Likes, Budgets, and Equilibria: Designing Contests for Socially Optimal Advertising

Sayantika Mandal
Indian Institute of Technology Bombay
22d0379@iitb.ac.in

Swaprava Nath
Indian Institute of Technology Bombay
swaprava@iitb.ac.in

Abstract

We study competitive advertising on a social network. Each of m firms allocates a budget across consumers, and a platform decides, through a contest success function (CSF), what share of each consumer's attention every firm captures from the budgets placed on that consumer; awareness then spreads through the network. We model awareness and budgets on two timescales and bound how many fast steps awareness needs to settle between two budget updates. On the slow timescale, we show that best-response dynamics converge at a geometric rate to a Nash equilibrium whenever the firms' utilities are strongly concave and smooth, and that the standard ratio-form CSFs satisfy these conditions. Their equilibrium, however, need not maximize welfare. A firm's spending lowers its rivals' shares of attention, leading to the individual optima being different from the social optima, and we show that a constant per-unit charge on budget cannot close this gap in general. We then give a sufficient condition for the equilibrium to coincide with the welfare optimum. To meet it, we multiply each firm's ratio share by a function of the total spending at the consumer, a class of CSFs that we call scaled ratio CSFs (SRCSFs), and we show that within this class the condition holds exactly when this function is a power of the total spending; we call the resulting CSF the Power RCSF. The Power RCSF alone still falls short in two ways. Its equilibrium is optimal only on instances where the firms exhaust their budgets, and we remove this restriction with a constant per-unit charge that is budget-balanced. Its curvature also becomes unbounded as spending at a consumer approaches zero, which breaks our convergence guarantee, and we restore the guarantee by requiring every firm to spend a small minimum amount on each consumer. With these two additions, best-response dynamics converge to a unique, symmetric equilibrium that maximizes p-mean social welfare for every $p \leqslant 1$ at the same time, so the Price of Anarchy is unity.

1 Introduction

The rise of AI-powered platforms has created a new class of competitive markets in which several providers compete for user attention through a common intermediary. Consider an AI marketplace where providers of competing coding assistants, search tools, or recommendation engines each invest resources to be visible to users. For each user, the platform decides whose content to surface and how prominently, based on how much each provider has invested in reaching that user. A user who adopts a product then recommends it to friends and contacts, so awareness spreads further through the social network. The same structure appears in any market where an intermediary controls initial exposure and peer influence shapes further adoption.

The platform's allocation rule therefore plays a dual role. For a provider, it determines how much of an investment turns into user attention. For a regulator or a platform designer, it is a design instrument, and a natural question is whether the rule can be chosen so that competition among providers produces collectively desirable outcomes instead of a wasteful race for visibility.

We study this question formally. Each user is a vertex of a social network, and each provider's share of a user's attention is an *awareness level* in $[0, 1]$. Awareness comes from direct investment by the provider and from peer influence through the network. Following Skaperdas [41], we model the platform's rule as a *contest success function* (CSF), which maps the providers' investments at a user into the shares of that user's attention they capture. Providers play a budget allocation game, and a planner chooses the CSF. Our central question is:

> *can a planner choose the CSF so that the competitive game has a Nash equilibrium that is socially optimal, and that the providers actually reach?*

In practice, AI-driven ad systems such as Google's Performance Max, Meta's Advantage+, and the sponsored answers now shown in ChatGPT and Copilot decide what share of a user's attention each advertiser receives as a function of competing budgets, which we model as a CSF.

Our answer is a qualified yes, and the qualifications are the substance of the paper. The standard ratio-form CSFs of contest theory lead to well-behaved competition, but they cannot guarantee an efficient outcome: part of what a firm gains from its own spending is taken from a rival, so a firm's private return to spending exceeds the social return. We show that this gap does not have to vanish for the equilibrium to be optimal, and we identify a CSF, together with two simple instruments, under which equilibrium and optimum coincide on every instance.

1.1 Our Contributions

We extend the awareness model of Bimpikis et al. [12] to m firms in a two-timescale setting, where consumers update their awareness quickly and firms revise their budgets slowly, and we bound the number of fast steps that awareness needs to settle (Theorem 1). On the slow timescale, we show that best-response dynamics (BRD) converge at a geometric rate to a Nash equilibrium (NE) whenever the induced utilities are strongly concave and smooth (Theorem 2). We then turn to the planner's problem. We first identify a subclass of the standard ratio-form CSFs, which we call Regular RCSFs, that satisfies these conditions. Their equilibrium need not maximize welfare, because a firm's spending lowers its rivals' attention shares, and we show that a constant per-unit charge does not fix this (Theorem 3, Remark 3). We next give a sufficient condition for the equilibrium to coincide with the welfare optimum (Proposition 1). To satisfy it, we scale each firm's ratio share by a function of the total spending at the consumer; we call this class *scaled ratio CSFs* (SRCSFs), and we show that within it the condition holds exactly when the scaling function is a power of the total spending, which gives the CSF we call the *Power RCSF* (Lemma 5, Proposition 2). The Power RCSF still needs two additions. Its equilibrium is optimal only when the firms exhaust their budgets, which the designer cannot control, so we add a constant per-unit charge whose revenue is returned to the firms (Proposition 3, Proposition 4). Its curvature is also unbounded near zero spending, which breaks the convergence guarantee, so we require every firm to spend a small minimum amount on each consumer, and we show that this floor leaves the equilibrium unchanged (Proposition 5, Lemma 7, Lemma 8). With both additions, BRD converge to a unique, symmetric equilibrium that maximizes p-mean social welfare for every $p \leqslant 1$ at the same time, so the Price of Anarchy (PoA) equals one (Theorem 4, Corollary 1). Table 1 summarizes how each stage of the design adds one guarantee.

2 Related Work

Our work draws on several literatures: competition for influence on networks, the theory of contests, models of competitive marketing, the efficiency of proportional allocation rules, and learning

Design	Sufficient for BRD convergence	Unique, symmetric NE	NE is welfare optimal	Results
Regular RCSF	✓	✓	✗ not guaranteed; a constant charge does not fix it	Lemmas 3 and 4, Theorem 3, Remark 3
Power RCSF	✗ curvature unbounded near zero spending	✓	only if budgets are exhausted at the optimum; never when $n \leqslant mC$	Propositions 1 to 3 and 5, Lemma 6
Power RCSF + charge τ and BB payment	✗	✓	✓	Proposition 4
Power RCSF + charge τ and BB payment + participation floor $\varepsilon < \bar{\varepsilon}$	✓	✓	✓ for every $p \leqslant 1$, so PoA $= 1$	Lemmas 7 and 8, Theorem 4, Corollary 1

Table 1: Guarantees obtained at each stage of the CSF design. The fast-timescale dynamics converge for every CSF (Theorem 1), and best-response dynamics converge whenever the induced utilities have the SB property (Theorem 2). Each row adds one instrument to remove the failure in the row above it. BB denotes the budget-balanced redistribution of the charge in Equation (20).

dynamics in games. We discuss them in turn.

Competitive influence on networks. Influence maximization was introduced by Kempe et al. [28], and its competitive version studies firms that choose seed consumers in a shared network [2, 11, 25]. Tzoumas et al. [44] analyze a two-player diffusion game under a linear threshold model, and Zuo et al. [48] study an online version with bandit feedback. Goyal and Kearns [25] also bound the inefficiency of equilibria in their contagion model. These works model competition through the choice of seeds; we instead study the allocation of continuous budgets.

Budget allocation and advertising on networks. A parallel line of work studies firms that spread budgets over consumers. Fazeli and Jadbabaie [20, 21], Fazeli et al. [22] study the trade-off between seeding and product quality, and Masucci and Silva [32, 33], Varma et al. [45, 46], and Ansari et al. [3] study related allocation games. Maehara et al. [31] show that a multi-advertiser budget allocation game is a potential game with a Price of Anarchy of two. Closest to us, Bimpikis et al. [12] model targeted advertising by two firms through the awareness dynamics that we extend to m firms, and they find that in equilibrium firms invest inefficiently high, with an inefficiency that grows with the centrality of the consumers they target. Our impossibility result for the Regular RCSF is consistent with this finding; our contribution is to ask how the allocation rule can be designed to remove the inefficiency.

Welfare and fairness in networked marketing. Several works add welfare or fairness objectives to influence maximization. Lu et al. [30] and Wang et al. [47] study fair allocation of seeds or budgets among advertisers, Rahmattalabi et al. [38] and Becker et al. [8] design fair influence

maximization through welfare objectives and randomization, and Banerjee et al. [4, 5] maximize consumer welfare under diffusion models. These works optimize allocations from the point of view of a planner or a host, and do not study how firms that choose their own spending can be led to an efficient equilibrium.

Contest success functions and welfare. Ratio-form contests go back to Tullock [43], and Skaperdas [41] characterizes them axiomatically; Konrad [29] surveys the theory. Corchón and Dahm [15] derive CSFs from the decision problem of a contest administrator, and Szidarovszky and Okuguchi [42] establish existence and uniqueness of equilibrium in rent-seeking games. A recurring conclusion of this literature is that competition for a fixed prize dissipates resources. The closest work to our design question is Chung [14], who studies contests in which the prize grows with aggregate effort, as the market-expansion factor $f(B_i)$ of our SRCSFs does; he shows that the resulting positive externality does not offset the negative one, and that aggregate effort remains socially excessive. We show that a specific functional form, the iso-elastic one, makes the two forces exactly offsettable by a constant charge. In industrial organization, Dixit and Norman [16] show that competitive advertising is socially excessive, and the classical remedy for such externalities is a corrective charge [7, 37]. Our charge is of this type, with a budget-balanced redistribution similar to that of Murhekar et al. [36].

Market-share attraction models. In marketing, a brand's market share is often modeled as its attraction divided by the total attraction of all brands [9], which is the ratio form of contest theory. Game-theoretic analyses of these models characterize equilibrium marketing efforts [27, 35], and Federgruen and Yang [23] study generalized attraction models in which total demand also responds to aggregate attraction. Our SRCSF has the same factorized structure, with f playing the role of total demand. To our knowledge, the observation that a constant elasticity of f is what aligns the competitive equilibrium with the welfare optimum does not appear in this literature.

Efficiency of proportional allocation. Johari and Tsitsiklis [26] show that when a divisible resource is allocated in proportion to bids, the equilibrium attains at least three quarters of the optimal aggregate utility, and that this bound is tight. Our setting differs in that the total amount of attention is itself endogenous, which is what allows a suitable design to reach full efficiency.

Dynamics and uniqueness in games and contests. Uniqueness of equilibrium in concave games under diagonal strict concavity goes back to Rosen [39], and projection methods for the resulting variational inequalities converge linearly under strong monotonicity and Lipschitz continuity [19]. Murhekar et al. [36] use gradient dynamics of this kind to reach welfare-maximizing equilibria in federated learning. In contests, Ghosh [24] and Elkind et al. [17] analyze best-response dynamics in Tullock contests with convex costs, and Ewerhart and Quartieri [18] prove uniqueness of equilibrium in contests with incomplete information. We contribute convergence guarantees for m firms in a two-timescale model, and identify the regularity failure that makes a participation floor necessary for the iso-elastic CSF.

3 Preliminaries

We consider a market with m competing firms (players) denoted by $\mathscr{M} = \{1, \ldots, m\}$, each having an equal and fixed marketing budget $C > 0$[1] that they strategically allocate across n vertices (consumers) denoted by $\mathscr{N} = \{1, 2, \ldots, n\}$, in a directed weighted social network $\mathscr{G} = (\mathscr{N}, \mathscr{E})$. The adjacency matrix $\mathscr{E} = (e_{ij})_{i,j \in \mathscr{N}}$ is row-stochastic, with $e_{ii} = 0$ for all i, and e_{ij} representing the influence of vertex j on vertex i. We consider *awareness* as a cardinal metric of a consumer's attention towards a particular firm. For each firm $s \in \mathscr{M}$ and each consumer $i \in \mathscr{N}$, we define an *awareness* $x_{s,i} \in [0, 1]$ that captures how prominently firm s occupies consumer i's attention relative to its competitors. We allow $\sum_s x_{s,i} \leqslant 1$, with the residual mass $1 - \sum_s x_{s,i}$ corresponding to the consumer being unaware of any of the m firms. Each firm aims to maximize the overall awareness of its brand across the network, subject to its budget constraints.

3.1 A Two-Timescale Communication Model

Real advertising markets evolve at two distinct rates. Consumers interact with each other (via word-of-mouth, social-media reshares, and direct content consumption) on a fast timescale, often hourly or daily. Firms revise their advertising budgets and targeting strategies on a slower timescale, often quarterly or annually. Hence, we model the interaction between firms' strategic decisions and the evolution of customer awareness using a two-timescale framework.

Slow timescale (strategy updates). The slow timescale is indexed by $k = 0, 1, 2, \ldots$. At the start of each period k, firm s commits to a budget vector $\mathbf{b}_s(k) = (b_{s,1}(k), \ldots, b_{s,n}(k))^\top \in \mathbb{R}^n_{\geqslant 0}$ representing the allocation of its marketing budget across the n vertices in the network, and this allocation remains fixed for the duration of period k. The set of feasible budget vectors for firm s is $\mathcal{B}_s = \{\mathbf{b}_s \in \mathbb{R}^n_{\geqslant 0} : \sum_i b_{s,i} \leqslant C\}$. The joint feasible set $\mathcal{B} = \prod_s \mathcal{B}_s$ is a convex and compact and we denote the joint strategy profile of all firms by the matrix $\mathbf{b} \coloneqq (\mathbf{b}_s, \mathbf{b}_{-s})$ where $\mathbf{b}_{-s}$ denotes the allocation vectors of all competing firms other than s.

Fast timescale (awareness dynamics). Within each period k, awareness evolves on a fast timescale indexed by $t = 0, 1, \ldots, T$. At every fast-timescale step, each consumer receives an influence that mixes two sources: the awareness of its neighbours, with weight $\alpha \in (0, 1)$, and the direct effect of the firms' advertising, with weight $1 - \alpha$. We call α the *peer-influence weight*. Our model is inspired by the awareness framework of Bimpikis et al. [12], but unlike theirs it is deterministic. The awareness of vertex i toward firm s at fast-timescale step t of period k is updated as

$$x_{s,i}(t) = \frac{1}{t}\Big[\alpha \sum_{j \in \mathscr{N}_i} e_{ij}\, x_{s,j}(t-1) + (1-\alpha)\, h_{s,i}(\mathbf{b}(k))\Big] + \Big(1 - \frac{1}{t}\Big) x_{s,i}(t-1), \qquad t \geqslant 1, \tag{1}$$

where $\mathscr{N}_i$ is the set of neighbours of i, and $x_{s,i}(0) \in [0, 1]$ is the awareness at the start of the period. The *contest success function* (CSF) $h_{s,i}(\mathbf{b}(k)) \coloneqq h(b_{s,i}(k), \mathbf{b}_{-s,i}(k))$ specifies the share of consumer i's attention that firm s captures through direct advertising, given the budgets all firms have invested in i; in contest-theoretic terms, it is firm s's success probability at i. We require $\sum_{s \in \mathscr{M}} h_{s,i}(\mathbf{b}(k)) \leqslant 1$, and we use the shorthand $h_{s,i}(\mathbf{b})$ when the dependence on k is clear.

The update (1) gives weight $1 - 1/t$ to the awareness accumulated so far and weight $1/t$ to the newest influence. This has a simple interpretation. Writing $y_{s,i}(t')$ for the bracketed term in (1) at step t' and multiplying (1) by t gives $t\, x_{s,i}(t) = y_{s,i}(t) + (t-1)\, x_{s,i}(t-1)$, and unrolling this identity

[1] The equal-budgets assumption ensures that the equilibrium can be the welfare maximizer. Our convergence results do not use it. See Remark 5 for more details.

down to $t = 1$ yields

$$x_{s,i}(t) = \frac{1}{t} \sum_{t'=1}^{t} y_{s,i}(t').$$

Thus the awareness at step t is the average of all the influences received so far, each counting equally. We use this to model awareness as cumulative exposure: an influence received early in the period is not forgotten when newer ones arrive. The dynamics also respect the interpretation of awareness as a share of attention. Each $y_{s,i}(t')$ is a convex combination of the numbers $x_{s,j}(t'-1)$ and $h_{s,i}$, because the weights αe_{ij} and $1-\alpha$ are non-negative and sum to one by row-stochasticity of $\mathscr{E}$. Summing over firms, $\sum_s y_{s,i}(t') \leqslant \alpha \max_j \sum_s x_{s,j}(t'-1) + (1-\alpha)\sum_s h_{s,i} \leqslant 1$ whenever $\sum_s x_{s,j}(t'-1) \leqslant 1$ for all j. By induction on t, if every $x_{s,i}(0)$ lies in $[0,1]$ and $\sum_s x_{s,i}(0) \leqslant 1$ for every i, then the same holds at every step t.

Steady-state awareness. Fix a slow-timescale period k and a firm s, and write $\mathbf{h}_s := \mathbf{h}_s(\mathbf{b}(k))$ for the vector $(h_{s,i}(\mathbf{b}(k)))_{i \in \mathscr{N}}$, which is constant throughout the period. In vector form, (1) reads $\mathbf{x}_s(t) = \frac{1}{t}\big[\alpha \mathscr{E} \mathbf{x}_s(t-1) + (1-\alpha)\mathbf{h}_s\big] + \big(1 - \frac{1}{t}\big)\mathbf{x}_s(t-1)$. A vector $\mathbf{x}$ is left unchanged by this update if and only if $\mathbf{x} = \alpha \mathscr{E}\mathbf{x} + (1-\alpha)\mathbf{h}_s$, since the terms involving $1/t$ cancel. Because $\mathscr{E}$ is row-stochastic, the spectral radius of $\alpha\mathscr{E}$ is at most $\alpha < 1$, so $I - \alpha\mathscr{E}$ is invertible and this equation has exactly one solution,

$$\mathbf{x}_s^*(k) = (1-\alpha)(I - \alpha\mathscr{E})^{-1}\mathbf{h}_s\big(\mathbf{b}_s(k), \mathbf{b}_{-s}(k)\big), \tag{2}$$

We call $\mathbf{x}_s^*(k)$ the *steady-state awareness* of firm s in period k. It has the same form as the steady state derived by Bimpikis et al. [12, Theorem 1] for two firms. The next theorem shows that the fast dynamics converge to it from any starting point, and it quantifies how many fast-timescale steps this takes.

Theorem 1 (Convergence of the fast-timescale dynamics)**.** *Fix a period k and a firm s, and let $\mathbf{x}_s^*$ be the steady state in Equation (2). For every initial awareness $\mathbf{x}_s(0) \in [0,1]^n$, the iterates of Equation (1) satisfy the following:*

(a) Convergence with an explicit rate. *For every $T \geqslant 1$,*

$$\|\mathbf{x}_s(T) - \mathbf{x}_s^*\|_\infty \;\leqslant\; P_T \,\|\mathbf{x}_s(0) - \mathbf{x}_s^*\|_\infty, \qquad \textit{where} \quad P_T := \prod_{t=1}^{T}\Big(1 - \frac{1-\alpha}{t}\Big) \;\leqslant\; (T+1)^{-(1-\alpha)}.$$

In particular, $\mathbf{x}_s(T) \to \mathbf{x}_s^*$ as $T \to \infty$.

(b) Number of steps. *For every accuracy $\epsilon_0 \in (0,1]$, running $T \geqslant \epsilon_0^{-1/(1-\alpha)}$ fast-timescale steps guarantees $\|\mathbf{x}_s(T) - \mathbf{x}_s^*\|_\infty \leqslant \epsilon_0$. Thus $T = O\big(\epsilon_0^{-1/(1-\alpha)}\big)$ steps suffice.*

(c) Tightness. *The factor P_T cannot be improved: it equals $\Gamma(T+\alpha)/\big(\Gamma(\alpha)\Gamma(T+1)\big)$, which behaves like $T^{-(1-\alpha)}/\Gamma(\alpha)$ as $T \to \infty$, and the bound in (a) holds with equality whenever $\mathbf{x}_s(0) - \mathbf{x}_s^*$ has equal coordinates.*

Proof sketch. Subtracting the fixed-point equation from (1) shows that the error $\mathbf{e}(t) := \mathbf{x}_s(t) - \mathbf{x}_s^*$ obeys the linear recursion $\mathbf{e}(t) = A(t)\,\mathbf{e}(t-1)$ with $A(t) := (1 - \frac{1}{t})I + \frac{\alpha}{t}\mathscr{E}$. Because $\mathscr{E}$ is non-negative and row-stochastic, every row of $A(t)$ is non-negative and sums to $1 - (1-\alpha)/t$, which is therefore its ℓ_∞ operator norm. Multiplying these norms over $t = 1, \ldots, T$ gives the factor P_T. Using $\log(1-x) \leqslant -x$ and the bound $\sum_{t=1}^{T} 1/t \geqslant \log(T+1)$ on the harmonic sum then gives

$P_T \leqslant (T+1)^{-(1-\alpha)}$, and (b) follows because both $\mathbf{x}_s(0)$ and $\mathbf{x}_s^*$ lie in $[0,1]^n$. For (c), a vector with equal coordinates is an eigenvector of every $A(t)$, and a Gamma-function product identity evaluates P_T exactly. The complete proof is in Appendix A.1. □

In the rest of the paper, we take the awareness of firm s at the end of period k to be the steady state $\mathbf{x}_s^*(k)$, and Theorem 1 quantifies the error of this idealization after T fast steps. The bound also gives explicit guidance for spacing the slow-timescale epochs. When peer influence is weak (small α), the exponent $1/(1-\alpha)$ is close to 1, and about $1/\epsilon_0$ fast steps suffice. As $\alpha \to 1$, the exponent grows, so more fast steps are needed for the same accuracy: stronger peer influence means that each new influence moves awareness less, and the dynamics settle more slowly. For every α, the required number of steps grows as the desired accuracy ϵ_0 decreases.

