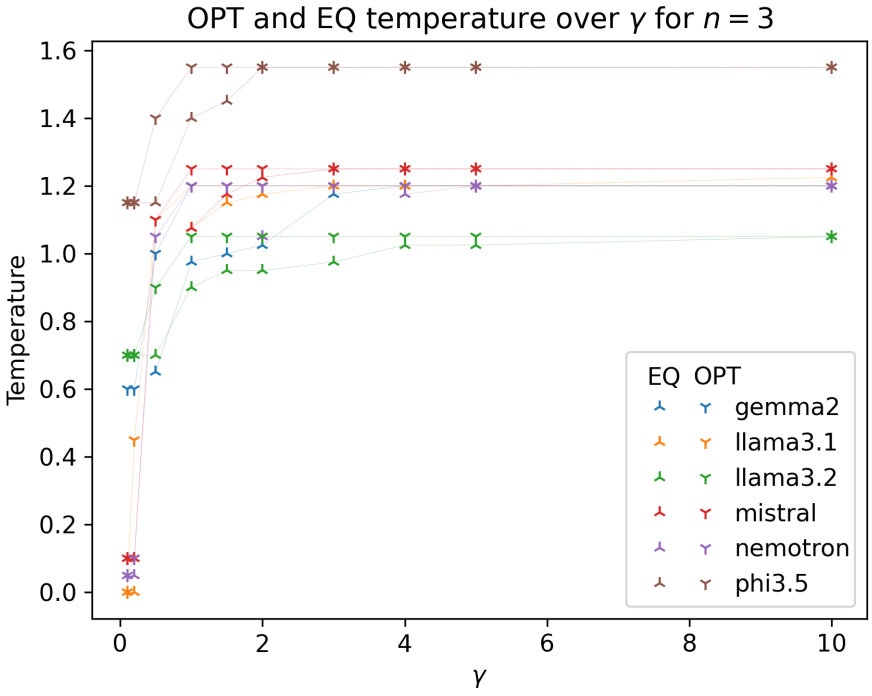


Figure 8: Socially optimal and equilibrium temperatures for each language model as a function of γ. Observe that (1) all curves are increasing in γ, and (2) equilibrium temperatures are lower than their socially optimal counterparts. Optimal and equilibrium temperatures converge for large γ, as predicted by Lemma 4.9. Note that llama3.1 appears to have higher equilibrium temperature than optimal temperature for $\gamma \geq 5$. This is an artifact of discretization of the temperature space. $n = 3$.

C.2 Experimental details

Using samples optimally. Given a finite number of samples for each (language model, temperature, instance) combination, we could naively estimate utilities in various competitive settings by resampling from that sample and simulating games. Critically, we must resample *without replacement* when simulating a game; if a particular answer appears once in our sample, resampling with replacement would allow two competing players to both sample it, which would bias our estimates of utility downwards. To simulate n players using a finite sample $\mathcal{K}$ of size $|\mathcal{K}| = S$ drawn i.i.d. from $\mathbf{p}$, we would randomly partition the sample into n subsets, and simulate $\lfloor S/n \rfloor$ games where we use each sample at most once. Of course, we could repeat this random partitioning step to reduce variance.

In fact, a variant of this scheme turns out to be optimal. Let $K_1, \ldots, K_n$ be random variables representing realized types of the n players in our simulation, where each K_i is drawn without replacement from $\mathcal{K}$. Let $\mathbb{E}[g(K_1, \ldots, K_n)]$ be the average utility of the n