Remark 1 (On the equal weighting of past influences). The polynomial rate in Theorem 1 is a consequence of the equal-weight memory of Equation (1). If recent influences were instead weighted more heavily, for instance by replacing $1/t$ with a constant weight $\gamma \in (0,1]$, the same argument would give the error factor $\big(1-\gamma(1-\alpha)\big)^T$, that is, geometric convergence in $O(\log(1/\epsilon_0))$ steps. The equal-weight model is therefore the harder case for the separation of timescales, and the guarantee of Theorem 1 carries over, with a better rate, to discounted variants.

3.2 Firms' Utilities and the Awareness-Competition Game (ACG)

We define the utility (payoff) of a firm as the difference between its consolidated awareness over the network and the cost incurred through advertising [17, 24]. This formulation captures the trade-off between maximizing awareness and minimizing expenditure. The utility of firm s at the end of period k is

$$u_s(\mathbf{b}(k)) = \mathbf{1}^\top \mathbf{x}_s^*(k) - \mathbf{1}^\top \mathbf{b}_s(k) = \sum_{i \in \mathscr{N}} M_i \, h_{s,i}(\mathbf{b}(k)) - \sum_{i \in \mathscr{N}} b_{s,i}(k), \tag{3}$$

where we drop the period index k when it is clear, and

$$M_i := (1-\alpha)\big[\mathbf{1}^\top (I - \alpha \mathscr{E})^{-1}\big]_i. \tag{4}$$

The weight M_i is a scaled version of the *absorption centrality* of vertex i used by Bimpikis et al. [12]: it measures how much of the attention captured at i eventually reaches the whole network. A firm s strategically allocates its budget across vertices to maximize this utility by balancing awareness gains against advertising costs in a competitive environment. We call the induced game between firms the *awareness competition game* (ACG).

3.3 Design Desiderata

We now define the equilibrium concept, the planner's objective, and the measure of inefficiency that connects them.

Definition 1 (Pure-strategy Nash equilibrium, PSNE)**.** A budget profile $\mathbf{b}^* = (\mathbf{b}_s^*, \mathbf{b}_{-s}^*) \in \mathcal{B}$ is a *PSNE* if $u_s(\mathbf{b}_s^*, \mathbf{b}_{-s}^*) \geqslant u_s(\mathbf{b}_s, \mathbf{b}_{-s}^*)$ for every $s \in \mathscr{M}$ and every $\mathbf{b}_s \in \mathcal{B}_s$.

We shorten this to Nash equilibrium (NE) where it is clear from context. While the competing firms reach an NE, the goal of a planner (network policymaker or regulator) is to ensure some *social welfare*. We consider the generalized p-mean welfare function, which provides a unifying framework for different welfare criteria.

Definition 2 (p-mean social welfare, 6)**.** Let $\mathcal{B}_{\geqslant 0} := \{\mathbf{b} \in \mathcal{B} : u_s(\mathbf{b}) \geqslant 0 \ \forall s \in \mathscr{M}\}$. For $\mathbf{b} \in \mathcal{B}_{\geqslant 0}$, the p-mean social welfare is

$$W_p(\mathbf{b}) = \left(\frac{1}{m}\sum_{s\in\mathscr{M}} u_s(\mathbf{b})^p\right)^{\frac{1}{p}}, \tag{5}$$

where $p \in (-\infty, 1]$ controls the trade-off between equity and efficiency. For $p < 0$ we set $W_p(\mathbf{b}) := 0$ whenever some $u_s(\mathbf{b}) = 0$, the limiting value.

Special cases include (a) utilitarian welfare ($p = 1$): $W_1(\mathbf{b}) = \frac{1}{m}\sum_s u_s(\mathbf{b})$, (b) egalitarian welfare ($p \to -\infty$): $W_{\text{egal}}(\mathbf{b}) = \min_s u_s(\mathbf{b})$, and (c) Nash social welfare ($p \to 0$): $W_{\text{NSW}}(\mathbf{b}) = (\prod_s u_s(\mathbf{b}))^{1/m}$. The planner optimizes over $\mathcal{B}_{\geqslant 0}$ because u_s^p is not real for $u_s < 0$ and non-integer p and a firm can guarantee $u_s \geqslant 0$ by spending nothing, so every NE of the ACG lies in $\mathcal{B}_{\geqslant 0}$. The set $\mathcal{B}_{\geqslant 0}$ is compact and W_p is continuous on it, so the maxima exist. Note that W_1 is defined on all of $\mathcal{B}$, and we use it there.

Definition 3 (Price of Anarchy)**.** Let $\mathcal{B}^\star \subseteq \mathcal{B}_{\geqslant 0}$ be the set of Nash equilibria of the ACG. The *Price of Anarchy* for W_p is

$$\text{PoA} = \frac{\max_{\mathbf{b}\in\mathcal{B}_{\geqslant 0}} W_p(\mathbf{b})}{\min_{\mathbf{b}\in\mathcal{B}^\star} W_p(\mathbf{b})},$$

the ratio of the optimal welfare to the welfare at the worst equilibrium. A larger value means more inefficiency from competition, and PoA $= 1$ means that every equilibrium is optimal.

A Price of Anarchy of one is only meaningful if the firms can find the equilibrium. Before asking which CSFs make the equilibrium efficient, we therefore study how firms reach an equilibrium in the first place.

4 Best Response Dynamics of ACGs and Convergence to Nash

We now ask how the firms reach an equilibrium. A firm does not know its rivals' future budgets, and in practice it observes only how its own utility responds to small changes in its spending. A natural model of play is therefore an adjustment process. In each slow-timescale period, every firm observes the current budget profile, and the awareness it induces, and moves its own budget in a direction that increases its utility, taking its rivals' budgets as given. This is the idea behind *best-response dynamics* (BRD) and we use a gradient-based form of it, in which every firm simultaneously takes a small step along the gradient of its own utility.

In terms of the two timescales of Section 3, at the start of period k each firm commits to $\mathbf{b}_s(k)$, awareness evolves on the fast timescale to the steady state in Equation (2), which takes the number of fast-timescale steps given by Theorem 1, and at the start of period $k+1$ each firm updates its budget using the utilities realized in period k. We show in this section that under certain conditions, the ACG converges to an NE when firms play BRD. This section makes no reference to a planner and focuses solely on the behavior of firms competing under a CSF whose induced utilities satisfy these conditions. The CSFs which make that equilibrium efficient are addressed in Sections 5 to 7.

Feasible sets. For a lower boundary $\ell \geqslant 0$ with $n\ell \leqslant C$, define

$$\mathcal{D}_s^\ell := \Big\{\mathbf{b}_s \in \mathbb{R}^n : \ b_{s,i} \geqslant \ell \ \ \forall i \in \mathscr{N}, \ \ \textstyle\sum_{i\in\mathscr{N}} b_{s,i} \leqslant C\Big\}, \qquad \mathcal{D}^\ell := \textstyle\prod_{s\in\mathscr{M}} \mathcal{D}_s^\ell. \tag{6}$$

Throughout this section $\ell = 0$, so that $\mathcal{D}_s^\ell = \mathcal{B}_s$ and $\mathcal{D}^\ell = \mathcal{B}$. Each $\mathcal{D}_s^\ell$ is non-empty, closed and convex, and hence so is $\mathcal{D}^\ell$.

Algorithm 1 Best Response Dynamics

Input: Initial budgets $\mathbf{b}_s(0) \in \mathcal{D}_s^\ell \ \forall s \in \mathscr{M}$; step size $\gamma > 0$; number of iterations K_ϵ.

1: **for** each iteration $k = 0, 1, \ldots, K_\epsilon - 1$ **do**

2: Compute $\mathbf{g}_s(\mathbf{b}(k), \boldsymbol{\mu}_s(k)) = \nabla_{\mathbf{b}_s} u_s(\mathbf{b}(k)) + \boldsymbol{\mu}_s(k)$ for all $s \in \mathscr{M}$ ▷ see Equation (7)

3: Update: $\mathbf{b}_s(k+1) \leftarrow \mathbf{b}_s(k) + \gamma\, \mathbf{g}_s(\mathbf{b}(k), \boldsymbol{\mu}_s(k))$, for all $s \in \mathscr{M}$

4: **end for**

Output: $\mathbf{b}^* \leftarrow \mathbf{b}(K_\epsilon)$ ▷ Nash equilibrium

Algorithm 1 implements BRD for the ACG. Every firm moves its budget in the direction of the gradient of its own utility, corrected by a term $\boldsymbol{\mu}_s(k)$ that keeps the new budget feasible. The corrected gradient of firm s is $\mathbf{g}_s(\mathbf{b}(k), \boldsymbol{\mu}_s(k)) \in \mathbb{R}^n$ with ith component

$$g_{s,i}(\mathbf{b}(k), \boldsymbol{\mu}_s(k)) = \frac{\partial u_s(\mathbf{b}(k))}{\partial b_{s,i}} + \mu_{s,i}(k), \tag{7}$$

where $\boldsymbol{\mu}_s(k) \in \mathbb{R}^n$ is the smallest correction, in Euclidean norm, for which the update in line 3 stays feasible:

$$\boldsymbol{\mu}_s(k) \in \arg\min \Big\{ \|\boldsymbol{\mu}\|_2 : \ \mathbf{b}_s(k) + \gamma\big(\nabla_{\mathbf{b}_s} u_s(\mathbf{b}(k)) + \boldsymbol{\mu}\big) \in \mathcal{D}_s^\ell \Big\}. \tag{8}$$

Lemma 1 (proof in Appendix B.1) below shows that the correction in (8) is well defined, and it explains what the correction does. In particular, if the uncorrected gradient step is feasible, then $\boldsymbol{\mu}_s(k) = \mathbf{0}$; otherwise, $\boldsymbol{\mu}_s(k)$ provides only the correction needed to restore feasibility.

Lemma 1. *For every $\mathbf{b} \in \mathcal{D}^\ell$ and every $s \in \mathscr{M}$, the minimization in (8) admits a unique solution, denoted by $\boldsymbol{\mu}_s(\mathbf{b})$. The correction used in Algorithm 1 is therefore given by $\boldsymbol{\mu}_s(k) = \boldsymbol{\mu}_s(\mathbf{b}(k))$.*

Proof sketch. Since the new budget equals the uncorrected step plus $\gamma\boldsymbol{\mu}$, minimizing $\|\boldsymbol{\mu}\|_2$ over feasible updates is the same as finding the feasible budget nearest to the uncorrected step, that is, the Euclidean projection onto the non-empty, closed, convex set $\mathcal{D}_s^\ell$, which exists and is unique. See Appendix B.1. □

For analytical convenience, we vectorize the joint budget matrix $\mathbf{b}$ into a column vector $\mathbf{b} \in \mathbb{R}^{mn}$ and stack the corrections and corrected gradients as $\boldsymbol{\mu}(k) := [\boldsymbol{\mu}_s(k)]_{s\in\mathscr{M}}$ and $\mathbf{g}(\mathbf{b}(k), \boldsymbol{\mu}(k)) := [\mathbf{g}_s(\mathbf{b}(k), \boldsymbol{\mu}_s(k))]_{s\in\mathscr{M}} \in \mathbb{R}^{mn}$. The following Lemma 2 (proof in Appendix B.2) relates the stationary points of BRD to Nash equilibria.

Lemma 2. *Suppose that each u_s is concave and continuously differentiable in $\mathbf{b}_s$. Then a profile $\mathbf{b}^\star \in \mathcal{D}^\ell$ satisfies $\mathbf{g}_s(\mathbf{b}^\star, \boldsymbol{\mu}_s(\mathbf{b}^\star)) = \mathbf{0}$ for all $s \in \mathscr{M}$ if and only if $\mathbf{b}^\star$ is a pure-strategy Nash equilibrium of the game played on $\mathcal{D}^\ell$.*

Proof sketch. The corrected gradient vanishes exactly when the projected step returns to $\mathbf{b}_s^\star$, which, by the characterization of projections onto convex sets, is the first-order optimality condition for firm s's best-response problem. Concavity makes that condition sufficient as well as necessary. See Appendix B.2. □

The above lemma clarifies that the proposed best-response dynamics cannot become stationary at a non-equilibrium profile, where some firm still has a profitable deviation, including a reallocation of its budget across consumers, as the dynamics do not converge to a trivial zero-gradient state.

4.1 Convergence of Best Response Dynamics

In this section, we present the conditions on the utility functions and analyze the convergence of the BRD to an NE. We define the block Hessian $G(\mathbf{b}) = [G_{s_1,s_2}]_{s_1,s_2 \in \mathscr{M}} \in \mathbb{R}^{mn \times mn}$ by $[G_{s_1,s_2}]_{i,j} = \partial^2 u_{s_1}(\mathbf{b}) / (\partial b_{s_2,j}\, \partial b_{s_1,i})$.

Definition 4 (SB property)**.** The utilities $\{u_s(\mathbf{b})\}_{s \in \mathscr{M}}$ satisfy the *SB property* on a convex and compact set $\mathcal{D}^\ell$ if there exist constants $\lambda, L > 0$ such that, for every $\mathbf{b} \in \mathcal{D}^\ell$,

1. *Strong concavity:* The matrix $G(\mathbf{b})$ is λ-negative semi-definite, i.e., the symmetric part of $G(\mathbf{b})$ satisfies $\frac{1}{2}\big(G(\mathbf{b}) + G(\mathbf{b})^\top\big) + \lambda I_{mn} \preceq 0$.
2. *Bounded operator norm:* $\|G(\mathbf{b})\|_{\mathrm{op}} \leqslant L$, where $\|G\|_{\mathrm{op}} := \inf\{c > 0 : \|G\mathbf{v}\| \leqslant c\|\mathbf{v}\|\ \forall \mathbf{v}\}$.

These are the canonical strong-concavity and smoothness conditions used in the analysis of best-response dynamics and gradient-based algorithms [36]. *Strong concavity* ensures each firm's utility is strongly concave in its own strategy, yielding a stable equilibrium, while a *bounded norm* guarantees that utility gradients do not change abruptly, so that fixed-step gradient updates converge. Since the SB property is a property of the induced utilities, it must be verified separately for each CSF considered.

With agents having SB property, the BRD converges to an NE, as formalized in the following theorem. Note that the convergence of BRD to a point where $\|\mathbf{g}(\mathbf{b}, \boldsymbol{\mu})\|_2$ is arbitrarily small implies convergence to an NE in the limit (Lemma 2).

Theorem 2 (Convergence to NE)**.** *Suppose the agent utilities satisfy the SB property (Definition 4) on a convex feasible set $\mathcal{D}^\ell$ as in (6) containing all iterates. Then Algorithm 1 converges to a Nash equilibrium $\mathbf{b}^\star$. More precisely, for any $\epsilon > 0$ and any initial point $\mathbf{b}(0) \in \mathcal{D}^\ell$, after $K_\epsilon > \frac{2L^2}{\lambda^2} \log\big(\|\mathbf{g}(\mathbf{b}(0), \boldsymbol{\mu}(0))\|_2 / \epsilon\big)$ iterations the gradients satisfy $\|\mathbf{g}(\mathbf{b}(K_\epsilon), \boldsymbol{\mu}(K_\epsilon))\|_2 < \epsilon$, when the step size is $\gamma = \lambda / L^2$.*

Proof sketch. All iterates lie in $\mathcal{D}^\ell$ by the construction of (8). We track the ℓ_2 norm of the update direction $\mathbf{g}(\mathbf{b}(k), \boldsymbol{\mu}(k))$. Keeping the correction $\boldsymbol{\mu}(k)$ fixed, a first-order expansion in integral form gives $\mathbf{g}(\mathbf{b}(k+1), \boldsymbol{\mu}(k)) = (I + \gamma \bar{G}(k))\, \mathbf{g}(\mathbf{b}(k), \boldsymbol{\mu}(k))$, where $\bar{G}(k)$ is the average of G along the segment from $\mathbf{b}(k)$ to $\mathbf{b}(k+1)$, which lies in $\mathcal{D}^\ell$ by convexity and so inherits both SB bounds (Definition 4). Substituting these bounds and choosing $\gamma = \lambda / L^2$ gives the contraction inequality $\|\mathbf{g}(\mathbf{b}(k), \boldsymbol{\mu}(k))\|_2^2 \leqslant \big(1 - \lambda^2/L^2\big)^k \|\mathbf{g}(\mathbf{b}(0), \boldsymbol{\mu}(0))\|_2^2$, implying exponential convergence. Thus, we can conclude that Algorithm 1 reaches a Nash equilibrium $\mathbf{b}^\star$ with $\|\mathbf{g}(\mathbf{b}(K_\epsilon), \boldsymbol{\mu}(K_\epsilon))\|_2 < \epsilon$ after $K_\epsilon = \mathcal{O}\big(\log \frac{1}{\epsilon}\big)$ iterations. The complete proof is deferred to Appendix B.3. □

Since each step of Algorithm 1 is γ times the corrected gradient, Theorem 2 implies that the step lengths shrink geometrically, so the iterates form a Cauchy sequence and converge to some $\mathbf{b}^\star$ in the closed set $\mathcal{D}^\ell$. The corrected gradient is continuous in $\mathbf{b}$, so it vanishes at $\mathbf{b}^\star$, and by Lemma 2, $\mathbf{b}^\star$ is a Nash equilibrium. In particular, the SB property guarantees that a Nash equilibrium exists.

Theorem 2 alone does not produce a welfare guarantee. An equilibrium may exist and be reached by BRD while remaining inefficient relative to the social optimum. The following sections examine this welfare gap and develops the conditions under which it can be eliminated.

5 Equilibrium Need Not Be Efficient

We now introduce the planner, whose goal is to choose a CSF that shapes firms' incentives so that their strategic behavior leads to a welfare-maximizing outcome. We first show by example

that competition under a natural CSF can settle at a very inefficient equilibrium. Section 6 then shows that the standard ratio-form CSFs, although they make BRD converge to a unique symmetric equilibrium, cannot guarantee efficiency, and identifies the reason. From that reason we derive a sufficient condition for efficiency (Section 6.1), find the CSF that satisfies it (Section 7), and add two instruments that make it work on every instance: a budget-balanced charge (Section 7.1) and a participation floor (Section 7.2). Section 8 then brings these pieces together to establish that the resulting CSF admits a unique, symmetric, welfare-maximizing Nash equilibrium.

Example 1 (An inefficient equilibrium)**.** Consider two firms A and B competing over a fully connected network of five agents, with the influence matrix $\mathscr{E}$ in Table 2. The peer-influence weight $\alpha = 0.5$, which gives the centralities $M = (0.810, 1.210, 1.131, 0.958, 0.891)$. Each firm has budget $C = 2$, and the platform uses the ratio-form CSF [41]:

$$h_{s,i}(\mathbf{b}) = \frac{e^{2b_{s,i}} - 1}{\sum_{r \in \mathscr{M}} \left(e^{2b_{r,i}} - 1\right) + 0.5}.$$

Table 2: Influence matrix $\mathscr{E}$ for Example 1.

	1	2	3	4	5
1	0	0.3901	0.3003	0.2456	0.0640
2	0.0669	0	0.3715	0.2578	0.3037
3	0.0149	0.7005	0	0.1534	0.1313
4	0.1407	0.2334	0.4025	0	0.2234
5	0.4340	0.0989	0.2072	0.2599	0

Let $\left(\mathbf{b}_A^{\text{opt}}, \mathbf{b}_B^{\text{opt}}\right) = \arg\max_{\mathbf{b}} W_p(\mathbf{b})$ denote the optimal budgets. The NE is the symmetric profile $\mathbf{b}_A^\star = \mathbf{b}_B^\star = (0.230, 0.448, 0.399, 0.302, 0.268)$, with utilities $u_A(\mathbf{b}^\star) = u_B(\mathbf{b}^\star) = 0.320$, so that $W_p(\mathbf{b}^\star) = 0.320$ for every $p \in (-\infty, 1]$. The utilitarian welfare-maximizing profiles are $\mathbf{b}_A^{\text{opt}} = (0, 0.368, 0.349, 0, 0.279)$ and $\mathbf{b}_B^{\text{opt}} = (0.253, 0, 0, 0.300, 0)$, which give $W_1 = 0.813$. At the egalitarian optimum $\mathbf{b}_A^{\text{opt}} = (0.162, 0, 0.348, 0, 0.279)$ and $\mathbf{b}_B^{\text{opt}} = (0.067, 0.369, 0, 0.300, 0)$, the utilities $u_A \approx u_B \approx 0.803$, so $W_p \approx 0.803$ for every $p \in (-\infty, 1]$. The Nash social welfare optimum $\mathbf{b}_A^{\text{opt}} = (0.253, 0, 0, 0.300, 0.279)$ and $\mathbf{b}_B^{\text{opt}} = (0, 0.369, 0.348, 0, 0)$ similarly gives $W_{\text{NSW}} = 0.811$.

In this example, both firms engage in strictly positive spending at equilibrium and contest every consumer, while the planner avoids this duplicated spending by assigning each consumer to one firm. The resulting duplication leads to a substantial welfare loss, with PoA $\gtrsim 2.5$. This inefficiency stems from the structure of the existing CSF, which encourages firms to spend resources that do not contribute to higher total welfare. To overcome this, we next explore how an appropriately designed CSF can align equilibrium behavior with collective social welfare, starting with the most widely used class.

6 First Attempt: the Regular RCSF

The natural place to look for a good CSF is the class most commonly studied in contest theory. We focus on the *ratio-form* CSF, in which a firm's success at a vertex is its own score relative to the total score of all firms competing for that vertex. This is the canonical axiomatic CSF of Skaperdas [41], characterized by monotonicity, anonymity, and consistency, and is the standard choice in the

contest literature. So our first attempt is to find whether a CSF from this class ensures both BRD convergence to NE and welfare maximization.

Following Bimpikis et al. [12], we additively regularize the denominator by a constant $\delta > 0$, which models an outside option that absorbs unallocated attention and keeps the CSF well defined even when all budgets vanish. This gives the *Ratio CSF* (RCSF),

$$h_{s,i}(\mathbf{b}) = \frac{f(b_{s,i})}{\sum_{r \in \mathscr{M}} f(b_{r,i}) + \delta}, \tag{9}$$

where f is a scalar *generator* applied to each firm's budget allocation $b_{s,i}$. A platform that scores ads by f(budget) and allocates impressions in proportion to scores implements an RCSF.

Not every RCSF is well behaved. We single out a subclass, defined by three conditions on f over the interval $[0, C]$ of possible budgets, which suffices for BRD to converge to a unique symmetric equilibrium. The CSF of Example 1 is itself an RCSF, but its generator $e^{kb_{s,i}} - 1$ is convex, so it does not belong to this subclass.

Definition 5 (Regular RCSF)**.** An RCSF is *Regular* if its generator f satisfies the following conditions, where all suprema are over budget $b \in [0, C]$.

(R1) *Concave increasing generator:* $f(0) = 0$, $f'(b) > 0$, and $f''(b) < 0$.

(R2) *Bounded slope and curvature:* $\sup f' < \infty$ and $\sup |f''| < \infty$.

(R3) *Curvature dominates coupling:*

$$\sup \frac{f'}{-f''} \cdot \sup f' \;<\; \frac{\delta}{m-1}. \tag{10}$$

The three conditions are easy to check. Condition (R1) asks for a concave, increasing, normalized generator, Condition (R2) keeps the second derivatives of the utilities bounded, and Condition (R3) asks the generator's own curvature to be strong enough, relative to its slope and the number of rivals, to dominate the coupling between firms. For the saturating exponential $f(b) = 1 - e^{-\beta b}$ with $\beta > 0$, we have $f'/(-f'') \equiv 1/\beta$ and $\sup f' = \beta$, so the left side of Equation (10) equals 1 and the CSF is Regular whenever $\delta > m - 1$. The logarithmic generator $f(b) = \log(1 + \beta b)$ is Regular when $\delta > (m-1)(1 + \beta C)$. The power generator $f(b) = b^q$ [43] satisfies Condition (R1), but for $q < 1$ its slope and curvature are unbounded near zero, and for $q = 1$, since $f'' = 0$, it has no curvature at all.

The next two lemmas establish the equilibrium properties of the Regular RCSF. Lemma 3 (proof in Appendix C.2) shows that BRD converges, and Lemma 4 (proof in Appendix C.3) characterizes the equilibrium it reaches.

Lemma 3 (Regular RCSF induces the SB property)**.** *If the CSF is a Regular RCSF (Definition 5), then the induced utilities satisfy the SB property (Definition 4) on $\mathcal{D}^\ell = \mathcal{B}$. Consequently, by Theorem 2 the best response dynamics (Algorithm 1) converges to a Nash equilibrium $\mathbf{b}^\star$.*

Proof sketch. Each firm's success at consumer i depends only on the budgets placed on i, so after reordering coordinates the block Hessian is block diagonal, with one $m \times m$ block per consumer. Condition (R3) makes every block strictly diagonally dominant with a negative diagonal, which by Gershgorin's circle theorem gives strong concavity, and Condition (R2) bounds every entry, which gives the norm bound. See Appendix C.2. □

Lemma 4 (Unique and symmetric NE for Regular RCSF)**.** *For an RCSF whose generator f satisfies Condition (R1), the ACG has a unique Nash equilibrium, and it is symmetric:* $\mathbf{b}_1^\star = \cdots = \mathbf{b}_m^\star$.

Proof sketch. Strict concavity of each firm's utility, together with symmetry, compactness, and convexity of the strategy spaces, gives existence of a symmetric NE by [13]. To rule out asymmetric NE, the Regular RCSF satisfies the marginal dominance inequality: if $b_{p,i} > b_{q,i}$, then $\partial h_{p,i}/\partial b_{p,i} < \partial h_{q,i}/\partial b_{q,i}$. Combined with the first-order conditions of the firm $M_i \partial h_{s,i}/\partial b_{s,i} = 1 + \lambda_s - \mu_{s,i}$, this implies that a firm spending more at some vertex has a strictly smaller budget multiplier. There are two cases for an asymmetric profile, and in both cases the first-order conditions together with the marginal dominance inequality give a contradiction. Hence every NE is symmetric. At a symmetric profile, the marginal return $\partial h_{s,i}/\partial b_{s,i}$ is strictly decreasing in the common budget, so the symmetric first-order conditions have exactly one solution, giving a unique symmetric NE. See Appendix C.3. □

The Regular RSCF therefore secures a well-behaved equilibrium. Nevertheless, this class is fundamentally incapable of implementing socially optimal equilibria. Before establishing this impossibility result in the following Theorem 3, we introduce the notation and auxiliary quantities used in its proof and in the analysis that follows.

Definition 6 (Private and social marginal returns, and the wedge)**.** At a symmetric profile with common value $b_{s,i} = t_i$ for every firm $s \in \mathscr{M}$, define

$$a(t_i) := \left.\frac{\partial h_{s,i}}{\partial b_{s,i}}\right|_{\text{sym}}, \qquad d(t_i) := \left.\frac{\partial h_{r,i}}{\partial b_{s,i}}\right|_{\text{sym}} (r \neq s), \qquad \psi(t_i) := \sum_{r \in \mathscr{M}} \left.\frac{\partial h_{r,i}}{\partial b_{s,i}}\right|_{\text{sym}} = a(t_i) + (m-1)d(t_i).$$

Here $a(t_i)$ is the *private marginal return* to firm s from its own spending at vertex i, $d(t_i)$ is the marginal effect of that spending on a single rival, and $\psi(t_i)$ is the *social marginal return*, the marginal effect on the sum of all m firms' success probabilities at i. The *externality wedge* is

$$w(t_i) := a(t_i) - \psi(t_i) = -(m-1)\, d(t_i). \tag{11}$$

By symmetry, these quantities do not depend on which firm is considered. The private and social marginal returns determine the respective first-order conditions of the firm and the planner, and the wedge $w(t_i)$ measures the difference between these marginal incentives. Theorem 3 shows that this wedge can prevent the Nash equilibrium from being welfare optimal.

Theorem 3 (Regular RCSFs cannot ensure that equilibrium is welfare optimal)**.** *Let h be a Regular RCSF. Then there exist instances of the awareness-competition game under this CSF for which the unique symmetric Nash equilibrium $\mathbf{b}^\star$ of Lemma 4 fails to maximize the utilitarian welfare W_1 over the feasible set $\mathcal{B}$.*

Proof sketch. Utilitarian welfare W_1 depends on the budgets only through the per-consumer totals of budgets $b_{s,i}$ and scores $f(b_{s,i})$, and by Jensen's inequality and concavity of f, symmetrizing any feasible profile weakly increases W_1 while preserving aggregate spending. Hence it suffices to compare the symmetric NE with the welfare optimum among symmetric profiles. At a symmetric profile, the firm and planner's first-order conditions involve the private marginal return $a(t_i)$ and social marginal return $\psi(t_i)$, and for $t_i > 0$ are $M_i a(t_i) = 1 + \lambda_s$ and $M_i \psi(t_i) = 1 + \hat{\lambda}$, respectively. For the Regular RCSF, the wedge $w(t_i) = a(t_i) - \psi(t_i) > 0$, and the ratio $a(t_i)/\psi(t_i) = 1 + (m-1)f(t_i)/\delta$ is strictly increasing in t_i. If the equilibrium funds vertices with different centralities, the two first-order conditions therefore cannot hold simultaneously. If all centralities are equal, we

construct an instance whose welfare optimum is positive and leaves the budget unexhausted. Then both budget multipliers are zero, which would require $a(t) = \psi(t)$, contradicting the strictly positive wedge. Thus, in either case, the unique symmetric NE is not welfare-optimal. The complete proof is in Appendix C.4. □

Remark 2 (The failure is structural). The wedge is the familiar *business-stealing externality*: part of what a firm gains from its spending is taken from a rival, so it counts in the firm's own utility but cancels in total welfare. Theorem 3 shows that it is strictly positive for every Regular RCSF and every symmetric profile with positive budgets, so the failure does not depend on a particular generator f or regularizer δ. The wedge would vanish only for $\delta = 0$, but then $\sum_s h_{s,i} \equiv 1$ and, whenever budgets are exhausted, total welfare no longer depends on the allocation, which makes the design problem vacuous.

Theorem 3 shows that the canonical Regular RCSF cannot guarantee that the Nash equilibrium reached by BRD is welfare-optimal. It already fails for the utilitarian welfare criterion $p = 1$. Since our design goal is stronger, namely an equilibrium maximizing W_p for every $p \in (-\infty, 1]$, ruling out the utilitarian case is sufficient to rule out the full-family requirement. We therefore characterize the conditions under which the firms' and planner's incentives can align.

6.1 The alignment criterion of social optimality and NE

The externality wedge (11) need not vanish for the planner's and firms' incentives to be aligned as requiring $w \equiv 0$ would imply $d \equiv 0$, meaning that a firm's spending has no effect on any rival, which is no longer a contest. For the equilibrium to coincide with the welfare optimum, the stationarity conditions in the KKT conditions of each firm's best-response problem must be satisfied by the profile that satisfies the stationarity conditions in the KKT conditions of the planner's problem. This requires an alignment condition between the private and social marginal returns. The following proposition gives the exact condition for this alignment in terms of the social and private marginal returns $\psi(t_i)$ and $a(t_i)$ at symmetric profiles,[2] as defined in Definition 6.

Proposition 1 (Alignment criterion for social optimality and NE)**.** *Consider a CSF for which the utility of a firm s, i.e. u_s is concave in $\mathbf{b}_s$, and whose private and social marginal returns (Definition 6) at symmetric profiles $b_{s,i} = t_i$ satisfy the proportional-marginals condition*

$$\psi(t_i) = c\, a(t_i), \tag{12}$$

for a constant $c \in (0, 1)$. Let $\mathbf{b}^{\text{opt}}$ be a symmetric profile that maximizes W_1, $b^{\text{opt}}_{s,i} = t^{\text{opt}}_i$ for every firm $s \in \mathscr{M}$. Then:

(i) if $\sum_i t^{\text{opt}}_i = C$, that is, every firm exhausts its budget, then $\mathbf{b}^{\text{opt}}$ is a Nash equilibrium;

(ii) conversely, if $\mathbf{b}^{\text{opt}} \neq \mathbf{0}$ and $\mathbf{b}^{\text{opt}}$ is a Nash equilibrium, then $\sum_i t^{\text{opt}}_i = C$.

Proof sketch. Substituting $\psi = ca$ into the planner's condition $M_i\psi(t_i) - 1 = \hat{\lambda} - \hat{\mu}_i$ and dividing by c yields the firm's condition $M_i a(t_i) - 1 = \lambda_s - \mu_{s,i}$ with the relabeled multipliers $\lambda_s = (\hat{\lambda} + 1 - c)/c$ and $\mu_{s,i} = \hat{\mu}_i/c$. All requirements on these multipliers are inherited from the planner's, except complementary slackness for the budget: since $c < 1$, the multiplier λ_s is strictly positive, so

[2] The proof of Theorem 3 shows that, for the Regular RCSF, replacing a feasible profile with its corresponding symmetric profile weakly increases welfare. This motivates considering symmetric profiles in the alignment analysis. The symmetry restriction is subsequently verified for the CSF identified by this analysis.

the budget must be exhausted. Concavity makes the firm's conditions sufficient. For (ii), an unexhausted budget would force both budget multipliers to zero, and then $\psi = a$ at any funded consumer, which contradicts $c < 1$. The complete proof is in Appendix C.5. □

Remark 3. Since $\psi = a+(m-1)d$, the condition $\psi = ca$ is equivalent to $d(t_i)/a(t_i) = (c-1)/(m-1)$ for all t_i. For $c < 1$ the cross effect d stays negative, so the contest retains genuine rivalry. The condition therefore requires the ratio of the cross effect to the own effect to remain constant across spending levels. For the Regular RCSF, this ratio is $-f(t_i)/[(m-1)f(t_i)+\delta]$, which depends on t_i for every $\delta > 0$, with the only constant-ratio case $\delta = 0$ rendering the design problem vacuous as noted in Remark 2. This gives a second view of why the Regular RCSF fails.

Proposition 1 separates the design of a CSF that aligns NE and social optimality, into two tasks. The first is to construct a CSF with proportional private and social marginal returns, which is determined by the functional form of the CSF and is addressed in the next section. The second is to ensure that the firms exhaust their full budget at the welfare optimum, which, as we will see, depends on the underlying instance rather than on the CSF and cannot be controlled by the designer.

7 Second Attempt: Scaled ratio CSFs

We now construct a CSF satisfying the alignment criterion in Proposition 1. Since a constant ratio d/a is unattainable inside the Ratio CSF form (9), we enlarge the class while preserving its ratio structure as it makes the rule a contest. Specifically, we scale each firm's contest share by a function of the total budget spent on the consumer.

Definition 7 (Scaled Ratio CSF)**.** A *Scaled Ratio CSF* (SRCSF) has the form

$$h_{s,i}(\mathbf{b}) = f(B_i) \cdot \frac{b_{s,i}}{B_i}, \qquad B_i := \sum_{r \in \mathscr{M}} b_{r,i}, \tag{13}$$

for an increasing, concave function $f : [0, mC] \to [0, 1]$ with $f(0) = 0$. When $B_i = 0$ we set $h_{s,i}(\mathbf{b}) := 0$, which is the continuous extension.

Summing (13) over firms gives $\sum_s h_{s,i} = f(B_i)$, so $f(B_i)$ is the total attention the platform releases at vertex i when firms spend B_i there. The second factor is the classical Tullock share [43]. The two factors separate the two economic forces cleanly: f governs how much attention competition creates in total, and the share governs how it is divided. At a symmetric profile, $b_{s,i} = t_i = B_i/m$, we derive ψ, a (Definition 6) as follows:

$$\psi(t_i) = f'(B_i), \qquad a(t_i) = \frac{f'(B_i)}{m} + \frac{(m-1)f(B_i)}{mB_i}, \tag{14}$$

and their ratio is

$$\frac{\psi(t_i)}{a(t_i)} = \frac{m\, f'(B_i)\, B_i}{f'(B_i)\, B_i + (m-1)f(B_i)}, \tag{15}$$

which can be expressed in terms of the elasticity of f, which we define below.

Definition 8 (Elasticity and iso-elasticity)**.** The *elasticity* of a differentiable positive function f at B_i is the ratio of the relative change in $f(B_i)$ to the relative change in B_i [40]:

$$e(B_i) := \frac{d \log f(B_i)}{d \log B_i} = \frac{f'(B_i)\, B_i}{f(B_i)}$$

We call f *iso-elastic* if $e(B_i)$ does not depend on B_i.

In words, $e(B_i)$ measures the percentage change in total attention $f(B_i)$ released at a vertex i resulting from a one-percent change in aggregate budget spending B_i. Iso-elasticity means that this percentage is constant across budget levels.

Lemma 5 (Iso-elasticity is necessary and sufficient)**.** *Within the SRCSF class* (13)*, the proportional-marginals condition* (12) *of Proposition 1 holds if and only if f is iso-elastic, that is, $f(B_i) = \kappa B_i^{\eta}$ for constants $\kappa > 0$ and $\eta \in (0, 1]$.*

Proof sketch. Dividing Equation (15) through by $f(B_i)$ gives $\psi/a = m\, e(B_i)/(e(B_i) + m - 1)$, a strictly increasing function of the elasticity, so the ratio is constant exactly when the elasticity is. Solving $f'/f = \eta/B_i$ gives $f = \kappa B_i^{\eta}$, and the requirements on f force $\eta \in (0, 1]$. The complete proof is in Appendix C.6. □

Lemma 5 identifies the functional form required for alignment within the SRCSF class. In particular, the total attention released by the platform at a vertex must have constant elasticity with respect to aggregate budget spending at that vertex. We therefore restrict attention to this iso-elastic subclass and we name the resulting CSF after its power-function form.

Definition 9 (Power RCSF)**.** For $\eta \in (0, 1)$ and $0 < \kappa \leqslant (mC)^{-\eta}$, the *Power-scaled Ratio CSF* (Power RCSF) is the SRCSF (13) with $f(B_i) = \kappa B_i^{\eta}$, that is

$$h_{s,i}(\mathbf{b}) = \kappa\, b_{s,i}\, B_i^{\eta-1}. \tag{16}$$

The cap on κ ensures $\sum_s h_{s,i} = \kappa B_i^{\eta} \leqslant \kappa (mC)^{\eta} \leqslant 1$ since $B_i \leqslant mC$ and $\eta > 0$,. The restriction $\eta < 1$ keeps the rule a contest, because then a rival's spending strictly reduces a firm's success:

$$\frac{\partial h_{s,i}}{\partial b_{s,i}} = \kappa B_i^{\eta-2}\big[B_i + (\eta - 1) b_{s,i}\big] > 0, \qquad \frac{\partial h_{s,i}}{\partial b_{r,i}} = \kappa(\eta - 1)\, b_{s,i}\, B_i^{\eta-2} < 0 \quad (r \neq s), \tag{17}$$

The parameter η interpolates between two degenerate extremes. As $\eta \to 0$, the total attention κB_i^{η} becomes constant and the rule reduces to the classical Tullock contest, where firms' spending only redistributes attention among competitors. As $\eta \to 1$, we get $h_{s,i} = \kappa b_{s,i}$, so a firm's success depends only on its own spending, eliminating rivalry among firms.

Proposition 2 (Power RCSF has proportional marginals)**.** *For the Power RCSF (Definition 9), the induced utility of each firm is concave in its own strategy, and the proportional marginals condition* (12) *of Proposition 1 holds with*

$$c = \frac{m\eta}{m + \eta - 1} \in (0, 1). \tag{18}$$

Proof sketch. Concavity follows from the sign of the second derivative of $h_{s,i}$ in $b_{s,i}$. Since the elasticity of κB_i^{η} is identically η, the formula in the proof of Lemma 5 gives $c = m\eta/(\eta + m - 1)$, which lies in $(0, 1)$ because $\eta < 1$. The complete proof is in Appendix C.7. □

Remark 4. For the Power RCSF, the total attention $\sum_s h_{s,i} = f(B_i) = \kappa B_i^{\eta}$ depends on aggregate spending when $\eta \in (0, 1)$, ensuring total welfare remains allocation-dependent. Unlike Regular RCSF, the Power RCSF therefore attains the constant ratio c in Equation (18) and hence the constant ratio $d/a = (c - 1)/(m - 1)$ without falling into the degenerate case of $\delta = 0$ flagged in Remarks 2 and 3.

Proposition 2 establishes the proportional marginals condition identified in Proposition 1. The remaining condition to be verified is full budget exhaustion at the optimum. The next proposition characterizes exactly when full budgets are exhausted for the Power RCSF, and shows that the condition depends on the primitives of the game instance such as the number of vertices n, the total budget C and the scaled network centralities M_i, none of which the designer controls.

Proposition 3 (When are budgets exhausted?)**.** *Fix the Power RCSF with parameters $\eta \in (0,1)$ and $\kappa \leqslant (mC)^{-\eta}$, and write $\theta := 1/(1-\eta) > 1$ and $\|M\|_\theta := (\sum_i M_i^\theta)^{1/\theta}$. Then:*

(i) *The symmetric welfare-maximizing profile exhausts every firm's budget if and only if*

$$\sum_{i \in \mathcal{N}} (\kappa \eta M_i)^\theta \;\geqslant\; mC. \tag{19}$$

(ii) *If (19) holds for some admissible κ, then $\eta \|M\|_\theta \geqslant mC$.*

(iii) *Consequently, if $n \leqslant mC$, then* no *admissible pair (κ, η) satisfies (19): on such instances no Power RCSF can align equilibrium with the welfare optimum.*

Proof sketch. Utilitarian welfare depends only on the totals B_i and equals $\frac{1}{m}[\kappa \sum_i M_i B_i^\eta - \sum_i B_i]$, whose unconstrained maximizer is $B_i = (\kappa \eta M_i)^\theta$; the budget is exhausted exactly when these totals add up to at least mC, which is (i). Part (ii) substitutes the largest admissible κ. For (iii), $\|M\|_\theta \leqslant \sum_i M_i = n$ by Lemma 10, so (ii) fails when $n \leqslant mC$; the optimum is then unexhausted and non-zero, and part (ii) of Proposition 1 rules out alignment. The complete proof is in Appendix C.8. □

The above proposition shows that full budget exhaustion cannot be treated as a general consequence of choosing the Power RCSF as condition (19) depends on the underlying instance, rather than the CS design parameters. Thus, even after identifying the appropriate CSF within the SRCSF class, alignment of welfare and NE cannot be guaranteed for all instances, as shown in Proposition 3(iii). This motivates a different approach to restore alignment.

7.1 Adding a Corrective Charge

We have seen that the Power RCSF can align private and social marginal returns (Proposition 2), but the full budget exhaustion condition (19) needed for alignment depends on the ACG instance and cannot be guaranteed by the CSF design alone (Proposition 3). We therefore take a different route and modify the firms' incentives directly. In this paper, we therefore restore alignment via a corrective charge on advertising expenditure, following the standard *Pigouvian* logic of pricing that is prevalent in the literature [7, 37]. The firm's marginal incentive can be aligned with the planner's social marginal return without requiring the firms to exhaust their budgets.