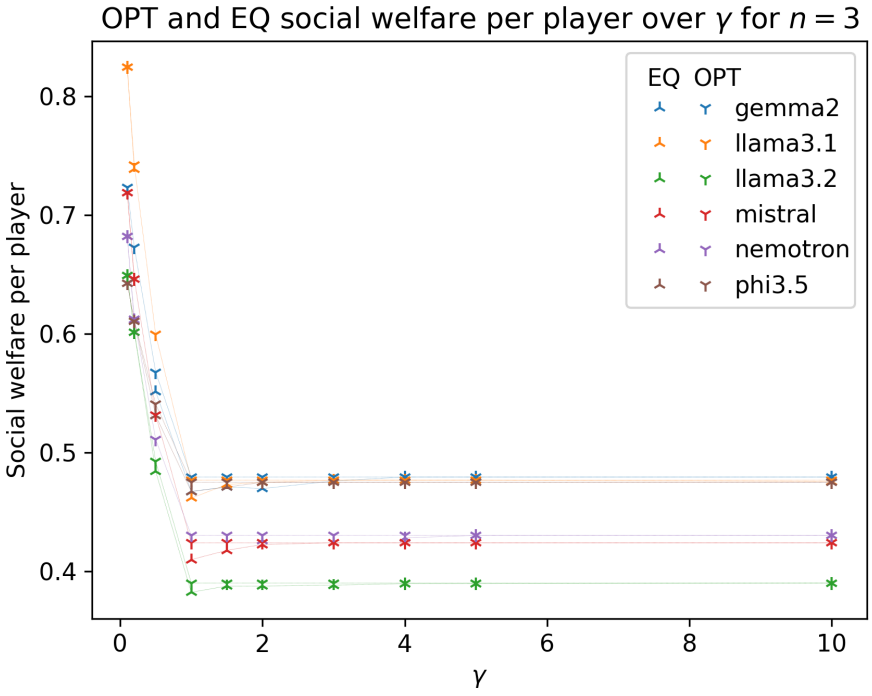


Figure 9: Social welfare for each language model at optimal and equilibrium temperatures as a function of γ. Note that for $\gamma \geq 1$, optimal social welfare is constant. Observe that (1) social welfare is decreasing in γ for $\gamma \leq 1$, and (2) equilibrium welfare is not too far from optimal. The nonmonotonicity at $\gamma = 1$ is due to the fact that our social welfare changes at this point. Interestingly, social welfare for gemma2, phi3.5, and llama3.1 are nearly identical for large values of γ despite the fact that they yield very different per-player utilities (Figure 10). $n = 3$.

players with these sampled responses and parameters $\mathbf{d}, s$. Note that g is symmetric in its arguments. We are interested in the quantity

$$U(\mathbf{p}; n, \mathbf{d}, s) = \mathbb{E}_{K_1,\ldots,K_n \sim \mathbf{p}} \left[g(K_1, \ldots, K_n) \right],$$

but we cannot compute it without access to $\mathbf{p}$. However, given a finite sample $\mathcal{K}$ drawn i.i.d. from $\mathbf{p}$, the minimum-variance [12, 76] estimator for $U(\mathbf{p})$ is the U-statistic [37]

$$\widehat{U}(\mathcal{K}) \triangleq \mathbb{E}_{K_1,\ldots,K_n \overset{\text{w/o rep.}}{\sim} \mathcal{K}} \left[g(K_1, \ldots, K_n) \right].$$

Note that the K_i's are drawn without replacement from $\mathcal{K}$. We can directly compute $\widehat{U}(\mathcal{K})$ as follows. For each type k, let a_k be the number of times k appears in $\mathcal{K}$. Suppose that player 1 samples type $\hat{k}$, i.e., $K_1 = \hat{k}$. The probability that exactly ℓ of the remaining $n-1$ players also sample type $\hat{k}$ is

$$\frac{\binom{a_{\hat{k}}-1}{\ell}\binom{S-a_{\hat{k}}}{n-1-\ell}}{\binom{S-1}{n-1}}.$$