Concretely, let the platform levy a per-unit charge $\tau \geqslant 0$ on advertising spend and redistribute the resulting charge revenue. Let ρ_s be the *rebate* received by firm s and require the scheme to be *budget balanced*,

$$\sum_{s \in \mathcal{M}} \rho_s(\mathbf{b}) \;=\; \tau \sum_{s \in \mathcal{M}} \sum_i b_{s,i}, \tag{20}$$

so that the platform retains no charge revenue. Budget balance is an aggregate condition and does not require each firm to receive its own charge revenue. A concrete scheme satisfying this condition is

$$\rho_s(\mathbf{b}_{-s}) := \frac{\tau}{m-1} \sum_{r \neq s} \sum_i b_{r,i}, \tag{21}$$

which credits firm s with an equal share of of the charge payments made by its rivals. Thus each firm's payment is fully redistributed across the other $m-1$ firms. Schemes of this kind are used for the same purpose by Murhekar et al. [36]. Importantly, ρ_s depends only on rivals' expenditures and is therefore a lump sum from firm s's perspective. Firm s consequently maximizes the *taxed utility*

$$\widetilde{u}_s(\mathbf{b}) = \sum_i M_i h_{s,i}(\mathbf{b}) - (1+\tau)\sum_i b_{s,i} + \rho_s(\mathbf{b}_{-s}), \tag{22}$$

From here on, whenever a charge is in place, the utilities in the welfare function of Definition 2, and in the set $\mathcal{B}_{\geqslant 0}$, are the taxed utilities $\widetilde{u}_s$. By budget balance, the fees and rebates cancel in the sum, $\sum_s \widetilde{u}_s(\mathbf{b}) = \sum_s u_s(\mathbf{b})$, so utilitarian welfare, and with it the planner's first-order conditions, is unchanged, while each firm now faces the unit cost $1+\tau$ instead of 1. The charge therefore moves the firms' first-order conditions toward the planner's without requiring budgets to be exhausted. The next proposition, which is the analogue of Proposition 1, establishes the corresponding alignment result.

Proposition 4 (Alignment of welfare and NE with a corrective charge)**.** *Consider a CSF whose induced taxed utility $\widetilde{u}_s$ (22) is concave in $\mathbf{b}_s$, and which satisfies the proportional-marginals condition (12) for a constant $c \in (0,1)$. Let the firms face a budget-balanced charge at the rate*

$$\tau \;=\; \frac{1}{c} - 1. \tag{23}$$

Then every symmetric profile that maximizes W_1 is a Nash equilibrium, whether or not budgets are exhausted.

Proof sketch. With the charge, the firm's first order condition is $M_i a(t_i) - (1+\tau) = \lambda_s - \mu_{s,i}$. Substituting $\psi = ca$ into the planner's condition and dividing by c gives exactly this, with $1+\tau = 1/c$ and the multipliers $\lambda_s = \hat{\lambda}/c$ and $\mu_{s,i} = \hat{\mu}_i/c$. Unlike in Proposition 1, the budget multiplier now carries no additive offset, so every complementary slackness condition is inherited from the planner's. The complete proof is in Appendix C.9. □

Note that the Regular RCSF fails the proportional marginals condition (12) (Remark 3), and the argument of Theorem 3 goes through with $1+\tau$ in place of 1, so no constant charge restores alignment for that class.

Two features of the charge (23) are useful. It is a constant per-unit rate, which follows from the constant ratio c in Equation (12), and for the Power RCSF it depends only on m and η, not on the network or the budgets, so the platform can compute it without knowing the instance. Before turning to the convergence dynamics, we first characterize the equilibrium under the Power RCSF under a common charge in Lemma 6 below.

Lemma 6 (Unique and symmetric Nash equilibrium for the Power RCSF)**.** *Consider the Power RCSF (Definition 9) with every firm facing the same unit advertising charge $1+\tau \geqslant 1$. Then the induced game admits a unique and symmetric Nash equilibrium. Writing $\theta := 1/(1-\eta)$ and $K := \kappa\, m^{\eta-2}(m+\eta-1)$, the equilibrium $\mathbf{b}^\star$ is*

$$b^\star_{s,i} = t^\star_i = \beta_i(\Gamma^\star), \quad \forall s \in \mathscr{M}; \qquad \beta_i(\Gamma) := \left(\frac{M_i K}{\Gamma}\right)^{\theta}, \tag{24}$$

where $\Gamma^\star \geqslant 1+\tau$ is the unique value with $S(\Gamma^\star) = \min\{S(1+\tau),\, C\}$ and $S(\Gamma) := \sum_i \beta_i(\Gamma)$.

Proof sketch. The argument has the same two parts as Lemma 4, with the unit cost 1 replaced by $1+\tau$, and three details change. First, the marginal dominance inequality is re-derived directly from (17): because swapping two firms' budgets at a vertex leaves B_i unchanged, the firm bidding strictly more there has the strictly smaller marginal return, and the same case analysis on multipliers then rules out every asymmetric profile. Second, at a symmetric profile $a(t) = Kt^{\eta-1}$, so $a'(t) = K(\eta-1)t^{\eta-2} < 0$ and a is strictly decreasing, hence invertible, which gives (24). Third, $a(t) \to +\infty$ as $t \to 0^+$ because $\eta < 1$, so no coordinate of the symmetric candidate can sit at the lower boundary and S is strictly decreasing on all of $(0,\infty)$, whence $\Gamma^\star$ is unique. The complete proof is in Appendix C.10. □

The above lemma establishes the uniqueness and symmetry of the equilibrium under the Power RCSF, and gives an explicit expression for the equilibrium budgets used in the results that follow. We now turn to the question of whether this equilibrium can be reached through best-response dynamics (Algorithm 1).

7.2 BRD to NE: Challenges and the need for a participation floor

Theorem 2 guarantees convergence of BRD only when the utilities have the SB property, and in particular when their second derivatives are bounded. This subsection shows that the Power RCSF fails the SB property on the unrestricted feasible set $\mathcal{B}_s$. The failure arises near $B_i = 0$ from the same constant-elasticity feature of the Power RCSF, $f(B_i) = \kappa B_i^\eta$, that enables alignment of welfare and NE. Since $\eta < 1$, both $f'(B_i) = \kappa\eta B_i^{\eta-1}$ and the curvature of f become unbounded as $B_i \to 0$. The proposition below formalizes this failure and motivates the participation floor introduced thereafter.

Proposition 5 (Unbounded curvature without a floor)**.** *Let $n \geqslant 2$, $m \geqslant 2$, $\eta \in (0,1)$, $\kappa > 0$, and let $\mathcal{B} = \prod_s \{\mathbf{b}_s \geqslant \mathbf{0} : \sum_i b_{s,i} \leqslant C\}$ be the unrestricted feasible set. Let $H_i \in \mathbb{R}^{m\times m}$ denote the vertex-i Hessian block, $(H_i)_{sr} = M_i\, \partial^2 h_{s,i}/(\partial b_{r,i}\partial b_{s,i})$. Then for every vertex i,*

$$\sup_{\mathbf{b}\in\mathcal{B}} \|H_i(\mathbf{b})\|_{\mathrm{op}} = +\infty.$$

Consequently the induced utilities do not satisfy the SB property (Definition 4) on $\mathcal{B}$: no finite bound L can satisfy the bounded-operator-norm condition, and the step size $\gamma = \lambda/L^2$ of Theorem 2 is not defined.

Proof sketch. Let every firm place B_i/m on consumer i and spend the rest elsewhere. A diagonal entry of H_i then has magnitude proportional to $B_i^{\eta-2}$, which diverges as $B_i \to 0$ because $\eta - 2 < 0$. The complete proof is in Appendix C.11. □

The singularity arises as $B_i = \sum_s b_{s,i} \to 0$, when all firms simultaneously do not spend any budget at vertex i, a standard technical difficulty in Tullock-style (ratio-form) contests [43]. By contrast, a zero budget spent by one firm does not create a singularity when at least one rival maintains positive spending. Previous work handles this difficulty either by modifying the best-response correspondence at zero effort [17, 24] or by introducing a temporary positive lower bound and subsequently letting it vanish [18]. In our setting, the singularity is more consequential because the Power RCSF has unbounded curvature as aggregate spending B_i approaches zero. We therefore introduce a participation constraint to keep the process away from this singular region, following the spirit of minimal-effort participation constraints used as a substantive design feature in Megidish and Sela [34].

Definition 10 (Floored game)**.** Every firm must invest at least $\varepsilon > 0$ on every consumer, which can be read as an upfront participation fee of $n\varepsilon$ charged by the platform, and allocate the rest of its budget freely. Each firm's feasible set becomes the *floored set* $\mathcal{B}_s^\varepsilon := \mathcal{D}_s^\varepsilon$ of Equation (6), with $\mathcal{B}^\varepsilon := \prod_s \mathcal{B}_s^\varepsilon$, and Algorithm 1 is run with $\mathcal{D}_s^\ell = \mathcal{B}_s^\varepsilon$, that is, with the lower boundary $\ell = \varepsilon$ in (6) in place of $\ell = 0$. On $\mathcal{B}^\varepsilon$, every consumer receives total spending $B_i \in [m\varepsilon, mC]$.

Restricting the strategy space can alter both the equilibrium and the optimum, and therefore may affect the alignment between them as established in Proposition 4. The next lemma shows that, for a sufficiently small floor, the equilibrium remains unaffected, and provides the corresponding threshold. We say the floor is *inactive* at a profile if every coordinate satisfies $b_{s,i} > \varepsilon$, so that the constraint $b_{s,i} \geqslant \varepsilon$ does not bind and its multiplier is zero.

Lemma 7 (Threshold for an inactive floor)**.** *Let $\theta = 1/(1-\eta)$, $K = \kappa m^{\eta-2}(m+\eta-1)$, and let $\mathbf{b}^\star$ be the equilibrium of Lemma 6 on the unrestricted set. Since the equilibrium is symmetric, write $b_{s,i}^\star = t_i^\star$, $\forall s \in \mathscr{M}$ and define*

$$\bar{\varepsilon} \;=\; \min\left\{ \left(\frac{K\,\min_i M_i}{1+\tau}\right)^{\theta},\;\; C\,\frac{\min_i M_i^\theta}{\sum_j M_j^\theta} \right\} \;>\; 0. \tag{25}$$

Then for every $\varepsilon \in (0, \bar{\varepsilon})$, we have $t_i^\star > \varepsilon$ for all i, so the floor is inactive at the equilibrium. Moreover $\mathcal{B}^\varepsilon$ is non-empty, since $n\varepsilon < C$.

Proof sketch. By Lemma 6, equilibrium spending is proportional to M_i^θ across consumers, so $\min_i t_i^\star$ equals the total equilibrium spending $S(\Gamma^\star)$ times $\min_i M_i^\theta / \sum_j M_j^\theta$. If the budget is not exhausted, the total is $S(1+\tau)$ and this gives the first term of Equation (25); if it is exhausted, the total is C and this gives the second. In each case, the relevant term is also the smaller one. The complete proof is in Appendix C.12. □

Why a floor. The floor is one of several possible remedies for the unbounded curvature of Proposition 5. The role of the participation floor, including the alternatives considered, is discussed in detail in Appendix C.13. The next lemma shows that the Power RCSF satisfies the SB Property (Definition 4) for the floored game (Definition 10).

Lemma 8 (The Power RCSF satisfies SB on the floored set)**.** *On $\mathcal{B}^\varepsilon$ and with any common unit advertising charge $1 + \tau \geqslant 1$, the utilities induced by the Power RCSF satisfy the SB property (Definition 4, with $\mathcal{D}^\ell = \mathcal{B}^\varepsilon$) with the constants*

$$\widetilde{\lambda} := \kappa(1-\eta)\,\chi(m,\eta)\,(mC)^{\eta-2}\min_i M_i \;>\; 0, \qquad \widetilde{L} := 2m\,\kappa(1-\eta)\,(m\varepsilon)^{\eta-2}\max_i M_i \;<\; \infty,$$

where $\chi(m,\eta) := \dfrac{1}{2}\Big[m+\eta-\sqrt{m(m-2\eta+\eta^2)}\Big] > 0$ *for* $\eta \in (0,1)$, $m \geqslant 2$.

Proof sketch. As for the Regular RCSF, it suffices to bound each block H_i. For the Power RCSF, every row of H_i has equal off-diagonal entries, which gives $H_i = c_0(\mathbf{v}\mathbf{1}^\top + B_i I_m)$ with $c_0 := M_i\kappa(\eta-1)B_i^{\eta-3} < 0$, $v_s := B_i - (2-\eta)b_{s,i}$. The symmetric part of $\mathbf{v}\mathbf{1}^\top$ has rank at most two, with eigenvalues $\frac{1}{2}(\mathbf{1}^\top\mathbf{v} \pm \sqrt{m}\|\mathbf{v}\|_2)$ and 0; bounding $\|\mathbf{v}\|_2$ using $\sum_s b_{s,i}^2 \leqslant B_i^2$ yields $\lambda_{\min}\big(\frac{1}{2}(H_i + H_i^\top)/c_0\big) \geqslant \chi(m,\eta)B_i$, and multiplying by $c_0 < 0$ with $B_i \leqslant mC$ gives $\widetilde{\lambda}$. For the norm, every entry of H_i is at most $2M_i\kappa(1-\eta)B_i^{\eta-2}$ in magnitude, and $B_i \geqslant m\varepsilon$ gives $\widetilde{L}$. The complete proof is in Appendix C.14. □

The SB property is therefore restored by imposing a sufficiently small participation floor, without changing the equilibrium characterization. Its only cost is a deterioration in the convergence rate as the resulting smoothness bound $\widetilde{L}$ grows like $\varepsilon^{\eta-2}$ as the floor shrinks, so in practice one takes ε close to, but below, $\bar{\varepsilon}$. We now bring together the above results to establish the main convergence and welfare result for the Power RCSF in the next section.

8 BRD convergence and welfare-optimality: Best of both worlds

We now combine the pieces. The Power RCSF has proportional private and social marginal returns with the constant $c = m\eta/(m+\eta-1)$ (Proposition 2), and within the SRCSF class it is the only CSF with this property (Lemma 5). The budget-balanced charge at rate $\tau = 1/c - 1$ makes the welfare optimum an equilibrium without requiring budgets to be exhausted (Proposition 4), which removes the obstacle of Proposition 3. The participation floor restores the SB property (Lemma 8), and below the threshold $\bar{\varepsilon}$ it is inactive at the equilibrium (Lemma 7), which is established to be unique and symmetric (Lemma 6). The theorem below combines these results to provide the three guarantees for the Power RCSF.

Theorem 4 (Welfare-optimal NE under Power RCSF)**.** *Let h be the Power RCSF (Definition 9) played on the floored set $\mathcal{B}^{\varepsilon}$ (Definition 10) with a budget-balanced charge (Equation (20)), and suppose*

(F) Participation floor. $0 < \varepsilon < \bar{\varepsilon}$, *with $\bar{\varepsilon}$ as in Equation (25); and*

(T) Charge. $\tau = \dfrac{1}{c} - 1 = \dfrac{(m-1)(1-\eta)}{m\eta}$, *with c as in Equation (18).*

Then the following hold:

(a) Convergence. *The best-response dynamics (Algorithm 1, run with $\mathcal{D}^{\ell} = \mathcal{B}^{\varepsilon}$) with step size $\gamma = \widetilde{\lambda}/\widetilde{L}^2$ converges to a Nash equilibrium $\mathbf{b}^{\star}$ of the floored game, reaching $\|\mathbf{g}(\mathbf{b}(K_\epsilon), \boldsymbol{\mu}(K_\epsilon))\|_2 < \epsilon$ after $K_\epsilon = O\big((\widetilde{L}/\widetilde{\lambda})^2 \log(1/\epsilon)\big)$ iterations, with $\widetilde{\lambda}, \widetilde{L}$ as in Lemma 8.*

(b) Uniqueness and symmetry. *The floored game has a unique Nash equilibrium $\mathbf{b}^{\star}$, which is symmetric and coincides with the unique Nash equilibrium of the unrestricted game. Moreover, $\widetilde{u}_s(\mathbf{b}^{\star}) > 0$ for every s, so $\mathbf{b}^{\star} \in \mathcal{B}_{\geqslant 0}$.*

(c) Welfare optimality. *For every $p \in (-\infty, 1]$, $\mathbf{b}^{\star}$ maximizes the p-mean welfare W_p over $\mathcal{B}_{\geqslant 0}$, and hence also over $\mathcal{B}_{\geqslant 0} \cap \mathcal{B}^{\varepsilon}$.*

Proof. **(a)** By Lemma 8, the induced utilities satisfy the SB property on $\mathcal{B}^{\varepsilon}$ with the constants $\widetilde{\lambda}$ and $\widetilde{L}$, so Theorem 2 applies with $\mathcal{D}^{\ell} = \mathcal{B}^{\varepsilon}$. $\mathcal{B}^{\varepsilon}$ is convex and compact and is non-empty, because $n\varepsilon < C$ by Lemma 7. Every iterate of Algorithm 1 lies in $\mathcal{B}^{\varepsilon}$ by construction of the correction $\boldsymbol{\mu}_s$ in Equation (8), and the iterates converge to a Nash equilibrium $\mathbf{b}^{\star}$ of the floored game at the stated rate $\gamma = \widetilde{\lambda}/\widetilde{L}^2$; the iteration count follows from the bound in Theorem 2, since $2(\widetilde{L}/\widetilde{\lambda})^2 \log(\|\mathbf{g}(\mathbf{b}(0), \boldsymbol{\mu}(0))\|_2/\epsilon) = O\big((\widetilde{L}/\widetilde{\lambda})^2 \log(1/\epsilon)\big)$

(b) Lemma 6 gives uniqueness and symmetry on the unrestricted set $\mathcal{B}$. The same argument applies verbatim on $\mathcal{B}^{\varepsilon}$ with the lower boundary 0 replaced by ε: in *Part I*, the marginal dominance inequality in Equation $(\star)$ does not involve the boundary value, the multiplier $\mu_{s,i}$ now enforces $b_{s,i} \geqslant \varepsilon$ with slackness $\mu_{s,i}(b_{s,i} - \varepsilon) = 0$, and $b^{\star}_{p,i_0} > b^{\star}_{q,i_0} \geqslant \varepsilon$ gives $b^{\star}_{p,i_0} > \varepsilon$, so $\mu^{\star}_{p,i_0} = 0$ and the case analysis is unchanged; and in *Part II*, $\beta_i(\Gamma)$ is replaced by $\max\{(M_i K/\Gamma)^{\theta}, \varepsilon\}$, so that the total

budget $S_\varepsilon(\Gamma) = \sum_i \max\{\beta_i(\Gamma), \varepsilon\}$ is continuous and non-increasing in Γ, and is strictly decreasing whenever at least one coordinate lies above the floor. Since $n\varepsilon < C$ by Lemma 7, if $S_\varepsilon(1+\tau) > C$, then $\lim_{\Gamma\to\infty} S_\varepsilon(\Gamma) = n\varepsilon < C$, hence the level C is crossed at a finite $\Gamma^\star > 1+\tau$ in the strictly decreasing region, which gives a unique $\Gamma^\star$; if $S_\varepsilon(1+\tau) \leqslant C$, then $\Gamma^\star = 1+\tau$. Thus in either case the floored game has a unique symmetric equilibrium.

By assumption **(F)** and Lemma 7, the unrestricted unique and symmetric solution $\mathbf{b}_s^\star$ already lies strictly above ε at every vertex, i.e. $b_{s,i}^\star = t_i^\star > \varepsilon$, and hence $\mathbf{b}^\star \in \mathcal{B}^\varepsilon$ and no floor constraint is active at it. Each $\mathbf{b}_s^\star$ is a best response to $\mathbf{b}_{-s}^\star$ over $\mathcal{B}_s$, and since it lies in the smaller set $\mathcal{B}_s^\varepsilon \subseteq \mathcal{B}_s$, it is also a best response over $\mathcal{B}_s^\varepsilon$. So $\mathbf{b}^\star$ is also a NE of the floored game. By the uniqueness established above, it is the unique NE of the floored game. Hence the floor does not affect the equilibrium.

For positivity of the taxed utility (22), note that at the symmetric profile the rebate (21) returns to each firm exactly what it paid, $\rho_s = \tau \sum_i t_i^\star$, so $\widetilde{u}_s(\mathbf{b}^\star) = \kappa m^{\eta-1} \sum_i M_i (t_i^\star)^\eta - \sum_i t_i^\star$. Stationarity gives $M_i K (t_i^\star)^{\eta-1} = \Gamma^\star$, i.e. $M_i (t_i^\star)^\eta = \Gamma^\star t_i^\star / K$, and substituting with $K = \kappa m^{\eta-2}(m+\eta-1)$ gives

$$\widetilde{u}_s(\mathbf{b}^\star) = \widetilde{u}(\mathbf{b}^\star) = \Big(\sum_i t_i^\star\Big)\Big[\frac{m\,\Gamma^\star}{m+\eta-1} - 1\Big], \quad \forall s \in \mathscr{M} \tag{26}$$

By Lemma 6, $\Gamma^\star \geqslant 1+\tau = 1/c = (m+\eta-1)/(m\eta)$, and since $\eta < 1$ we have $(m+\eta-1)/(m\eta) > (m+\eta-1)/m$, so the bracket in (26) is strictly positive. Hence $\widetilde{u}(\mathbf{b}^\star) > 0$ for every s and by Definition 2, we can conclude that $\mathbf{b}^\star$ lies in $\mathcal{B}_{\geqslant 0}$.

(c) *Step 1: the utilitarian optimum without the floor.* For the taxed utilities $\widetilde{u}_s$, the charge and rebate cancel in the sum of the utilties, so the utilitarian welfare (W_p for $p = 1$) is

$$W_1(\mathbf{b}) = \frac{1}{m}\sum_s \widetilde{u}_s(\mathbf{b}) = \frac{1}{m}\Big[\kappa \sum_i M_i B_i^\eta - \sum_i B_i\Big],$$

which depends on $\mathbf{b}$ only through the total budget $B_i = \sum_s b_{s,i}$. Over $\mathcal{B}$ the attainable totals are exactly the vectors $\{\mathbf{B} : B_i \geqslant 0,\ \sum_i B_i \leqslant mC\}$: any feasible $\mathbf{b}$ produces such totals and any such totals are realized by the symmetric split $b_{s,i} = B_i/m$, which is feasible for each firm. This objective is separable and strictly concave in $\mathbf{B}$, and we call its maximizer $\mathbf{B}^{\mathrm{opt}}$: let $\mathbf{b}^{\mathrm{opt}}$ be its symmetric split, with $t_i^{\mathrm{opt}} = B_i^{\mathrm{opt}}/m$. Then $W_1(\mathbf{b}) \leqslant W_1(\mathbf{b}^{\mathrm{opt}})$ for every $\mathbf{b} \in \mathcal{B}$, since W_1 doesn't depend on how the total budget B_i^{opt} is split among firms.

Step 2: $\mathbf{b}^{\mathrm{opt}}$ *is the equilibrium.* By Proposition 2 the Power RCSF has proportional marginals with constant c, and each u_s, hence each $\widetilde{u}_s$, is concave in $\mathbf{b}_s$. Every total B_i^{opt} is strictly positive, because in the planner's first-order condition $M_i\psi(t_i) - 1 = \hat{\lambda} - \hat{\mu}_i$, the marginal welfare $M_i\psi(t_i) = M_i\kappa\eta B_i^{\eta-1}$ grows without bound as $B_i \to 0$ and $M_i \geqslant 1-\alpha > 0$ by Lemma 10; so the planner's first-order conditions are well defined at $\mathbf{b}^{\mathrm{opt}}$. By assumption **(T)** the charge satisfies $1+\tau = 1/c$. Proposition 4 therefore applies and $\mathbf{b}^{\mathrm{opt}}$ is a Nash equilibrium of the game without the floor. By part (b) the Nash equilibrium is unique, so $\mathbf{b}^\star = \mathbf{b}^{\mathrm{opt}}$, maximizes W_1 over $\mathcal{B}$, and also over $\mathcal{B}^\varepsilon \subseteq \mathcal{B}$, to which it belongs.

Step 3: extension to all $p \in (-\infty, 1]$. For $\mathbf{b} \in \mathcal{B}_{\geqslant 0}$, $W_p(\mathbf{b})$ is the power mean of order p of the non-negative numbers $\widetilde{u}_1(\mathbf{b}), \ldots, \widetilde{u}_m(\mathbf{b})$. By the power mean inequality, a power mean of order $p \leqslant 1$ never exceeds the arithmetic mean, with the limiting conventions of Definition 2 for $p \leqslant 0$, so $W_p(\mathbf{b}) \leqslant W_1(\mathbf{b})$. At the equilibrium this inequality is an equality. Since $\mathbf{b}^\star$ is symmetric, every firm holds an identical allocation and receives a rebate equal to its own payment ($\rho_s = \tau \sum_i t_i^\star$), so all utilities coincide at the common value $\tilde{u}(\mathbf{b}^\star) > 0$ of Equation (26) in part (b); and any

power mean of m equal numbers equals that number, so $W_p(\mathbf{b}^\star) = W_1(\mathbf{b}^\star)$. Steps 1 and 2 give $W_1(\mathbf{b}) \leqslant W_1(\mathbf{b}^{\text{opt}}) = W_1(\mathbf{b}^\star)$ for every $\mathbf{b} \in \mathcal{B}$. Combining all three inequalities together, we obtain

$$W_p(\mathbf{b}) \;\leqslant\; W_1(\mathbf{b}) \;\leqslant\; W_1(\mathbf{b}^\star) \;=\; W_p(\mathbf{b}^\star) \qquad \forall\, \mathbf{b} \in \mathcal{B}_{\geqslant 0},\; p \in (-\infty, 1].$$

Since $\mathbf{b}^\star \in \mathcal{B}_{\geqslant 0} \cap \mathcal{B}^\varepsilon$ by part (b), it maximizes W_p over $\mathcal{B}_{\geqslant 0}$ and over $\mathcal{B}_{\geqslant 0} \cap \mathcal{B}^\varepsilon$. By Definition 2, the planner's problem is restricted to $\mathcal{B}_{\geqslant 0}$, so profiles outside this set need not be considered. □

Corollary 1. *Under the conditions of Theorem 4,* PoA $= 1$ *for every* $p \in (-\infty, 1]$ *simultaneously.*

Remark 5 (Unequal budgets). Suppose firm s has its own budget C_s, and replace mC by $\sum_s C_s$ in the cap on κ and in the constant $\tilde{\lambda}$ of Lemma 8, and C by $\min_s C_s$ in the requirement $n\varepsilon < C$ that keeps the floored set non-empty. Theorem 1 does not involve budgets, Theorem 2 only needs $\mathcal{D}^\ell$ to be convex and compact, and Lemma 8 holds with this replacement; hence best-response dynamics still converge to a Nash equilibrium of the floored game. What fails is symmetry: the case analysis in the proofs of Lemmas 4 and 6 uses equal budgets, so the equilibrium is in general asymmetric. Characterizing equilibrium and welfare in this setting is left as future work.

9 Conclusion

We developed a two-timescale framework to study how multiple competing firms allocate advertising budgets over social networks to maximize brand awareness, with convergence-rate guarantees on the fast timescale that formally justify the separation of dynamics. Under standard conditions, firms' best-response dynamics (BRD) converge to a pure-strategy Nash equilibrium (NE), which need not coincide with the social optimum. We therefore study whether a platform can design a contest success function (CSF), which determines how attention is allocated among firms, so that competition leads to a socially optimal outcome.

For standard ratio-form CSFs, BRD converges to a unique symmetric equilibrium, but equilibrium behaviour cannot in general be aligned with social welfare. We show that this failure is structural: firms' private marginal returns to spending are strictly larger than the corresponding social marginal returns. We then show that this gap can be accounted for through a broader class of scaled ratio CSFs. For NE and welfare to align, it suffices that private and social marginal returns are proportional, and within the class of scaled ratio CSFs, this condition singles out the Power RCSF. The Power RCSF alone is not sufficient: a budget-balanced charge removes the need for firms to exhaust their budgets, while a small participation floor restores the regularity needed for convergence of BRD. Together, these ingredients yield dynamics that converge to a unique, symmetric equilibrium maximizing every p-mean welfare with $p \leqslant 1$, so the Price of Anarchy is one. Our results show that the allocation rule itself can be a useful design instrument: rather than eliminating competition, a platform can shape how competition translates into socially desirable outcomes.

A natural direction for future work is to allow firms to have heterogeneous budgets. The BRD and its convergence guarantee continue to apply, but the equilibrium need not be symmetric in general. Consequently, our welfare characterization does not extend directly to heterogeneous budgets. Characterizing equilibrium and welfare in this setting is an interesting direction for future work.

References

[1] Milton Abramowitz and Irene A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, volume 55. US Government Printing Office, 1948.

[2] Noga Alon, Michal Feldman, Ariel D. Procaccia, and Moshe Tennenholtz. A note on competitive diffusion through social networks. *Information Processing Letters*, 110(6):221–225, 2010.

[3] Amirhossein Ansari, Masoud Dadgar, Ali Hamzeh, Jörg Schlötterer, and Michael Granitzer. Competitive influence maximization: Integrating budget allocation and seed selection. *arXiv preprint arXiv:1912.12283*, 2019.

[4] Prithu Banerjee, Wei Chen, and Laks V. S. Lakshmanan. Maximizing welfare in social networks under a utility driven influence diffusion model. In *Proceedings of the 2019 International Conference on Management of Data*, pages 1078–1095, 2019.

[5] Prithu Banerjee, Wei Chen, and Laks V. S. Lakshmanan. Maximizing social welfare in a competitive diffusion model. *arXiv preprint arXiv:2012.03354*, 2020.

[6] Siddharth Barman, Umang Bhaskar, Anand Krishna, and Ranjani G Sundaram. Tight approximation algorithms for p-mean welfare under subadditive valuations. *arXiv preprint arXiv:2005.07370*, 2020.

[7] William J. Baumol. On taxation and the control of externalities. *The American Economic Review*, 62(3):307–322, 1972.

[8] Ruben Becker, Gianlorenzo D'Angelo, Sajjad Ghobadi, and Hugo Gilbert. Fairness in influence maximization through randomization. *Journal of Artificial Intelligence Research*, 73:1251–1283, 2022.

[9] David E. Bell, Ralph L. Keeney, and John D. C. Little. A market share theorem. *Journal of Marketing Research*, 12(2):136–141, 1975.

[10] Dimitri P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, MA, 2nd edition, 1999. ISBN 978-1886529007.

[11] Shishir Bharathi, David Kempe, and Mahyar Salek. Competitive influence maximization in social networks. In *Internet and Network Economics (WINE)*, pages 306–311. Springer, 2007.

[12] Kostas Bimpikis, Asuman Ozdaglar, and Ercan Yildiz. Competitive targeted advertising over networks. *Operations Research*, 64(3):705–720, 2016.

[13] Shih-Fen Cheng, Daniel M. Reeves, Yevgeniy Vorobeychik, and Michael P. Wellman. Notes on equilibria in symmetric games. In *Proceedings of the 6th Workshop on Game Theoretic and Decision Theoretic Agents*, pages 23–28, 2004.

[14] Tai-Yeong Chung. Rent-seeking contest when the prize increases with aggregate efforts. *Public Choice*, 87(1–2):55–66, 1996.

[15] Luis C. Corchón and Matthias Dahm. Foundations for contest success functions. *Economic Theory*, 43(1):81–98, 2010.

[16] Avinash Dixit and Victor Norman. Advertising and welfare. *The Bell Journal of Economics*, 9 (1):1–17, 1978.

[17] Edith Elkind, Abheek Ghosh, and Paul W. Goldberg. Continuous-time best-response and related dynamics in Tullock contests with convex costs. In *Web and Internet Economics (WINE 2024)*, Lecture Notes in Computer Science. Springer, 2024. arXiv:2402.08541.

[18] Christian Ewerhart and Federico Quartieri. Unique equilibrium in contests with incomplete information. *Economic Theory*, 70(1):243–271, 2020. doi: 10.1007/s00199-019-01209-4.

[19] Francisco Facchinei and Jong-Shi Pang. *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer, New York, 2003.

[20] Arastoo Fazeli and Ali Jadbabaie. Game theoretic analysis of a strategic model of competitive contagion and product adoption in social networks. In *51st IEEE Conference on Decision and Control (CDC)*, pages 74–79. IEEE, 2012.

[21] Arastoo Fazeli and Ali Jadbabaie. Targeted marketing and seeding products with positive externality. In *50th Annual Allerton Conference on Communication, Control, and Computing*, pages 1111–1117. IEEE, 2012.

[22] Arastoo Fazeli, Amir Ajorlou, and Ali Jadbabaie. Competitive diffusion in social networks: Quality or seeding? *IEEE Transactions on Control of Network Systems*, 4(3):665–675, 2016.

[23] Awi Federgruen and Nan Yang. Competition under generalized attraction models: Applications to quality competition under yield uncertainty. *Management Science*, 55(12):2028–2043, 2009.

[24] Abheek Ghosh. Best-response dynamics in Tullock contests with convex costs. In *Web and Internet Economics (WINE 2023)*, Lecture Notes in Computer Science. Springer, 2023. arXiv:2310.03528.

[25] Sanjeev Goyal and Michael Kearns. Competitive contagion in networks. In *Proceedings of the 44th Annual ACM Symposium on Theory of Computing*, pages 759–774, 2012.

[26] Ramesh Johari and John N. Tsitsiklis. Efficiency loss in a network resource allocation game. *Mathematics of Operations Research*, 29(3):407–435, 2004.

[27] Aneel Karnani. Strategic implications of market share attraction models. *Management Science*, 31(5):536–547, 1985.

[28] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 137–146, 2003.

[29] Kai A. Konrad. *Strategy and Dynamics in Contests*. Oxford University Press, Oxford, 2009.

[30] Wei Lu, Francesco Bonchi, Amit Goyal, and Laks V. S. Lakshmanan. The bang for the buck: Fair competitive viral marketing from the host perspective. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 928–936, 2013.

[31] Takanori Maehara, Akihiro Yabe, and Ken-ichi Kawarabayashi. Budget allocation problem with multiple advertisers: A game theoretic view. In *Proceedings of the 32nd International Conference on Machine Learning*, pages 428–437, 2015.

[32] Antonia Maria Masucci and Alonso Silva. Strategic resource allocation for competitive influence in social networks. In *52nd Annual Allerton Conference on Communication, Control, and Computing*, pages 951–958. IEEE, 2014.

[33] Antonia Maria Masucci and Alonso Silva. Advertising competitions in social networks. In *2017 American Control Conference (ACC)*, pages 4619–4624. IEEE, 2017.

[34] Reut Megidish and Aner Sela. Allocation of prizes in contests with participation constraints. *Journal of Economics & Management Strategy*, 22(4):713–727, 2013. doi: 10.1111/jems.12030.

[35] George E. Monahan. The structure of equilibria in market share attraction models. *Management Science*, 33(2):228–243, 1987.

[36] Aniket Murhekar, Zhuowen Yuan, Bhaskar Ray Chaudhury, Bo Li, and Ruta Mehta. Incentives in federated learning: Equilibria, dynamics, and mechanisms for welfare maximization. *Advances in Neural Information Processing Systems*, 36, 2024.

[37] Arthur C. Pigou. *The Economics of Welfare*. Macmillan, London, 1920.

[38] Aida Rahmattalabi, Shahin Jabbari, Himabindu Lakkaraju, Phebe Vayanos, Max Izenberg, Ryan Brown, Eric Rice, and Milind Tambe. Fair influence maximization: A welfare optimization approach. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 11630–11638, 2021.

[39] J. B. Rosen. Existence and uniqueness of equilibrium points for concave N-person games. *Econometrica*, 33(3):520–534, 1965.

[40] Carl P Simon, Lawrence Blume, and Lawrence Blume. *Mathematics for economists*, volume 7. Norton New York, 1994.

[41] Stergios Skaperdas. Contest success functions. *Economic Theory*, 7:283–290, 1996.

[42] Ferenc Szidarovszky and Koji Okuguchi. On the existence and uniqueness of pure Nash equilibrium in rent-seeking games. *Games and Economic Behavior*, 18(1):135–140, 1997.

[43] Gordon Tullock. Efficient rent seeking. *Toward a Theory of the Rent-Seeking Society*, pages 97–112, 1980.

[44] Vasileios Tzoumas, Christos Amanatidis, and Evangelos Markakis. A game-theoretic analysis of a competitive diffusion process over social networks. In *International Workshop on Internet and Network Economics*, pages 1–14. Springer, 2012.

[45] Vineeth S. Varma, Irinel-Constantin Morărescu, Samson Lasaulce, and Samuel Martin. Marketing resource allocation in duopolies over social networks. *IEEE Control Systems Letters*, 2 (4):593–598, 2018.

[46] Vineeth S. Varma, Samson Lasaulce, Julien Mounthanyvong, and Irinel-Constantin Morărescu. Allocating marketing resources over social networks: A long-term analysis. *IEEE Control Systems Letters*, 3(4):1002–1007, 2019.

[47] Pengzi Wang, Yiming Zhu, Kai Han, Zhizhuo Yin, Qing Xiu, and Pan Hui. Fairness-aware algorithms for seed allocation in social advertising. In *IEEE 24th International Conference on High Performance Computing and Communications*, pages 464–473. IEEE, 2022.

[48] Jinhang Zuo, Xutong Liu, Carlee Joe-Wong, John CS Lui, and Wei Chen. Online competitive influence maximization. In *International conference on artificial intelligence and statistics*, pages 11472–11502. PMLR, 2022.

A Appendix to Section 3.1

A.1 Proof of Theorem 1: Convergence of the Fast-Timescale Dynamics

Throughout the proof we fix the slow-timescale period k and the firm s, and we suppress both from the notation where no confusion arises. In particular, we write $\mathbf{h} := \mathbf{h}_s(\mathbf{b}(k))$ for the vector of CSF values, which stays constant during the period, $\mathbf{x}(t)$ for $\mathbf{x}_s(t)$, and $\mathbf{x}^*$ for the steady state (2). In vector form, the dynamics (1) read

$$\mathbf{x}(t) = \frac{1}{t}\Big[\alpha\mathscr{E}\,\mathbf{x}(t-1) + (1-\alpha)\,\mathbf{h}\Big] + \Big(1-\frac{1}{t}\Big)\mathbf{x}(t-1), \qquad t \geqslant 1. \tag{27}$$

We work with the ℓ_∞ norm $\|\mathbf{v}\|_\infty = \max_i |v_i|$ on $\mathbb{R}^n$. Its induced operator norm on matrices is the maximum absolute row sum, $\|A\|_\infty = \max_i \sum_j |a_{ij}|$, and it is sub-multiplicative: $\|A\mathbf{v}\|_\infty \leqslant \|A\|_\infty \|\mathbf{v}\|_\infty$ for every matrix A and vector $\mathbf{v}$.

Step 1: the steady state lies in $[0,1]^n$. We first record a property of $\mathbf{x}^*$ that is used in part (b). Since $\mathscr{E}$ is row-stochastic, the spectral radius of $\alpha\mathscr{E}$ is at most $\alpha\|\mathscr{E}\|_\infty = \alpha < 1$, so the Neumann series $\sum_{\ell\geqslant 0}(\alpha\mathscr{E})^\ell$ converges to $(I-\alpha\mathscr{E})^{-1}$. Every term of this series has non-negative entries, so the matrix $M := (1-\alpha)(I-\alpha\mathscr{E})^{-1}$ is non-negative. Its row sums equal one. Indeed, $\mathscr{E}\mathbf{1} = \mathbf{1}$ gives $(\alpha\mathscr{E})^\ell\mathbf{1} = \alpha^\ell\mathbf{1}$ for every ℓ, and summing the geometric series yields $M\mathbf{1} = (1-\alpha)\sum_{\ell\geqslant 0}\alpha^\ell\,\mathbf{1} = \mathbf{1}$. Hence every coordinate of $\mathbf{x}^* = M\mathbf{h}$ is a weighted average of the coordinates of $\mathbf{h}$, and since $\mathbf{h} \in [0,1]^n$, we conclude that $\mathbf{x}^* \in [0,1]^n$.

Step 2: the error obeys a linear recursion. Define the error $\mathbf{e}(t) := \mathbf{x}(t) - \mathbf{x}^*$. By construction, $\mathbf{x}^*$ satisfies the fixed-point equation

$$\mathbf{x}^* = \alpha\mathscr{E}\,\mathbf{x}^* + (1-\alpha)\,\mathbf{h}. \tag{28}$$

We subtract $\mathbf{x}^*$ from both sides of (27). On the right-hand side, we write $\mathbf{x}^* = \frac{1}{t}\mathbf{x}^* + (1-\frac{1}{t})\mathbf{x}^*$ and replace the first copy by the right-hand side of (28). Grouping the terms that multiply $1/t$ and $1-1/t$, the constant term $(1-\alpha)\mathbf{h}$ cancels, and we obtain

$$\mathbf{e}(t) = \frac{1}{t}\Big[\alpha\mathscr{E}\big(\mathbf{x}(t-1)-\mathbf{x}^*\big)\Big] + \Big(1-\frac{1}{t}\Big)\big(\mathbf{x}(t-1)-\mathbf{x}^*\big).$$

In other words, the error evolves by the linear recursion

$$\mathbf{e}(t) = A(t)\,\mathbf{e}(t-1), \qquad \text{where} \quad A(t) := \Big(1-\frac{1}{t}\Big)I + \frac{\alpha}{t}\,\mathscr{E}. \tag{29}$$

Step 3: the norm of each step. For $t \geqslant 1$ the coefficients $1-1/t$ and α/t are non-negative, and $\mathscr{E}$ has non-negative entries, so every entry of $A(t)$ is non-negative. Its row sums can therefore be computed without absolute values. Using the row-stochasticity of $\mathscr{E}$, that is, $\sum_j e_{ij} = 1$ for every i, the ith row sum is

$$\sum_j\Big[\Big(1-\frac{1}{t}\Big)\delta_{ij} + \frac{\alpha}{t}\,e_{ij}\Big] = 1 - \frac{1}{t} + \frac{\alpha}{t} = 1 - \frac{1-\alpha}{t},$$

where δ_{ij} is the Kronecker delta. All rows have the same sum, so

$$\|A(t)\|_\infty = 1 - \frac{1-\alpha}{t} \qquad \text{and, in fact,} \qquad A(t)\,\mathbf{1} = \Big(1-\frac{1-\alpha}{t}\Big)\mathbf{1}. \tag{30}$$

Since $\alpha \in (0,1)$, this number lies in $(0,1)$ for every $t \geqslant 1$, so each step strictly contracts the error.