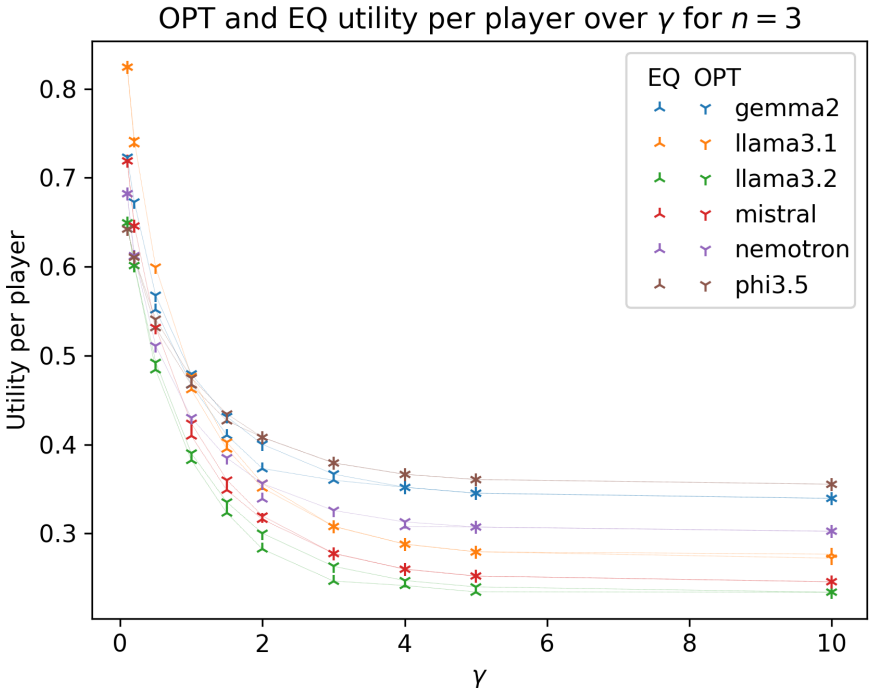


Figure 10: Socially optimal and equilibrium utility (per player) for each language model as a function of γ. Note that for $\gamma > 1$, social welfare (Figure 9) is strictly larger than the sum of player utilities. Observe that (1) utility is decreasing in γ, and (2) equilibrium welfare is not too far from optimal. Note that phi3.5 and gemma2 seem to perform relatively better in the presence of stronger competition. $n = 3$.

By symmetry,

$$\begin{aligned}\widehat{U}(\mathcal{K}) &= \sum_{k\in K} \mathbf{d}_k \mathbb{E}_{K_1,\ldots,K_n \overset{\text{w/o rep.}}{\sim} \mathcal{K}} \left[\frac{\mathbb{1}\left[K_1 = k\right]}{s\left(\sum_{i\in[n]} \mathbb{1}\left[K_i = k\right]\right)} \right] \\ &= \sum_{k\in\mathcal{K}} \frac{a_k \mathbf{d}_k}{S} \sum_{\ell=0}^{n-1} \frac{1}{s(1+\ell)} \cdot \frac{\binom{a_k-1}{\ell}\binom{S-a_k}{n-1-\ell}}{\binom{S-1}{n-1}}.\end{aligned}$$

Note that we use $a_k - 1$ in the first binomial coefficient to account for the fact that, given that $K_1 = k$, there are $a_k - 1$ samples of type k remaining in $\mathcal{K}$. A near-identical approach yields a social welfare estimator. For $s \in \mathcal{S}_{\nearrow}$, social welfare is simply $n\times$ per-player utility, so

$$\widehat{W}(\mathcal{K}) = n\widehat{U}(\mathcal{K}).$$

For $s \in \mathcal{S}_{\searrow}$, social welfare coincides with the sum of player utilities under s_1, so