Step 4: proof of part (a). Applying the recursion (29) successively for $t = T, T-1, \dots, 1$ gives $\mathbf{e}(T) = A(T)A(T-1)\cdots A(1)\,\mathbf{e}(0)$. By sub-multiplicativity of the operator norm and (30),

$$\|\mathbf{e}(T)\|_\infty \leqslant \prod_{t=1}^{T} \|A(t)\|_\infty \, \|\mathbf{e}(0)\|_\infty = P_T \, \|\mathbf{e}(0)\|_\infty, \qquad P_T := \prod_{t=1}^{T} \Big(1 - \frac{1-\alpha}{t}\Big). \tag{31}$$

It remains to bound P_T. Every factor of P_T is positive, so we may take logarithms. By the elementary inequality $\log(1-x) \leqslant -x$, valid for every $x < 1$, applied with $x = (1-\alpha)/t$, we get

$$\log P_T = \sum_{t=1}^{T} \log\Big(1 - \frac{1-\alpha}{t}\Big) \leqslant -(1-\alpha) \sum_{t=1}^{T} \frac{1}{t}.$$

We now bound the harmonic sum from below by an integral. On each interval $[t, t+1]$ we have $1/t \geqslant 1/x$, so $\sum_{t=1}^{T} \frac{1}{t} \geqslant \int_1^{T+1} \frac{dx}{x} = \log(T+1)$. Substituting this, and using $1-\alpha > 0$, we obtain $\log P_T \leqslant -(1-\alpha)\log(T+1)$, that is,

$$P_T \leqslant (T+1)^{-(1-\alpha)}. \tag{32}$$

Together with (31), this proves the inequality in part (a). Since $(T+1)^{-(1-\alpha)} \to 0$ as $T \to \infty$, it also shows that $\mathbf{x}(T) \to \mathbf{x}^*$.

Step 5: proof of part (b). By hypothesis $\mathbf{x}(0) \in [0,1]^n$, and by Step 1 also $\mathbf{x}^* \in [0,1]^n$. The difference of two vectors in $[0,1]^n$ has every coordinate in $[-1,1]$, so $\|\mathbf{e}(0)\|_\infty \leqslant 1$. Combining this with part (a), we have $\|\mathbf{e}(T)\|_\infty \leqslant (T+1)^{-(1-\alpha)}$. The right-hand side is at most ϵ_0 exactly when $(T+1)^{1-\alpha} \geqslant 1/\epsilon_0$, that is, when $T+1 \geqslant \epsilon_0^{-1/(1-\alpha)}$. This holds in particular whenever $T \geqslant \epsilon_0^{-1/(1-\alpha)}$, which proves part (b).

Step 6: proof of part (c). We first show that the bound (31) can hold with equality. Suppose $\mathbf{e}(0) = c\,\mathbf{1}$ for some constant c. By the second identity in (30), the vector $\mathbf{1}$ is an eigenvector of every $A(t)$, with eigenvalue $1 - (1-\alpha)/t$. Applying the recursion (29) T times therefore gives $\mathbf{e}(T) = P_T\, c\,\mathbf{1}$ exactly, so $\|\mathbf{e}(T)\|_\infty = P_T \|\mathbf{e}(0)\|_\infty$. Hence no factor smaller than P_T can serve in part (a) for all initial conditions.

We next evaluate P_T exactly, to identify its rate of decay. Writing each factor as $(t-(1-\alpha))/t$ and setting $c_0 := 1-\alpha \in (0,1)$,

$$P_T = \frac{\prod_{t=1}^{T}(t-c_0)}{\prod_{t=1}^{T} t} = \frac{\prod_{t=1}^{T}(t-c_0)}{T!}.$$

The numerator is evaluated by the following identity.

Lemma 9 (Gamma product identity)**.** *For every $c_0 \in (0,1)$ and every positive integer T,* $\displaystyle\prod_{t=1}^{T}(t - c_0) = \frac{\Gamma(T+1-c_0)}{\Gamma(1-c_0)}$.

Proof. We use the functional equation of the Gamma function, $\Gamma(z+1) = z\,\Gamma(z)$ for every $z > 0$. Applying it first with $z = T - c_0$, then with $z = T-1-c_0$, and so on down to $z = 1-c_0$, we obtain

$$\Gamma(T+1-c_0) = (T-c_0)\,\Gamma(T-c_0) = (T-c_0)(T-1-c_0)\,\Gamma(T-1-c_0) = \cdots = \Gamma(1-c_0)\prod_{t=1}^{T}(t-c_0).$$

Every argument used is positive because $c_0 < 1$. Dividing both sides by $\Gamma(1-c_0) > 0$ gives the identity. □

Applying Lemma 9, and using $T! = \Gamma(T+1)$ and $1 - c_0 = \alpha$, we obtain

$$P_T = \frac{\Gamma(T+\alpha)}{\Gamma(\alpha)\,\Gamma(T+1)}.$$

Finally, the standard asymptotic ratio of Gamma functions, $\Gamma(T+a)/\Gamma(T+b) \sim T^{a-b}$ as $T \to \infty$ [1, Eq. 6.1.46], applied with $a = \alpha$ and $b = 1$, gives

$$P_T \sim \frac{T^{-(1-\alpha)}}{\Gamma(\alpha)} \qquad \text{as } T \to \infty.$$

So the error factor decays exactly at the polynomial rate $T^{-(1-\alpha)}$, and the upper bound (32) used in parts (a) and (b) is sharp up to the constant factor $1/\Gamma(\alpha)$, which lies in $(0,1)$ for $\alpha \in (0,1)$. □

B Appendix to Section 4

B.1 Proof of Lemma 1.

For $\mathbf{b} \in \mathcal{D}^\ell$ and a correction $\boldsymbol{\mu} \in \mathbb{R}^n$, from Algorithm 1 line 3, we denote by $\mathbf{b}_s^{\boldsymbol{\mu}} := \mathbf{b}_s + \gamma\big(\nabla_{\mathbf{b}_s} u_s(\mathbf{b}) + \boldsymbol{\mu}\big)$, the updated budget that $\boldsymbol{\mu}$ produces for firm $s \in \mathscr{M}$. In particular, $\mathbf{b}_s^{\mathbf{0}} = \mathbf{b}_s + \gamma \nabla_{\mathbf{b}_s} u_s(\mathbf{b})$ is the *uncorrected budget update*, and $\mathbf{b}_s^{\boldsymbol{\mu}_s(k)} = \mathbf{b}_s(k+1)$ when $\mathbf{b} = \mathbf{b}(k)$. By definition, the minimisation in (8) is the minimisation of $\|\boldsymbol{\mu}\|_2$ over the set of admissible corrections

$$\mathcal{K}_s := \big\{\boldsymbol{\mu} \in \mathbb{R}^n : \ \mathbf{b}_s^{\boldsymbol{\mu}} \in \mathcal{D}_s^\ell\big\},$$

that is, over the corrections whose updated budget is feasible.

Step 1: the correction in (8) is a projection. Since $\mathbf{b}_s^{\boldsymbol{\mu}} = \mathbf{b}_s^{\mathbf{0}} + \gamma\boldsymbol{\mu}$, consider the map $T : \mathcal{K}_s \to \mathcal{D}_s^\ell, \quad T(\boldsymbol{\mu}) := \mathbf{b}_s^{\mathbf{0}} + \gamma\boldsymbol{\mu} = \mathbf{b}_s^{\boldsymbol{\mu}}$, which maps each admissible correction to the budget it produces which lies in $\mathcal{D}_s^\ell$ by the definition of $\mathcal{K}_s$. The map T is a bijection, with inverse $T^{-1}(\mathbf{b}'_s) = \boldsymbol{\mu} = \frac{\mathbf{b}'_s - \mathbf{b}_s^{\mathbf{0}}}{\gamma} = \frac{\mathbf{b}'_s - \mathbf{b}_s}{\gamma} - \nabla_{\mathbf{b}_s} u_s(\mathbf{b})$, for all $\mathbf{b}'_s \in \mathcal{D}_s^\ell$. Thus the minimum-norm correction problem in (8), $\min_{\boldsymbol{\mu}\in\mathcal{K}_s} \|\boldsymbol{\mu}\|_2$ is equivalent to $\min_{\mathbf{b}'_s \in \mathcal{D}_s^\ell} \big\|\mathbf{b}'_s - \mathbf{b}_s^{\mathbf{0}}\big\|_2$, i.e. the Euclidean projection of $\mathbf{b}_s^{\mathbf{0}}$ onto the convex set $\mathcal{D}_s^\ell$: $\Pi_{\mathcal{D}_s^\ell}\big(\mathbf{b}_s^{\mathbf{0}}\big) := \arg\min_{\mathbf{b}'_s \in \mathcal{D}_s^\ell} \big\|\mathbf{b}'_s - \mathbf{b}_s^{\mathbf{0}}\big\|_2$.

Step 2: existence and uniqueness. The set $\mathcal{D}_s^\ell$ is non-empty, closed and convex. By the projection theorem [10, Prop. 2.1.3], the projection of any point onto a non-empty, closed and convex set exists and is unique. Hence $\Pi_{\mathcal{D}_s^\ell}(\mathbf{b}_s^{\mathbf{0}})$ exists and is unique. Since T is a bijection, exactly one admissible correction produces this budget, namely $\boldsymbol{\mu}_s(\mathbf{b}) = T^{-1}\big(\Pi_{\mathcal{D}_s^\ell}(\mathbf{b}_s^{\mathbf{0}})\big)$ which by Step 1, is the unique solution of (8).

Step 3: a characterisation of the minimiser. The projection theorem also characterises the projection:

$$\mathbf{b}_s^{\boldsymbol{\mu}} = \Pi_{\mathcal{D}_s^\ell}\big(\mathbf{b}_s^{\mathbf{0}}\big) \quad \Longleftrightarrow \quad \big(\mathbf{b}_s^{\mathbf{0}} - \mathbf{b}_s^{\boldsymbol{\mu}}\big)^\top \big(\mathbf{b}'_s - \mathbf{b}_s^{\boldsymbol{\mu}}\big) \leqslant 0 \qquad \forall\, \mathbf{b}'_s \in \mathcal{D}_s^\ell.$$

From the update $\mathbf{b}_s^{\boldsymbol{\mu}} = \mathbf{b}_s^{\mathbf{0}} + \gamma\boldsymbol{\mu}$ we have $\mathbf{b}_s^{\mathbf{0}} - \mathbf{b}_s^{\boldsymbol{\mu}} = -\gamma\boldsymbol{\mu}$. Substituting this into the inequality and dividing by $-\gamma < 0$, we conclude, using Step 2, that an admissible correction $\boldsymbol{\mu}$ is the minimiser $\boldsymbol{\mu}_s(\mathbf{b})$ if and only if

$$\boldsymbol{\mu}^\top \big(\mathbf{b}'_s - \mathbf{b}_s^{\boldsymbol{\mu}}\big) \geqslant 0 \qquad \forall\, \mathbf{b}'_s \in \mathcal{D}_s^\ell. \tag{33}$$

In words, the minimal correction, which points from the uncorrected step to its nearest feasible budget, makes a non-obtuse angle with every direction pointing from that budget into the feasible set; otherwise a small move towards some feasible budget would bring the updated budget closer to the uncorrected step, and the correction would not be minimal. □

B.2 Proof of Lemma 2.

($\Leftarrow$) *Every Nash equilibrium has zero corrected gradient.* Let $\mathbf{b}^\star$ be a Nash equilibrium of the game on $\mathcal{D}^\ell$, fix a firm s and any $\mathbf{b}'_s \in \mathcal{D}^\ell_s$. For $t \in (0,1]$, the budget $\mathbf{b}^\star_s + t(\mathbf{b}'_s - \mathbf{b}^\star_s)$ lies in $\mathcal{D}^\ell_s$ by convexity. Since $\mathbf{b}^\star_s$ is a best response to $\mathbf{b}^\star_{-s}$, this budget cannot increase firm s's utility:

$$u_s\big(\mathbf{b}^\star_s + t(\mathbf{b}'_s - \mathbf{b}^\star_s), \mathbf{b}^\star_{-s}\big) \leqslant u_s(\mathbf{b}^\star)$$

Dividing by t,

$$\frac{u_s\big(\mathbf{b}^\star_s + t(\mathbf{b}'_s - \mathbf{b}^\star_s), \mathbf{b}^\star_{-s}\big) - u_s(\mathbf{b}^\star)}{t} \leqslant 0 \qquad \text{for every } t \in (0,1].$$

As $t \downarrow 0$, the left-hand side converges to the directional derivative of $u_s(\cdot, \mathbf{b}^\star_{-s})$ at $\mathbf{b}^\star_s$ in the direction $\mathbf{b}'_s - \mathbf{b}^\star_s$, which equals $\nabla_{\mathbf{b}_s} u_s(\mathbf{b}^\star)^\top(\mathbf{b}'_s - \mathbf{b}^\star_s)$ because u_s is continuously differentiable. This gives the first-order condition

$$\nabla_{\mathbf{b}_s} u_s(\mathbf{b}^\star)^\top\big(\mathbf{b}'_s - \mathbf{b}^\star_s\big) \leqslant 0 \qquad \text{for all } \mathbf{b}'_s \in \mathcal{D}^\ell_s. \tag{34}$$

Now consider the candidate correction $\hat{\boldsymbol{\mu}} := -\nabla_{\mathbf{b}_s} u_s(\mathbf{b}^\star)$. Its updated budget is $\mathbf{b}^{\hat{\boldsymbol{\mu}}}_s = \mathbf{b}^\star_s + \gamma\big(\nabla_{\mathbf{b}_s} u_s(\mathbf{b}^\star) - \nabla_{\mathbf{b}_s} u_s(\mathbf{b}^\star)\big) = \mathbf{b}^\star_s \in \mathcal{D}^\ell_s$, so $\hat{\boldsymbol{\mu}}$ is admissible. Moreover, for every $\mathbf{b}'_s \in \mathcal{D}^\ell_s$,

$$\hat{\boldsymbol{\mu}}^\top\big(\mathbf{b}'_s - \mathbf{b}^{\hat{\boldsymbol{\mu}}}_s\big) = -\nabla_{\mathbf{b}_s} u_s(\mathbf{b}^\star)^\top\big(\mathbf{b}'_s - \mathbf{b}^\star_s\big) \geqslant 0$$

by (34); that is, $\hat{\boldsymbol{\mu}}$ satisfies Appendix B.1's condition (33) (Proof of Lemma 1). By Lemma 1, the minimizer is unique and characterised by (33), so $\boldsymbol{\mu}_s(\mathbf{b}^\star) = \hat{\boldsymbol{\mu}}$. Hence

$$\mathbf{g}_s\big(\mathbf{b}^\star, \boldsymbol{\mu}_s(\mathbf{b}^\star)\big) = \nabla_{\mathbf{b}_s} u_s(\mathbf{b}^\star) + \hat{\boldsymbol{\mu}} = \mathbf{0}.$$

($\Rightarrow$) *Zero corrected gradient implies a Nash equilibrium.* Suppose $\mathbf{g}_s(\mathbf{b}^\star, \boldsymbol{\mu}_s(\mathbf{b}^\star)) = \mathbf{0}$ for every $s \in \mathscr{M}$, and fix s. By the definition (7), $\nabla_{\mathbf{b}_s} u_s(\mathbf{b}^\star) + \boldsymbol{\mu}_s(\mathbf{b}^\star) = \mathbf{0}$, i.e. $\boldsymbol{\mu}_s(\mathbf{b}^\star) = -\nabla_{\mathbf{b}_s} u_s(\mathbf{b}^\star)$, and the corresponding updated budget is $\mathbf{b}^{\boldsymbol{\mu}_s(\mathbf{b}^\star)}_s = \mathbf{b}^\star_s + \gamma \cdot \mathbf{0} = \mathbf{b}^\star_s$. Since $\boldsymbol{\mu}_s(\mathbf{b}^\star)$ is the minimiser, it satisfies (33) from the proof of Lemma 1 (Appendix B.1):

$$-\nabla_{\mathbf{b}_s} u_s(\mathbf{b}^\star)^\top\big(\mathbf{b}'_s - \mathbf{b}^\star_s\big) \geqslant 0, \quad \text{i.e.} \quad \nabla_{\mathbf{b}_s} u_s(\mathbf{b}^\star)^\top\big(\mathbf{b}'_s - \mathbf{b}^\star_s\big) \leqslant 0 \qquad \text{for all } \mathbf{b}'_s \in \mathcal{D}^\ell_s.$$

Since u_s is concave and differentiable in $\mathbf{b}_s$, it lies below its tangent plane at $\mathbf{b}^\star_s$: for every $\mathbf{b}'_s \in \mathcal{D}^\ell_s$,

$$u_s(\mathbf{b}'_s, \mathbf{b}^\star_{-s}) \leqslant u_s(\mathbf{b}^\star) + \nabla_{\mathbf{b}_s} u_s(\mathbf{b}^\star)^\top\big(\mathbf{b}'_s - \mathbf{b}^\star_s\big) \leqslant u_s(\mathbf{b}^\star).$$

Hence $\mathbf{b}^\star_s$ is a best response to $\mathbf{b}^\star_{-s}$ over $\mathcal{D}^\ell_s$, i.e. $\mathbf{b}^\star$ is a Nash equilibrium. □

B.3 Proof of Theorem 2

We stack the own-gradients as $\mathbf{g}(\mathbf{b}) := \big[\nabla_{\mathbf{b}_s} u_s(\mathbf{b})\big]_{s \in \mathscr{M}} \in \mathbb{R}^{mn}$, so that $\mathbf{g}(\mathbf{b}, \boldsymbol{\mu}) = \mathbf{g}(\mathbf{b}) + \boldsymbol{\mu}$. By definition of the block Hessian, $[G_{s_1,s_2}(\mathbf{b})]_{i,j} = \partial^2 u_{s_1} / (\partial b_{s_2,j} \partial b_{s_1,i})$, which is the derivative of the (s_1, i) entry of $\mathbf{g}(\mathbf{b})$ with respect to $b_{s_2,j}$; that is, $G(\mathbf{b})$ is the Jacobian matrix of $\mathbf{g}(\mathbf{b})$.

Step 1: Each utility is concave in its own budget. Strong concavity in Definition 4 gives $\mathbf{v}^\top G(\mathbf{b})\,\mathbf{v} \leqslant -\lambda\|\mathbf{v}\|_2^2$ for every $\mathbf{v} \in \mathbb{R}^{mn}$. Consider a $\mathbf{v}$ that is zero outside the block of firm s, with block $\mathbf{v}_s$. Then $\mathbf{v}^\top G(\mathbf{b})\mathbf{v} = \mathbf{v}_s^\top G_{s,s}(\mathbf{b})\mathbf{v}_s$, and $G_{s,s}(\mathbf{b})$ is the Hessian of u_s with respect to $\mathbf{b}_s$. Thus, $\mathbf{v}_s^\top G_{s,s}(\mathbf{b})\mathbf{v}_s \leqslant -\lambda\|\mathbf{v}_s\|_2^2 < 0$ for all $\mathbf{v}_s \neq \mathbf{0}$ and all $\mathbf{b} \in \mathcal{D}^\ell$. Along any segment in the convex

set $\mathcal{D}_s^\ell$, the second derivative of $u_s(\cdot, \mathbf{b}_{-s})$ is therefore negative, so u_s is concave in $\mathbf{b}_s$ on $\mathcal{D}_s^\ell$. Hence Lemma 2 applies.

Step 2: All iterates are feasible. By assumption, $\mathbf{b}(0) \in \mathcal{D}^\ell$. Suppose $\mathbf{b}(k) \in \mathcal{D}^\ell$. By Lemma 1, each correction $\boldsymbol{\mu}_s(k)$ exists, and it satisfies the constraint in (8), which says precisely that $\mathbf{b}_s(k+1) = \mathbf{b}_s(k) + \gamma\, \mathbf{g}_s(k) \in \mathcal{D}_s^\ell$. Hence $\mathbf{b}(k+1) \in \mathcal{D}^\ell$, and by induction $\mathbf{b}(k) \in \mathcal{D}^\ell$ for all $k \geqslant 0$. Since $\mathcal{D}^\ell$ is convex, the whole segment between two consecutive iterates also lies in $\mathcal{D}^\ell$.

Step 3: Geometric decay of the corrected gradient. The convergence of the budget update rule in line 3 of Algorithm 1 is analyzed through the L^2 norm of the update direction. Keep the correction $\boldsymbol{\mu}(k)$ fixed and move from $\mathbf{b}(k)$ to $\mathbf{b}(k+1)$. Since $\boldsymbol{\mu}(k)$ does not depend on the point at which $\mathbf{g}(\mathbf{b})$ is evaluated,

$$\mathbf{g}(\mathbf{b}(k+1), \boldsymbol{\mu}(k)) - \mathbf{g}(\mathbf{b}(k), \boldsymbol{\mu}(k)) = \mathbf{g}(\mathbf{b}(k+1)) - \mathbf{g}(\mathbf{b}(k)). \tag{35}$$

Since $\mathbf{g}(\mathbf{b})$ is a vector-valued map, we use its integral form. Let $\phi(\sigma) := \mathbf{g}(\mathbf{b}(k) + \sigma(\mathbf{b}(k+1) - \mathbf{b}(k)))$ for $\sigma \in [0,1]$. The points $\mathbf{b}(k)+\sigma(\mathbf{b}(k+1)-\mathbf{b}(k))$ lie in $\mathcal{D}^\ell$ (Step 2), and $\mathbf{g}(\mathbf{b})$ is continuously differentiable there with Jacobian G, so $\phi'(\sigma) = G(\mathbf{b}(k) + \sigma(\mathbf{b}(k+1) - \mathbf{b}(k)))\,(\mathbf{b}(k+1) - \mathbf{b}(k))$. The fundamental theorem of calculus gives

$$\mathbf{g}(\mathbf{b}(k+1)) - \mathbf{g}(\mathbf{b}(k)) = \int_0^1 \phi'(\sigma)\, d\sigma = \bar{G}(k)\,(\mathbf{b}(k+1) - \mathbf{b}(k)),$$

where

$$\bar{G}(k) := \int_0^1 G\big(\mathbf{b}(k) + \sigma(\mathbf{b}(k+1) - \mathbf{b}(k))\big)\, d\sigma \tag{36}$$

is the average Hessian-like matrix along the path from $\mathbf{b}(k)$ to $\mathbf{b}(k+1)$. Since every point of the segment lies in $\mathcal{D}^\ell$, the averaged matrix $\bar{G}(k)$ inherits both SB bounds of Definition 4. For every $\mathbf{v} \in \mathbb{R}^{mn}$,

$$\mathbf{v}^\top \bar{G}(k)\mathbf{v} = \int_0^1 \mathbf{v}^\top G(\cdot)\, \mathbf{v}\, d\sigma \leqslant -\lambda\|\mathbf{v}\|_2^2, \qquad \|\bar{G}(k)\mathbf{v}\|_2 \leqslant \int_0^1 \|G(\cdot)\, \mathbf{v}\|_2\, d\sigma \leqslant L\|\mathbf{v}\|_2. \tag{37}$$

Thus, (35) becomes

$$\mathbf{g}(\mathbf{b}(k+1), \boldsymbol{\mu}(k)) = \mathbf{g}(\mathbf{b}(k), \boldsymbol{\mu}(k)) + \bar{G}(k)(\mathbf{b}(k+1) - \mathbf{b}(k)) \tag{38}$$

Substituting $\mathbf{b}(k+1) - \mathbf{b}(k)$ from line 3 of Algorithm 1, we get:

$$\mathbf{g}(\mathbf{b}(k+1), \boldsymbol{\mu}(k)) = \mathbf{g}(\mathbf{b}(k), \boldsymbol{\mu}(k)) + \gamma \bar{G}(k)\mathbf{g}(\mathbf{b}(k), \boldsymbol{\mu}(k))$$

Taking the squared norm of both sides and expanding, we have

$$\begin{aligned} &\|\mathbf{g}(\mathbf{b}(k+1), \boldsymbol{\mu}(k))\|_2^2 \;=\; \|\mathbf{g}(\mathbf{b}(k), \boldsymbol{\mu}(k)) + \gamma \bar{G}(k)\mathbf{g}(\mathbf{b}(k), \boldsymbol{\mu}(k))\|_2^2 \\ &= \|\mathbf{g}(\mathbf{b}(k), \boldsymbol{\mu}(k))\|_2^2 \;+\; 2\gamma \mathbf{g}(\mathbf{b}(k), \boldsymbol{\mu}(k))^T \bar{G}(k)\mathbf{g}(\mathbf{b}(k), \boldsymbol{\mu}(k)) + \gamma^2 \|\bar{G}(k)\mathbf{g}(\mathbf{b}(k), \boldsymbol{\mu}(k))\|_2^2 \end{aligned} \tag{39}$$

From the assumption of strong concavity of utility applied with $\mathbf{v} = \mathbf{g}(\mathbf{b}(k), \boldsymbol{\mu}(k))$,

$$\mathbf{g}(\mathbf{b}(k), \boldsymbol{\mu}(k))^T \bar{G}(k)\mathbf{g}(\mathbf{b}(k), \boldsymbol{\mu}(k)) \leqslant -\lambda \|\mathbf{g}(\mathbf{b}(k), \boldsymbol{\mu}(k))\|_2^2 \tag{40}$$

From the assumption that the operator norm of the averaged Hessian-like matrix is bounded, $\|\bar{G}(k)\|_{\text{op}} \leqslant L$ and using the standard property of operator norms $\|A\mathbf{v}\|_2 \leqslant \|A\|_{\text{op}}\|\mathbf{v}\|_2$, we obtain the following inequalities

$$\|\bar{G}(k)\mathbf{g}(\mathbf{b}(k), \boldsymbol{\mu}(k))\|_2 \leqslant L\|\mathbf{g}(\mathbf{b}(k), \boldsymbol{\mu}(k))\|_2$$

$$\|\bar{G}(k)\mathbf{g}(\mathbf{b}(k),\boldsymbol{\mu}(k))\|_2^2 \leqslant L^2\|\mathbf{g}(\mathbf{b}(k),\boldsymbol{\mu}(k))\|_2^2 \tag{41}$$

Substituting the bounds from (40) and (41) in (39):

$$\|\mathbf{g}(\mathbf{b}(k+1),\boldsymbol{\mu}(k))\|_2^2 \leqslant \|\mathbf{g}(\mathbf{b}(k),\boldsymbol{\mu}(k))\|_2^2(1-2\gamma\lambda+\gamma^2L^2) \tag{42}$$

For convergence, the coefficient of the right-hand side must be less than one, leading to the step-size condition

$$1-2\gamma\lambda+\gamma^2L^2<1 \quad \Rightarrow \gamma < \frac{2\lambda}{L^2}$$

Hence, choose the step size as $\gamma = \frac{\lambda}{L^2}$.

Updating the correction cannot increase the norm: From the proof of Lemma 1 (Appendix B.1), $\boldsymbol{\mu}_s(k)$ satisfies (33) with updated budget $\mathbf{b}_s(k+1)$, and $\boldsymbol{\mu}_s(k+1)$ satisfies it with updated budget $\mathbf{b}_s(k+2)$. Take $\mathbf{b}'_s = \mathbf{b}_s(k+2)$ in the first and $\mathbf{b}'_s = \mathbf{b}_s(k+1)$ in the second; both are feasible by Step 2. Since $\mathbf{b}_s(k+2) - \mathbf{b}_s(k+1) = \gamma\,\mathbf{g}_s(\mathbf{b}(k+1),\boldsymbol{\mu}_s(k+1))$ by line 3 (Algorithm 1), summing over s and dividing by $\gamma>0$ gives

$$\boldsymbol{\mu}(k+1)^\top\mathbf{g}(\mathbf{b}(k+1),\boldsymbol{\mu}(k+1)) \;\leqslant\; 0 \;\leqslant\; \boldsymbol{\mu}(k)^\top\mathbf{g}(\mathbf{b}(k+1),\boldsymbol{\mu}(k+1)).$$

Writing $\mathbf{g}(\mathbf{b}(k+1),\boldsymbol{\mu}(k+1)) = \mathbf{g}(\mathbf{b}(k+1)) + \boldsymbol{\mu}(k+1)$, applying these two inequalities and then the Cauchy–Schwarz inequality,

$$\begin{aligned}\|\mathbf{g}(\mathbf{b}(k+1),\boldsymbol{\mu}(k+1))\|_2^2 &\leqslant \mathbf{g}(\mathbf{b}(k+1),\boldsymbol{\mu}(k+1))^\top\big(\mathbf{g}(\mathbf{b}(k+1)) + \boldsymbol{\mu}(k)\big)\\ &\leqslant \|\mathbf{g}(\mathbf{b}(k+1),\boldsymbol{\mu}(k+1))\|_2\,\|\mathbf{g}(\mathbf{b}(k+1),\boldsymbol{\mu}(k))\|_2,\end{aligned}$$

and dividing by $\|\mathbf{g}(\mathbf{b}(k+1),\boldsymbol{\mu}(k+1))\|_2$ (the claim is trivial if it is zero) gives:

$$\|\mathbf{g}(\mathbf{b}(k+1),\boldsymbol{\mu}(k+1))\|_2 \leqslant \|\mathbf{g}(\mathbf{b}(k+1),\boldsymbol{\mu}(k))\|_2. \tag{43}$$

Thus, Eq. (42) simplifies to

$$\|\mathbf{g}(\mathbf{b}(k+1),\boldsymbol{\mu}(k+1))\|_2^2 \leqslant \|\mathbf{g}(\mathbf{b}(k),\boldsymbol{\mu}(k))\|_2^2\left(1-\frac{\lambda^2}{L^2}\right)$$

where $\left(1-\frac{\lambda^2}{L^2}\right)<1$. Applying this recursively over k iterations yields

$$\|\mathbf{g}(\mathbf{b}(k),\boldsymbol{\mu}(k))\|_2^2 \leqslant \left(1-\frac{\lambda^2}{L^2}\right)^k\|\mathbf{g}(\mathbf{b}(0),\boldsymbol{\mu}(0)\|_2^2 \tag{44}$$

Under the inequality $(1-x)^k \leqslant e^{-xk}$ and taking the square roots, we obtain an exponential rate of convergence

$$\|\mathbf{g}(\mathbf{b}(k),\boldsymbol{\mu}(k))\|_2 \leqslant e^{-\frac{\lambda^2}{2L^2}k}\|\mathbf{g}(\mathbf{b}(0),\boldsymbol{\mu}(0))\|_2 \tag{45}$$

Step 4: Iteration count. The number of iterations $k = K_\epsilon$ required to achieve a desired error tolerance $\epsilon>0$ is

$$\begin{aligned}&e^{-\frac{\lambda^2}{2L^2}K_\epsilon}\|\mathbf{g}(\mathbf{b}(0),\boldsymbol{\mu}(0))\|_2 = \epsilon\\ \Rightarrow K_\epsilon &= \frac{2L^2}{\lambda^2}\log\left(\frac{\|\mathbf{g}(\mathbf{b}(0),\boldsymbol{\mu}(0)\|_2}{\epsilon}\right)\end{aligned} \tag{46}$$

Step 5: Convergence of the iterates to a Nash equilibrium. By line 3 (Algorithm 1) and (44),

$$\|\mathbf{b}(k+1)-\mathbf{b}(k)\|_2 = \gamma\,\|\mathbf{g}(\mathbf{b}(k),\boldsymbol{\mu}(k))\|_2 \leqslant \gamma\, q^k\,\|\mathbf{g}(\mathbf{b}(0),\boldsymbol{\mu}(0))\|_2$$

with $q := (1-\lambda^2/L^2)^{1/2} < 1$.

Hence, for a step $l > k$, the triangle inequality gives

$$\|\mathbf{b}(l)-\mathbf{b}(k)\|_2 \leqslant \sum_{t=k}^{l-1}\|\mathbf{b}(t+1)-\mathbf{b}(t)\|_2 \leqslant \sum_{t=k}^{l-1}\gamma\, q^t\,\|\mathbf{g}(\mathbf{b}(0),\boldsymbol{\mu}(0))\|_2$$

and the geometric series sum ($\sum_{t=k}^{l-1} q^t = q^k(1-q^{(l-k)})/(1-q) \leqslant q^k/(1-q)$) gives

$$\|\mathbf{b}(l)-\mathbf{b}(k)\|_2 \leqslant \frac{\gamma q^k}{1-q}\|\mathbf{g}(\mathbf{b}(0),\boldsymbol{\mu}(0))\|_2 \xrightarrow[k\to\infty]{} 0.$$

Hence, for every $\epsilon > 0$, we can choose K sufficiently large such that

$$\frac{\gamma q^K}{1-q}\|\mathbf{g}(\mathbf{b}(0),\boldsymbol{\mu}(0))\|_2 \leqslant \epsilon$$

Then for all $l > k \geqslant K$, $\|\mathbf{b}(l)-\mathbf{b}(k)\|_2 \leqslant \epsilon$. So $(\mathbf{b}(k))_k$ is a Cauchy sequence, and by the completeness of $\mathbb{R}^{mn}$ it converges to some $\mathbf{b}^\star \in \mathbb{R}^{mn}$. Since every $\mathbf{b}(k)$ lies in the closed set $\mathcal{D}^\ell$ (Step 2), $\mathbf{b}^\star \in \mathcal{D}^\ell$.

To see that $\mathbf{b}^\star$ is a Nash equilibrium, for a firm s, let $\mathbf{b}'_s \in \mathcal{D}^\ell_s$. By the characterization (33) in the proof of Lemma 1, the correction $\boldsymbol{\mu}_s(k) = \mathbf{g}_s(\mathbf{b}(k),\boldsymbol{\mu}_s(k)) - \mathbf{g}_s(\mathbf{b}(k))$ satisfies (33) with updated budget $\mathbf{b}_s(k+1)$:

$$\big(\mathbf{g}_s(\mathbf{b}(k),\boldsymbol{\mu}_s(k)) - \mathbf{g}_s(\mathbf{b}(k))\big)^\top\big(\mathbf{b}'_s - \mathbf{b}_s(k+1)\big) \geqslant 0 \qquad \forall\, k.$$

As $k\to\infty$, the corrected gradient tends to $\mathbf{0}$ (Step 3), $\mathbf{g}_s(\mathbf{b}(k)) \to \mathbf{g}_s(\mathbf{b}^\star)$ by continuity of the gradient, and $\mathbf{b}_s(k+1) \to \mathbf{b}^\star_s$; hence

$$\mathbf{g}_s(\mathbf{b}^\star)^\top(\mathbf{b}'_s - \mathbf{b}^\star_s) \leqslant 0 \qquad \forall\, \mathbf{b}'_s \in \mathcal{D}^\ell_s$$

Since u_s is concave in $\mathbf{b}_s$ (Step 1), $u_s(\mathbf{b}'_s,\mathbf{b}^\star_{-s}) \leqslant u_s(\mathbf{b}^\star) + \mathbf{g}_s(\mathbf{b}^\star)^\top(\mathbf{b}'_s - \mathbf{b}^\star_s) \leqslant u_s(\mathbf{b}^\star)$. So every firm is best-responding at $\mathbf{b}^\star$, which is therefore a Nash equilibrium. Equivalently, by Lemma 2, $\mathbf{g}(\mathbf{b}^\star,\boldsymbol{\mu}(\mathbf{b}^\star)) = \mathbf{0}$. This also follows directly from the continuity of the corrected gradient: by the proof of Lemma 1, $\mathbf{g}_s(\mathbf{b},\boldsymbol{\mu}_s(\mathbf{b})) = \big[\Pi_{\mathcal{D}^\ell_s}\big(\mathbf{b}_s + \gamma\nabla_{\mathbf{b}_s}u_s(\mathbf{b})\big) - \mathbf{b}_s\big]/\gamma$, which is continuous in $\mathbf{b}$ because the projection onto a non-empty, closed, convex set is non-expansive and $\nabla_{\mathbf{b}_s}u_s$ is continuous; hence $\mathbf{g}(\mathbf{b}^\star,\boldsymbol{\mu}(\mathbf{b}^\star)) = \lim_k \mathbf{g}(\mathbf{b}(k),\boldsymbol{\mu}(k)) = \mathbf{0}$ by Step 3. Thus, the best response dynamics converges to a Nash Equilibrium $\mathbf{b}^\star$ in $K_\epsilon = \mathcal{O}\left(\log\frac{1}{\epsilon}\right)$ iterations. □

C Appendix to Section 5

C.1 Reduction of the SB property to vertex-wise blocks

Both Lemma 3 and Lemma 8 rest on the same reduction, which we record once. The block Hessian of Definition 4 has entries $[G_{s_1,s_2}]_{i,j} = \partial^2 u_{s_1}/(\partial b_{s_2,j}\partial b_{s_1,i}) = M_i\,\partial^2 h_{s_1,i}/(\partial b_{s_2,j}\partial b_{s_1,i})$, since the cost term is linear. Because $h_{s_1,i}$ depends only on the budgets allocated at vertex i, this vanishes whenever $j \neq i$. Reordering coordinates from firm-major order $\mathbf{b} = (b_{1,1},\dots,b_{1,n},\dots b_{m,1},\dots,b_{m,n})^\top$ to

vertex-major order $\widetilde{\mathbf{b}} = (b_{1,1}, \ldots, b_{m,1}, \ldots . b_{1,n}, \ldots, b_{m,n})^\top$ therefore turns G into the block-diagonal matrix

$$\widetilde{G} = \text{block-diag}(H_1, \ldots, H_n), \qquad (H_i)_{s_1,s_2} = M_i \frac{\partial^2 h_{s_1,i}}{\partial b_{s_2,i}\, \partial b_{s_1,i}}, \quad s_1, s_2 \in \mathscr{M}.$$

This reordering is an orthogonal permutation, so G and $\widetilde{G}$ have identical spectra and identical operator norms. Hence the SB property (Definition 4) for some constants $\widetilde{\lambda}, \widetilde{L} > 0$, can be rewritten as:

$$\tfrac{1}{2}\big(\widetilde{G}(\mathbf{b}) + \widetilde{G}(\mathbf{b})^\top\big) + \widetilde{\lambda} I_{mn} \preceq 0 \;\text{ (Strong Concavity)} \quad \text{and} \quad \|\widetilde{G}\|_{\text{op}} \leqslant \widetilde{L} \;\text{ (Bounded Norm).} \tag{47}$$

Since $\widetilde{G}$ is block-diagonal, the spectrum of its symmetric part is the union of the spectra of the symmetric parts of the blocks H_i and $\|\widetilde{G}\|_{\text{op}} = \max_{i \in \mathscr{N}} \|H_i\|_{\text{op}}$. Thus, (47) can be verified block by block: for each $i \in \mathscr{N}$, suppose: $\frac{1}{2}(H_i + H_i^\top) \preceq -\widetilde{\lambda}_i I_m$ and $\|H_i\|_{\text{op}} \leqslant \widetilde{L}_i$. The common SB constants can then be chosen as $\widetilde{\lambda} := \min_i \widetilde{\lambda}_i$ and $\widetilde{L} := \max_i \widetilde{L}_i$ so that the SB property can be stated uniformly as:

$$\tfrac{1}{2}(H_i + H_i^\top) \preceq -\widetilde{\lambda}\, I_m \quad \text{and} \quad \|H_i\|_{\text{op}} \leqslant \widetilde{L}, \qquad \forall\, i \in \mathscr{N}. \tag{48}$$

C.2 Proof of Lemma 3: the Regular RCSF induces the SB property

By Appendix C.1 it suffices to establish the two conditions in (48) for each vertex block H_i. A convenient *sufficient* test for the first condition of (48) (strong concavity) is strict diagonal dominance as defined below, which by *Gershgorin's circle theorem*, guarantees the required spectral bound. The converse does not hold in general.

Definition 11 (Strict diagonal dominance)**.** A matrix $H_i \in \mathbb{R}^{m \times m}$ is said to satisfy the $(\tilde{\lambda}, m)$-strict diagonal dominance condition if for some $\tilde{\lambda} > 0$,

$$|(H_i)_{ss}| - \sum_{r \neq s} |(H_i)_{sr}| \geqslant \tilde{\lambda}, \qquad \forall s \in \mathscr{M},$$

where $|(H_i)_{ss}|$ is the diagonal magnitude and $|(H_i)_{sk}|$ is the magnitude of the off diagonal entries.

For a vertex i and firm $r \in \mathscr{M}$, we will denote $f_r = f(b_{r,i})$, $f_r' = f'(b_{r,i})$, $f_r'' = f''(b_{r,i})$ and $S = \sum_{r \in \mathscr{M}} f_r + \delta$. The required second derivatives for the Regular RCSF are

$$\frac{\partial^2 h_{s,i}}{\partial b_{s,i}^2} = \frac{(S - f_s)}{S^2}\Big(f_s'' - \frac{2(f_s')^2}{S}\Big), \qquad \frac{\partial^2 h_{s,i}}{\partial b_{r,i}\, \partial b_{s,i}} = \frac{f_s'\, f_r'\, (2f_s - S)}{S^3}, \quad r \neq s. \tag{49}$$

Since f is increasing ($f_s' > 0$) and concave ($f_s'' < 0$) and $S - f_s = \sum_{r \neq s} f_r + \delta \geqslant \delta > 0$, $\frac{\partial^2 h_{s,i}}{\partial b_{s,i}^2} < 0$. We use the elementary bounds $S \geqslant \delta$ and $S - f_s \geqslant \delta$, both from $f \geqslant 0$.