$$\widehat{W}(\mathcal{K}) = n\widehat{U}(\mathcal{K}; s_1),$$

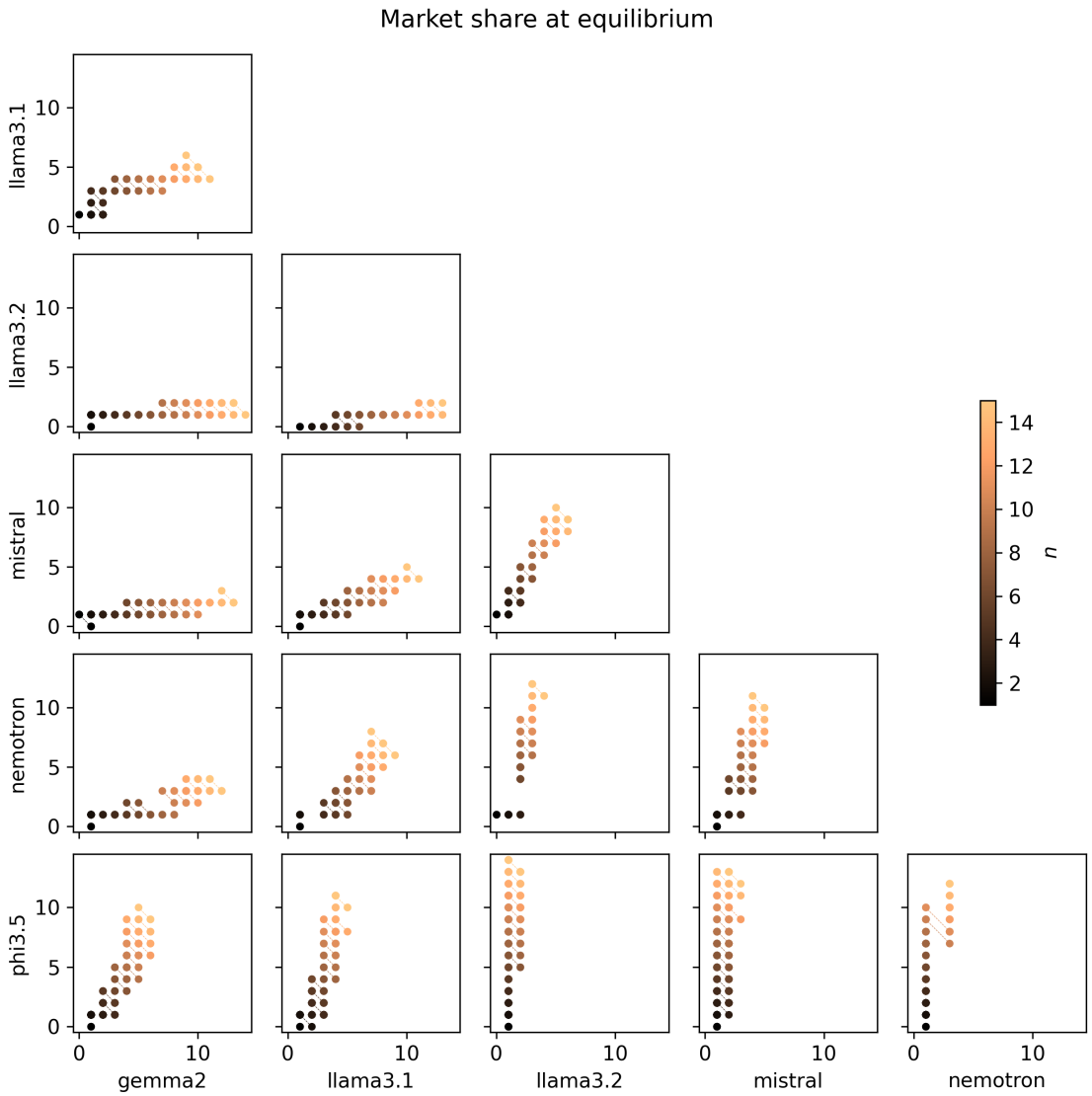


Figure 11: Pairwise equilibria for all pairs of GAITs. $\gamma = 1.0$.

where in a slight abuse of notation, we use $\widehat{U}(\mathcal{K}; s_1)$ to denote our U-statistic with the score function s_1.

To compute and verify equilibria, we also need to estimate quantities of the form $U(\mathbf{p}, \mathbf{p}')$. We can use a similar approach, where players are drawing from disjoint samples $\mathcal{K}$ and $\mathcal{K}'$. (e.g., player 1 is using a different temperature and/or language model from the remaining players). Letting a'_k be the number of times type k appears in $\mathcal{K}'$, an unbiased estimator for

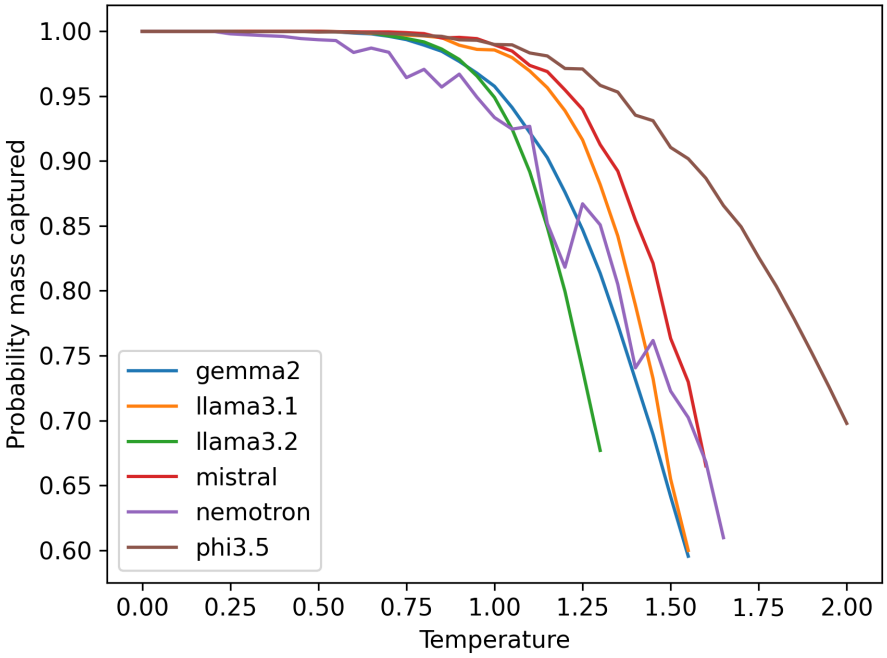


Figure 12: Probability mass captured by our sample at different temperatures.

$U(\mathbf{p}, \mathbf{p}')$ is

$$\begin{aligned}\widehat{U}(\mathcal{K}, \mathcal{K}') &= \mathbb{E}_{K_1 \sim \mathcal{K}; K_2, \ldots, K_n \overset{\text{w/o rep.}}{\sim} \mathcal{K}'} \left[\sum_{k \in [K]} \frac{\mathbf{d}_k \mathbb{1}\left[K_1 = k\right]}{s\left(\sum_{i \in [n]} \mathbb{1}\left[K_i = k\right]\right)} \right] \\ &= \sum_{k \in \mathcal{K}} \frac{a_k \mathbf{d}_k}{S} \sum_{\ell=0}^{n-1} \frac{1}{s(1+\ell)} \cdot \frac{\binom{a'_k}{\ell}\binom{S-a'_k}{n-1-\ell}}{\binom{S}{n-1}}.\end{aligned}$$

Estimating variance. To get an estimate of the variance of $\widehat{U}(\mathcal{K})$, we use the following fact [e.g. 56, Sec. 1.3, Thm. 3]:

$$\operatorname{Var}(\widehat{U}(\mathcal{K})) = \binom{S}{n}^{-1} \sum_{i=1}^{n} \binom{n}{i}\binom{S-n}{n-i}\sigma_i^2,$$

where σ_i^2 is the conditional variance

$$\sigma_i^2 \triangleq \operatorname{Var}\left(\mathbb{E}\left[g(K_1, \ldots, K_n) \mid K_1, \ldots, K_i\right]\right).$$

Model name	Hugging Face model ID	# params	Ref.	$\tau_{\max}$	$\lvert\mathcal{V}\rvert$
gemma2	google/gemma-2-2b-it	2.6B	[91, 92]	1.55	169,069
llama3.1	meta-llama/Llama-3.1-8B-Instruct	8.0B	[24]	1.55	72,399
llama3.2	meta-llama/Llama-3.2-1B-Instruct	1.2B	[24]	1.3	72,399
mistral	mistralai/Mistral-7B-Instruct-v0.3	7.2B	[47]	1.6	25,093
nemotron	nvidia/Nemotron-Mini-4B-Instruct	4.2B	[67, 72]	1.65	169,159
phi3.5	microsoft/Phi-3.5-mini-instruct	3.8B	[1]	2.0	24,158
qwen2.5	Qwen/Qwen2.5-7B-Instruct	7.6B	[93]	–	–

Table 1: Language model details. $\tau_{\max}$ is the maximum temperature at which we sample using that model. $\mathcal{V}$ is the set of tokens we allow (i.e., those that only contain alphanumeric characters plus the built in end-of-message character). We qwen2.5 only for answer verification, not for sample generation.

For large S, $\mathrm{Var}(\widehat{U}(\mathcal{K}))$ is dominated by the first term:

$$\begin{aligned}\binom{S}{n}^{-1}\sum_{i=1}^{n}\binom{n}{i}\binom{S-n}{n-i}\sigma_i^2 &= \sum_{i=1}^{n}\binom{n}{i}\frac{\binom{S-n}{n-i}}{\binom{S}{n}}\sigma_i^2\\ &= \sum_{i=1}^{n}\Theta(S^{-i})\sigma_i^2.\end{aligned}$$

Each σ_i^2 is independent of S, and they are all bounded from above by $\frac{1}{4}$ since utility is between 0 and 1. Therefore, for fixed n and sufficiently large S,

$$\begin{aligned}\mathrm{Var}(\widehat{U}(\mathcal{K})) &= \binom{n}{1}\frac{\binom{S-n}{n-1}}{\binom{S}{n}}\sigma_1^2 + \Theta(S^{-2})\\ &= n\frac{(S-n)!n!(S-n)!}{(S-2n+1)!(n-1)!S!}\sigma_1^2 + \Theta(S^{-2})\\ &= n^2\frac{(S-n)!(S-n)!}{(S-2n+1)!S!}\sigma_1^2 + \Theta(S^{-2})\\ &= \frac{n^2}{S}\sigma_1^2 + \Theta(S^{-2}).\end{aligned} \tag{43}$$

We will show that $\sigma_1^2 \le n^{-2}$. First, we take the case where $s \in \mathcal{S}_{\nearrow}$. Using the fact that the sum of player utilities is $\sum_{k\in[K]}\mathbf{d}_k C_k/s(C_k)$, where C_k is the number of players who sample

Letter	Category	Letter	Category
B	Fish	L	Things With Tails
C	Spicy Foods	M	Crimes
D	Things That Can Kill You	O	Things With Tails
D	Words With Double Letters	Q	Game Terms
E	Bad Habits	R	Junk Food
E	Birds	R	Software
E	Things Found in a Hospital	R	Something You Keep Hidden
G	Sports	U	Items in a Kitchen
I	Reptiles/Amphibians	Y	Television Stars
J	Things You Get in the Mail	Z	Articles of Clothing
K	Terms of Endearment	Z	Kinds of Soup
K	Types of Drinks	Z	Things You Replace
L	Colors		

Table 2: Instances in our sample

type k,

$$
\begin{aligned}
\sigma_1^2 &= \operatorname{Var}\left(\mathbb{E}\left[g(K_1,\ldots,K_n) \mid K_1\right]\right) \\
&= \mathbb{E}\left[\left(\mathbb{E}\left[g(K_1,\ldots,K_n) \mid K_1\right] - \mathbb{E}\left[\mathbb{E}\left[g(K_1,\ldots,K_n) \mid K_1\right]\right]\right)^2\right] \\
&= \mathbb{E}\left[\left(\mathbb{E}\left[g(K_1,\ldots,K_n) \mid K_1\right] - \mathbb{E}\left[g(K_1,\ldots,K_n)\right]\right)^2\right] \\
&= \sum_{k_1\in[K]} \mathbf{p}_{k_1} \left(\frac{1}{n} \sum_{k\in[K]} \mathbf{d}_k \mathbb{E}\left[\frac{C_k}{s(C_k)} \mid K_1 = k_1\right] - \frac{1}{n} \sum_{k\in[K]} \mathbf{d}_k \mathbb{E}\left[\frac{X(n,\mathbf{p}_k)}{s(X(n,\mathbf{p}_k)}\right] \right)^2 \\
&= \frac{1}{n^2} \sum_{k_1\in[K]} \mathbf{p}_{k_1} \left(\mathbf{d}_{k_1} \left(\mathbb{E}\left[\frac{1+X(n-1,\mathbf{p}_{k_1})}{s(1+X(n-1,\mathbf{p}_{k_1})}\right] - \mathbb{E}\left[\frac{X(n,\mathbf{p}_{k_1})}{s(X(n,\mathbf{p}_{k_1}))}\right] \right) \right. \\
&\qquad \left. + \sum_{k\neq k_1} \mathbf{d}_k \mathbb{E}\left[\frac{X(n-1,\mathbf{p}_k)}{s(X(n-1,\mathbf{p}_k))}\right] - \mathbf{d}_k \mathbb{E}\left[\frac{X(n,\mathbf{p}_k)}{s(X(n,\mathbf{p}_k)}\right] \right)^2 \\
&= \frac{1}{n^2} \sum_{k_1\in[K]} \mathbf{p}_{k_1} \left(\mathbf{d}_{k_1} \left((1-\mathbf{p}_{k_1}) \mathbb{E}\left[\frac{1+X(n-1,\mathbf{p}_{k_1})}{s(1+X(n-1,\mathbf{p}_{k_1}))} - \frac{X(n-1,\mathbf{p}_{k_1})}{s(X(n-1,\mathbf{p}_{k_1})}\right] \right) \right. \\
&\qquad \left. + \sum_{k\neq k_1} \mathbf{d}_k \mathbf{p}_k \mathbb{E}\left[\frac{X(n-1,\mathbf{p}_k)}{s(X(n-1,\mathbf{p}_k))} - \frac{1+X(n-1,\mathbf{p}_k)}{s(1+X(n-1,\mathbf{p}_k))}\right] \right)^2 \\
&= \frac{1}{n^2} \sum_{k_1\in[K]} \mathbf{p}_{k_1} \left(\mathbf{d}_{k_1} \left((1-\mathbf{p}_{k_1}) \mathbb{E}\left[\bar{w}(X(n-1,\mathbf{p}_{k_1}))\right] \right) - \sum_{k\neq k_1} \mathbf{d}_k \mathbf{p}_k \mathbb{E}\left[\bar{w}(X(n-1,\mathbf{p}_k))\right] \right)^2
\end{aligned}
$$

where

$$\bar{w}(x) \triangleq \frac{1+x}{s(1+x)} - \frac{x}{s(x)}.$$

By assumption **D.4**, either $\bar{w}(x) < 0$ (**D.4(a)**) or $\bar{w}(x) > 0$ (**D.4(b)**). In either case, the k_1 and $k \neq k_1$ terms have opposite signs. For any x, also by **D.4**, $|\bar{w}(x)| \leq 1$. Using the fact that if a and b have the same signs, then $|a - b| \leq \max\{|a|, |b|\}$,

$$\begin{aligned}
\sigma_1^2 &= \frac{1}{n^2} \sum_{k_1 \in [K]} \mathbf{p}_{k_1} \left(\mathbf{d}_{k_1} \left((1 - \mathbf{p}_{k_1}) \mathbb{E}\left[\bar{w}(X(n-1, \mathbf{p}_{k_1})) \right] \right) - \sum_{k \neq k_1} \mathbf{d}_k \mathbf{p}_k \mathbb{E}\left[\bar{w}(X(n-1, \mathbf{p}_k)) \right] \right)^2 \\
&\leq \frac{1}{n^2} \sum_{k_1 \in [K]} \mathbf{p}_{k_1} \max \left\{ \mathbf{d}_{k_1}(1 - \mathbf{p}_{k_1}), \sum_{k \neq k_1} \mathbf{d}_k \mathbf{p}_k \right\}^2 \qquad (|a-b| \leq \max\{|a|, |b|\},\ |\bar{w}(x)| \leq 1) \\
&\leq \frac{1}{n^2} \sum_{k_1 \in [K]} \mathbf{p}_{k_1} \max \left\{ (1 - \mathbf{p}_{k_1}), \sum_{k \neq k_1} \mathbf{p}_k \right\}^2 \qquad (\mathbf{d}_k \leq 1) \\
&= \frac{1}{n^2} \sum_{k_1 \in [K]} \mathbf{p}_{k_1} (1 - \mathbf{p}_{k_1})^2 \\
&\leq \frac{1}{n^2}.
\end{aligned}$$

Thus, by (43),

$$\mathrm{Var}(\widehat{U}(\mathcal{K})) \leq \frac{1}{S} + \Theta(S^{-2}) \approx \frac{1}{S}.$$

Because our overall estimator is the average of 25 such U-statistics, the standard errors on our utility estimates are on the order of

$$\sqrt{\frac{1}{2000 \cdot 25}} \approx .0045.$$

In principle, we could also directly compute confidence intervals for $\widehat{U}(\mathcal{K})$ using tail bounds for U-statistics (see Hoeffding [38, eq. (5.7)]).

Other implementation details. Our implementation uses nucleus sampling [39] with $p = 0.95$, meaning we discard tokens the maximal number of tokens whose total probability falls under $1 - p$. We set a limit of 6 new tokens, finding that around 5.5% of samples do not terminate (i.e., hit a `STOP` token). This is a conservative estimate, since some of those samples would have hit a `STOP` token on the 7th token, meaning that sample would have been terminal anyways.

We choose a maximum temperature $\tau_{\max}$ independently for each language model (see Table 1). As a rough guide, we use the curves in Figure 7 to determine the temperature at which a language model exhibits diminishing returns for welfare with $n = 15$. We ensure that in all downstream analyses, neither equilibrium nor optimal strategies require choosing the maximal available temperature.

Compute resources. We run all of our experiments on NVIDIA Tesla V100 GPUs. All told, generating 2,000 samples per (language model, temperature, instance) combination and

validating them with qwen2.5 takes around 1 week of GPU time (which can be parallelized at the (language model, instance) level).[16] Downstream analyses of equilibria take on the order of tens of hours of computation and are highly parallelizable on a CPU cluster.

C.3 Empirical distributional measures

Measuring weighted inversions To compute $\mathsf{WI}(\mathbf{p}^{(\tau)} \,\|\, \mathbf{p}^{(\tau')})$, we would need full access to $\mathbf{p}^{(\tau)}$ and $\mathbf{p}^{(\tau')}$. This would require us to compute the probability that a GAIT outputs each possible (multi-token) answer, which is computationally intractable. We instead sample from each $\mathbf{p}^{(\tau)}$, but we have more information than just the samples: when we compute a sample, we do so by explicitly taking the softmax for each successive token. We therefore know the precise probability of generating each answer in our sample. Let $\mathcal{K}^{(\tau)}$ be the set of tokens observed in the sample at temperature τ. For the purposes of measuring inversions, we can think of our sample as providing $\{\mathbf{p}_k : k \in \mathcal{K}^{(\tau)}\}$. In general, our sample is incomplete, meaning we are missing some of the probability mass. Figure 12 shows how much of the overall mass we capture in our samples.

Let $\hat{\mathbf{p}}^{(\tau)}$ be partial distribution covered by our sample:

$$\hat{\mathbf{p}}_k^{(\tau)} \triangleq \begin{cases} \mathbf{p}_k^{(\tau)} & k \in \mathcal{K}^{(\tau)} \\ \perp & \text{otherwise} \end{cases}$$

These probabilities won't sum to 1, since we don't observe the full distribution. But if we restrict ourselves to $\hat{\mathbf{p}}^{(\tau)}$, we can compute weighted inversions on these observed distributions. Critically, we can only do so for answers that appear in both $\hat{\mathbf{p}}^{(\tau)}$ and $\hat{\mathbf{p}}^{(\tau')}$, i.e., $k \in \mathcal{K}^{(\tau)} \cap \mathcal{K}^{(\tau')}$. By convention, let $d_{\mathrm{TV}}(\mathbf{p}, \mathbf{q})$ ignore all k such that $\mathbf{p}_k = \perp$ or $\mathbf{q}_k = \perp$. Then,

$$\widehat{\mathsf{WI}}(\hat{\mathbf{p}}^{(\tau)} \,\|\, \hat{\mathbf{p}}^{(\tau')}) = \frac{\min_{\mathbf{p} \in \Delta(\pi(\hat{\mathbf{p}}^{(\tau)}))} d_{\mathrm{TV}}(\mathbf{p}, \hat{\mathbf{p}}^{(\tau')})}{d_{\mathrm{TV}}(\hat{\mathbf{p}}^{(\tau)}, \hat{\mathbf{p}}^{(\tau')})}.$$

While $\widehat{\mathsf{WI}}$ can either be larger or smaller than WI, we believe it to be an underestimate because the contribution of each k to WI appears to be larger for smaller values of $\mathbf{p}_k^{(\tau)}$. We are more likely to fail to sample low-probability k's, so $\widehat{\mathsf{WI}}$ is likely to underestimate WI. The absolute value of the difference between WI and $\widehat{\mathsf{WI}}$ is bounded by the probability mass we fail to capture in $\mathbf{p}^{(\tau)}$ and $\mathbf{p}^{(\tau')}$, which we show in Figure 12.

We mitigate this to an extent as follows. Observe that for $\tau < \tau' < \tau''$, there can exist k such that $k \in \mathcal{K}^{(\tau)} \cap \mathcal{K}^{(\tau'')}$ but $k \notin \mathcal{K}^{(\tau')}$. When this occurs, we linearly interpolate to estimate $\hat{\mathbf{p}}_k^{(\tau')}$, which improves our coverage over answers. As long as τ and τ'' are fairly close to one another, the linear approximation is a reasonable one.

Finally, we need to actually compute $\min_{\mathbf{p} \in \Delta(\pi(\mathbf{p}^{(\tau)}))} d_{\mathrm{TV}}(\mathbf{p}, \hat{\mathbf{p}}^{(\tau')})$. This is effectively an isotonic regression problem with ℓ_1 loss, for which a number of efficient algorithms exist [3, 79, 89]. We choose the dynamic programming-based approach of Rote [79] for its simplicity.

[16] In principle, we could also parallelize across temperatures or even across independent samples for a given (language model, temperature, instance) combination, but our implementation caches language model calls for each (language model, instance) combination.

Inversions in d. To validate whether Assumption 2 holds in our Scattergories setting, we might hope to use a similar approach to measuring the extent to which $\mathbf{d}$ is decreasing for a given language model. In particular, let Π^* be the set of permutations on K that rank all valid answers above all invalid answers. Then let

$$\mathsf{D}(\mathbf{p}) \triangleq \min_{\pi \in \Pi^*, \mathbf{p}' \in \Delta(\pi)} d_{\mathrm{TV}}(\mathbf{p}, \mathbf{p}')$$

be a measure of distance between a distribution and the "nearest" distribution that yields a perfect ordering. As long as $\mathbf{p}$ assigns 0 probability to sufficiently many valid answers, this is approximately equal to the probability mass $\mathbf{p}$ assigns to k such that $\mathbf{d}_k = 0$. Thus, it suffices to simply estimate $\mathbb{E}_{k \sim \mathbf{p}}[\mathbf{d}_k]$, which we show in Figure 5b.

System

You are a helpful assistant. Answer in as few words as possible, with no explanations.

User

We are playing a game of Scattergories. I will give you a letter and a category, and you will give me a word or short phrase that starts with that letter and matches the category. For example, if I say "Fruit" and "A," you could say "Apple" or "Apricot."

Assistant

I understand

User

Letter: C
Category: Countries

Assistant

Canada

User

Letter: X
Category: Instruments

Assistant

Xylophone

User

Letter: {{letter}}
Category: {{category}}

Assistant

(a) Prompt used for answer generation

System

You are a helpful assistant. Answer in as few words as possible, with no explanations.

User

You are judging a Scattergories game. I will give you a category and an answer. You will tell me whether that answer fits the given category. You are strict but fair. You do not accept incomplete answers. Answer with either the word 'yes' or 'no'.

Assistant

I understand

User

Category: Things that are red
Answer: blood

Assistant

yes

User

Category: Animals
Answer: rock

Assistant

no

User

Category: {{category}}
Answer: {{answer}}

Assistant

(b) Prompt for answer verification

Figure 13: Prompts