Claim 1 (Strong concavity via diagonal dominance)**.** *Under conditions (R1) and (R3) of Definition 5, each block H_i is strictly diagonally dominant with negative diagonal, and there is a constant $\widetilde{\lambda} > 0$ with $\frac{1}{2}(H_i + H_i^\top) \preceq -\widetilde{\lambda} I_m$ for every i. Hence $\frac{1}{2}\big(\widetilde{G}(\mathbf{b}) + \widetilde{G}(\mathbf{b})^\top\big) + \widetilde{\lambda} I_{mn} \preceq 0$.*

Proof. Fix a row s. By (49) and $S - f_s \geqslant \delta$, the diagonal magnitude satisfies

$$\Big|\frac{\partial^2 h_{s,i}}{\partial b_{s,i}^2}\Big| = \frac{(S - f_s)}{S^2}\Big(\frac{2(f_s')^2}{S} - f_s''\Big) \geqslant \frac{(S - f_s)(-f_s'')}{S^2} \geqslant \frac{\delta(-f_s'')}{S^2}.$$

For the off-diagonal entries of row s, using $|2f_s - S| \leqslant S$ (as $0 \leqslant f_s \leqslant S$) and $\sum_{k\neq s} f'_k \leqslant (m-1)\sup_{\mathcal{B}} f'$,

$$\sum_{r\neq s}\Big|\frac{\partial^2 h_{s,i}}{\partial b_{r,i}\partial b_{s,i}}\Big| = \frac{f'_s|2f_s - S|}{S^3}\sum_{r\neq s} f'_r \leqslant \frac{f'_s(m-1)\sup_{\mathcal{B}} f'}{S^2}.$$

Subtracting,

$$\Big|\frac{\partial^2 h_{s,i}}{\partial b_{s,i}^2}\Big| - \sum_{r\neq s}\Big|\frac{\partial^2 h_{s,i}}{\partial b_{r,i}\partial b_{s,i}}\Big| \geqslant \frac{1}{S^2}\Big[\delta(-f''_s) - (m-1)f'_s \sup_{\mathcal{B}} f'\Big]. \tag{50}$$

Condition (R3) of Definition 5 gives

$$(m-1)\,\frac{f'_s}{-f''_s}\,\sup f' \leqslant (m-1)\sup\frac{f'}{-f''}\,\sup f' < \delta$$

$$\Longrightarrow\ (m-1)f'_s\,\sup f' < \delta\,(-f''_s)$$

Hence in (50), the RHS term

$$\delta\,(-f''_s) - (m-1)f'_s\,\sup f' > 0$$

On the compact set $\mathcal{B}$, we have $S \leqslant mf(C) + \delta$ (since $b_{s,i} \leqslant C$) so $\frac{1}{S^2} \geqslant \frac{1}{(mf(C)+\delta)^2}$. Define the smallest possible uniform diagonal-dominance gap:

$$\Delta := \min_{\substack{\mathbf{b}\in\mathcal{B}\\ s\in\mathscr{M},\, i\in\mathcal{N}}} \big\{\delta\big(-f''_s\big) - (m-1)f'_s \sup f'\big\}.$$

Substituting in Definition 11,

$$|(H_i)_{ss}| - \sum_{k\neq s}|(H_i)_{sk}| = M_i\Big[\,\Big|\frac{\partial^2 h_{s,i}}{\partial b_{s,i}^2}\Big| - \sum_{k\neq s}\Big|\frac{\partial^2 h_{s,i}}{\partial b_{k,i}\,\partial b_{s,i}}\Big|\,\Big] \geqslant \frac{M_i\Delta}{(mf(C)+\delta)^2}$$

Since $M_i \geqslant 1-\alpha > 0$ (Lemma 10) and $\Delta > 0$,

$$|(H_i)_{ss}| - \sum_{k\neq s}|(H_i)_{sk}| \geqslant \frac{(1-\alpha)\Delta}{(mf(C)+\delta)^2} =: \widetilde{\lambda} > 0$$

The same bounds apply to the column sums, because the off-diagonal entries of column s are bounded by the same quantity with the roles of f'_s and f'_r exchanged, and $f'_r \leqslant \sup f'$. So the symmetric part $\frac{1}{2}(H_i + H_i^\top)$ inherits strict diagonal dominance with negative diagonal. Consequently, all eigenvalues of $\frac{1}{2}(H_i + H_i^\top)$ are bounded above by $-\widetilde{\lambda}$ implying $\frac{1}{2}(H_i + H_i^\top) \preceq -\widetilde{\lambda}\, I_m$ ((48)). By the reordering above, $\frac{1}{2}\big(\widetilde{G}(\mathbf{b}) + \widetilde{G}(\mathbf{b})^\top\big) + \widetilde{\lambda} I_{mn} \preceq 0$. □

Claim 2 (Bounded operator norm)**.** *Under conditions (R1) and (R2) of Definition 5, every entry of H_i is bounded in magnitude by $M_i\varkappa$, where $\varkappa = 2(\sup f')^2/\delta^2 + \sup|f''|/\delta$. Hence $\|H_i\|_{\mathrm{op}} \leqslant \|H_i\|_F \leqslant mM_i\varkappa$, and $\|\widetilde{G}\|_{\mathrm{op}} \leqslant \widetilde{L}$ with $\widetilde{L} := m\varkappa \max_i M_i < \infty$.*

Proof. Using $S - f_s \leqslant S$, $1/S \leqslant 1/\delta$ and $1/S^2 \leqslant 1/\delta^2$ in (49),

$$\Big|\frac{\partial^2 h_{s,i}}{\partial b_{s,i}^2}\Big| \leqslant \frac{1}{S}\Big(\frac{2(f'_s)^2}{S} + |f''_s|\Big) \leqslant \frac{2(\sup f')^2}{\delta^2} + \frac{\sup|f''|}{\delta} = \varkappa,$$

while the off-diagonal entries satisfy $|\partial^2 h_{s,i}/(\partial b_{k,i}\partial b_{s,i})| \leqslant (\sup f')^2/S^2 \leqslant (\sup f')^2/\delta^2 \leqslant \varkappa$. Thus every entry of H_i is at most $M_i\varkappa$ in magnitude, so $\|H_i\|_F \leqslant mM_i\varkappa$ and $\|H_i\|_{\mathrm{op}} \leqslant \|H_i\|_F \leqslant m\varkappa\max_i M_i := \widetilde{L}$. Thus in (47), $\|\widetilde{G}\|_{\mathrm{op}} \leqslant \widetilde{L} < \infty$. □

Claims 1 and 2 establish Equation (48), so by Appendix C.1 the utilities have the SB property (Definition 4) on $\mathcal{D}^\ell = \mathcal{B}$ with the constants $\lambda = \widetilde{\lambda}$ and $L = \widetilde{L}$. Hence, by Theorem 2, the best-response dynamics (Algorithm 1) converges to a Nash equilibrium $\mathbf{b}^\star$ with step size $\gamma = \widetilde{\lambda}/\widetilde{L}^2$, reaching $\|\mathbf{g}(\mathbf{b}(K_\epsilon), \boldsymbol{\mu}(K_\epsilon))\|_2 < \epsilon$ after $K_\epsilon = O\big((\widetilde{L}/\widetilde{\lambda})^2 \log(1/\epsilon)\big)$ iterations. □

C.3 Proof of Lemma 4: uniqueness and symmetry for the RCSF

Part (I): every NE is symmetric.

Following [13],the game admits a pure strategy symmetric equilibrium because it is a symmetric game with compact and convex strategy spaces and continuous and strictly-concave utility function (from condition 1 of Definition 5, the Regular RCSF $h_{s,i}(\mathbf{b})$ is strictly concave which implies that each firm's utility $u_s(\mathbf{b})$ is also strictly concave in its own decision vector $\mathbf{b}_s$ since it combines a concave total awareness term $\sum_i M_i h_{s,i}(\mathbf{b})$ and the negative of a convex (linear) cost term $\sum_i b_{s,i}$. Now it remains to show that the game doesn't have any asymmetric equilibria. First, we derive an inequality which we call the *marginal dominance (MD) inequality* which shows that at any vertex, the firm bidding strictly more has the strictly smaller marginal return there. This inequality will be used crucial later in ruling out asymmetric equilibria from the game.

Step 1: marginal dominance inequality. Fix a vertex i and two distinct firms $p, q \in \mathscr{M}$. Write $b_{p,i} = x$, $b_{q,i} = y$ with $0 \leqslant y < x \leqslant C$, and let $S = \sum_{r \neq p,q} f(b_{r,i}) \geqslant 0$ denote the fixed sum of all other firms' budgets at i. Isolating firms p and q,

$$h_{p,i}(\mathbf{b}) = h_{p,i}(x, y; S) = \frac{f(x)}{f(x) + f(y) + \delta'}, \qquad \delta' := \delta + S > 0,$$

so that the marginal return for firm p is

$$\frac{\partial h_{p,i}(\mathbf{b})}{\partial b_{p,i}} = \frac{\partial h_{p,i}}{\partial x}(x, y; S) = \frac{f'(x)\big(f(y) + \delta'\big)}{D(x, y)^2}, \qquad D(x, y) := f(x) + f(y) + \delta' \tag{51}$$

and for firm q,

$$\frac{\partial h_{q,i}(\mathbf{b})}{\partial b_{q,i}} = \frac{\partial h_{q,i}}{\partial y}(x, y; S) = \frac{f'(y)\big(f(x) + \delta'\big)}{D(x, y)^2} \tag{52}$$

Since the denominators of (51) and (52) are identical, this simplifies to comparing

$$\frac{f'(x)}{f(x) + \delta'} \text{ and } \frac{f'(y)}{f(y) + \delta'}$$

Let $F(t) := f'(t)/(f(t) + \delta')$. Then

$$F'(t) = \frac{f''(t)\big(f(t) + \delta'\big) - f'(t)^2}{\big(f(t) + \delta'\big)^2}.$$

Since f is concave and increasing, $f'(t) > 0$ and $f''(t) < 0$, and $\delta' > 0$; thus $F'(t) < 0$, so F is strictly decreasing. Hence for $x > y$, $F(y) > F(x)$, which gives

$$\frac{\partial h_{p,i}}{\partial b_{p,i}} < \frac{\partial h_{q,i}}{\partial b_{q,i}}, \qquad \text{for } b_{p,i} > b_{q,i}. \tag{$\star$}$$

Step 2: KKT conditions. Since the utility is strictly concave in its own strategy, KKT conditions are necessary and sufficient. Firm s solves

$$\text{Maximize } u_s(\mathbf{b}) = \sum_i M_i h_{s,i}(\mathbf{b}) - \sum_i b_{s,i}, \qquad \text{subject to } \mathbf{b}_s \geqslant \mathbf{0},\ \sum_i b_{s,i} \leqslant C.$$

Let $\mu_{s,i}, \lambda_s$ be the Lagrange multipliers for the two constraints, with Lagrangian

$$\mathcal{L}(\mathbf{b}, \lambda_s, \mu_{s,i}) = u_s(\mathbf{b}) - \lambda_s\Big(\sum_i b_{s,i} - C\Big) - \sum_i \mu_{s,i}(-b_{s,i}).$$

The KKT conditions are:

$$\text{Stationarity:} \qquad M_i \frac{\partial h_{s,i}(\mathbf{b})}{\partial b_{s,i}} = 1 + \lambda_s - \mu_{s,i}, \tag{S}$$

$$\text{Dual feasibility:} \qquad \mu_{s,i} \geqslant 0, \quad \lambda_s \geqslant 0, \tag{D}$$

$$\text{Complementary slackness:} \qquad \mu_{s,i} b_{s,i} = 0, \quad \lambda_s\Big(C - \sum_i b_{s,i}\Big) = 0. \tag{CS}$$

Step 3: ruling out asymmetric equilibria. Suppose $\mathbf{b}^\star$ is an asymmetric PSNE, i.e. $\mathbf{b}_p^\star \neq \mathbf{b}_q^\star$ for some two firms $p, q \in \mathscr{M}$.

Suppose $b^\star_{p,i_0} > b^\star_{q,i_0} \geqslant 0$ at some vertex i_0. Since $b^\star_{p,i_0} > 0$, (CS) gives $\mu^\star_{p,i_0} = 0$, and (S) gives

$$M_{i_0} \frac{\partial h_{p,i_0}}{\partial b_{p,i_0}}(\mathbf{b}^\star) = 1 + \lambda_p. \tag{53}$$

For firm q, (CS) gives $\mu_{q,i_0} \geqslant 0$, so (S) gives

$$M_{i_0} \frac{\partial h_{q,i_0}}{\partial b_{q,i_0}}(\mathbf{b}^\star) = 1 + \lambda_q - \mu_{q,i_0} \leqslant 1 + \lambda_q. \tag{54}$$

By the marginal dominance inequality ($\star$) at i_0 (the other $m-2$ firms do not matter),

$$\frac{\partial h_{p,i_0}}{\partial b_{p,i_0}}(\mathbf{b}^\star) < \frac{\partial h_{q,i_0}}{\partial b_{q,i_0}}(\mathbf{b}^\star). \tag{55}$$

Combining (53), (54), (55),

$$1 + \lambda_p < 1 + \lambda_q \implies \lambda_p < \lambda_q. \tag{56}$$

That is: whoever bids strictly more at some vertex has a strictly smaller budget multiplier.

Case A: There exist two vertices i, j with $b^\star_{p,i} > b^\star_{q,i}$ and $b^\star_{q,j} > b^\star_{p,j}$. This case arises whenever the two firms have equal total spend, i.e. $\sum_i b^\star_{p,i} = \sum_i b^\star_{q,i}$. Applying (56) at vertex i gives $\lambda_p < \lambda_q$; applying it at vertex j (roles reversed) gives $\lambda_q < \lambda_p$. Contradiction.

Case B: $b^\star_{p,i} \geqslant b^\star_{q,i}$ $\forall i$, with strict inequality at some i_0. Then $\sum_i b^\star_{p,i} > \sum_i b^\star_{q,i}$ (unequal total spend). At i_0, (56) gives $\lambda_p < \lambda_q$.

- If $\sum_i b^\star_{p,i} = C$, then $\sum_i b^\star_{q,i} < C$, so (CS) gives $\lambda_q = 0$. Combined with $\lambda_p < \lambda_q = 0$, this forces $\lambda_p < 0$, violating (D).
- If $\sum_i b^\star_{p,i} < C$, then (CS) gives $\lambda_p = 0$. Then $\lambda_p < \lambda_q$ forces $\lambda_q > 0$, which by (CS) forces $\sum_i b^\star_{q,i} = C$. But this implies $\sum_i b^\star_{q,i} > \sum_i b^\star_{p,i}$, contradicting the case hypothesis.

Hence, in every case, no asymmetric NE is possible. Therefore every NE is symmetric.

Part (II): the symmetric NE is unique.

At the symmetric profile $b_{s,i} = t_i$, let $a(t_i) := \partial h_{s,i}/\partial b_{s,i}$ be the marginal-return function. By the chain rule,

$$a'(t_i) = \left.\frac{\partial^2 h_{s,i}}{\partial b_{s,i}^2}\right|_{\text{symm}} + (m-1)\left.\frac{\partial^2 h_{s,i}}{\partial b_{r,i}\,\partial b_{s,i}}\right|_{\text{symm}}. \tag{57}$$

Writing $D = \sum_s f_s + \delta$ with $f_s := f(b_{s,i})$, direct differentiation of $h_{s,i} = f_s/D$ gives

$$\frac{\partial^2 h_{s,i}}{\partial b_{s,i}^2} = \frac{(D-f_s)}{D^2}\left(f_s'' - \frac{2f_s'^2}{D}\right) < 0, \qquad \text{since } f_s'' < 0,\ f_s' > 0,\ D - f_s > 0.$$

Evaluated at the symmetric point ($D = mf(t_i) + \delta$),

$$\left.\frac{\partial^2 h_{s,i}}{\partial b_{s,i}^2}\right|_{\text{symm}} = \frac{\big((m-1)f(t_i)+\delta\big)}{D^2}\left(f''(t_i) - \frac{2f'(t_i)^2}{D}\right) < 0. \tag{58}$$

Similarly, for $r \neq s$,

$$\frac{\partial^2 h_{s,i}}{\partial b_{r,i}\,\partial b_{s,i}} = -f'(b_{s,i})f'(b_{r,i})\frac{D - 2f(b_{r,i})}{D^3},$$

which at the symmetric point gives

$$\left.\frac{\partial^2 h_{s,i}}{\partial b_{r,i}\,\partial b_{s,i}}\right|_{\text{symm}} = -\big(f'(t_i)\big)^2\,\frac{(m-2)f(t_i)+\delta}{D^3} < 0. \tag{59}$$

Substituting (58) and (59) into (57) gives $a'(t_i) < 0$; hence a is strictly decreasing in t_i, hence injective.

Rewriting the KKT conditions for a symmetric equilibrium profile $\mathbf{t}^\star = (t_i^\star)_{i\in\mathscr{N}}$ with common multipliers $\lambda \geqslant 0$ and $\mu_i \geqslant 0$ (shared across firms by symmetry) :

$$M_i\, a(t_i^\star) = 1 + \lambda - \mu_i$$

If $t_i^\star > 0$, then by complementary slackness, $\mu_i = 0$, the KKT becomes $M_i\, a(t_i^\star) = \lambda + 1 =: \Gamma$. If $t_i^\star = 0$, then $\mu_i \geqslant 0$ and we have $M_i\, a(0) = (\lambda+1) - \mu_i \leqslant \Gamma$.

Since a is continuous and strictly decreasing, so is $M_i\, a(\cdot)$ for each fixed $M_i > 0$, and it is therefore invertible. Define

$$t_i^\star = \beta_i(\Gamma) := \max\left\{a^{-1}\left(\frac{\Gamma}{M_i}\right),\ 0\right\}. \tag{60}$$

Then $\beta_i(\Gamma)$ is single-valued and (weakly, strictly wherever positive) decreasing in Γ; hence for any $\Gamma > 0$ the budget $t_i^\star$ is unique. Next we prove that this Γ is also unique for a given game instance.

Now consider the aggregate budget $S(\Gamma) := \sum_i \beta_i(\Gamma)$, and it is also continuous and strictly decreasing on the region where it is positive, with $S(\Gamma) \to 0$ as $\Gamma \to \infty$. Since the total budget constraint is $\sum_i t_i^\star \leqslant C$, two cases arise considering the smallest possible $\Gamma = 1$:

(1) $S(1) \leqslant C$: the unconstrained budget at the lowest possible marginal threshold $\Gamma = 1$ already satisfies the constraint, thus the budget multiplier is inactive, i.e. $\lambda = 0$. Thus there is a unique $\Gamma^\star = 1$, and $t_i^\star = \beta_i(1)$ is the unique solution.

(2) $S(1) > C$: the budget constraint is active, $\lambda > 0$. Since $\lim_{\Gamma\to\infty} S(\Gamma) = 0$, there exists a finite $\Gamma_1 > 1$ such that $S(\Gamma_1) < C$. Since S is continuous on $(0,\infty)$, it is continuous on the closed interval $[1,\Gamma_1]$ with $S(1) > C > S(\Gamma_1)$. The Intermediate Value Theorem then implies that there exists a $\Gamma^\star > 1$ with $S(\Gamma^\star) = C$, and strict monotonicity implies that the $\Gamma^\star$ is unique as a strictly monotone continuous function crosses any given level exactly once. Thus $t_i^\star = \beta_i(\Gamma^\star)$.

Thus in both the cases, there is a unique $\Gamma^\star \geqslant 1$ satisfying the budget constraint and hence at most one symmetric profile $\mathbf{t}^\star = (\beta_i(\Gamma^\star))_{i\in\mathcal{N}}$ satisfies the KKT system, so the symmetric equilibrium is unique. Together with **Part (I)**, the game admits a unique Nash equilibrium, and it is symmetric. □

C.4 Proof of Theorem 3

Step 1 (Reduction to symmetric profiles). The utility of a firm s is $u_s(\mathbf{b}) = \sum_i M_i h_{s,i}(\mathbf{b}) - \sum_i b_{s,i}$ (3) . Let $h_{s,i}(\mathbf{b}) = f(b_{s,i})/(F_i+\delta)$ with $F_i = \sum_r f(b_{r,i})$. Since $\sum_s f(b_{s,i})/(F_i+\delta) = F_i/(F_i+\delta)$ for every i, utilitarian welfare W_1 depends on $\mathbf{b}$ only through the aggregates $F_i(\mathbf{b}) := \sum_s f(b_{s,i})$ and $B_i(\mathbf{b}) := \sum_s b_{s,i}$:

$$W_1(\mathbf{b}) = \frac{1}{m}\sum_s u_s(\mathbf{b}) = \frac{1}{m}\left[\sum_i M_i \frac{F_i(\mathbf{b})}{F_i(\mathbf{b})+\delta} - \sum_i B_i(\mathbf{b})\right].$$

For $\mathbf{b}\in\mathcal{B}$, let $\bar{\mathbf{b}}$ be its *symmetrization*, $\bar{b}_{s,i} := \frac{1}{m}\sum_r b_{r,i}$ for every s. Then $\bar{\mathbf{b}}\in\mathcal{B}$, since $\sum_i \bar{b}_{s,i} = \frac{1}{m}\sum_r\sum_i b_{r,i} \leqslant C$, and symmetrization preserves aggregate spend exactly, $B_i(\bar{\mathbf{b}}) = B_i(\mathbf{b})$. By concavity of f (condition (R1) of Definition 5), Jensen's inequality gives $f(\bar{b}_{s,i}) \geqslant \frac{1}{m}\sum_r f(b_{r,i}) = F_i(\mathbf{b})/m$ for every s, so $F_i(\bar{\mathbf{b}}) = \sum_s f(\bar{b}_{s,i}) \geqslant F_i(\mathbf{b})$: symmetrization weakly increases the aggregate score. Since $t \mapsto t/(t+\delta)$ is strictly increasing, $F_i(\bar{\mathbf{b}}) \geqslant F_i(\mathbf{b})$ together with $B_i(\bar{\mathbf{b}}) = B_i(\mathbf{b})$ give $W_1(\bar{\mathbf{b}}) \geqslant W_1(\mathbf{b})$: symmetrization never decreases average welfare. Because $\mathcal{B}$ is compact and W_1 continuous, a maximizer of W_1 over $\mathcal{B}$ exists; its symmetrization is itself a maximizer, and is symmetric. Consequently, if a welfare-maximizing profile exists, there exists one on the symmetric slice, and it suffices to compare $\mathbf{b}^\star$ which is symmetric by Lemma 4, against a welfare-maximizing profile restricted to symmetric profiles.

Step 2 (The two KKT systems). Restricted to a symmetric profile $b_{s,i} = t_i$, firm s's stationarity condition is

$$M_i a(t_i) - 1 = \lambda_s - \mu_{s,i}, \qquad \mu_{s,i}\geqslant 0,\ \mu_{s,i}t_i = 0,\ \lambda_s \geqslant 0,\ \lambda_s\big(C - \textstyle\sum_i t_i\big) = 0,$$

while the planner, maximizing $U := mW_1 = \sum_r u_r$, has

$$M_i\psi(t_i) - 1 = \hat{\lambda} - \hat{\mu}_i, \qquad \hat{\mu}_i\geqslant 0,\ \hat{\mu}_i t_i = 0,\ \hat{\lambda} \geqslant 0,\ \hat{\lambda}\big(C - \textstyle\sum_i t_i\big) = 0,$$

where a and ψ are as in Definition 6. (The planner faces one budget constraint per firm; at a symmetric profile these carry a common multiplier, which we write as $\hat{\lambda}$ after rescaling by m.) A symmetric profile is simultaneously a Nash equilibrium and a welfare optimum only if some choice of multipliers satisfies both systems at once.

Step 3 (The wedge is strictly positive). Writing $D = \sum_r f_r + \delta$ with $f_r := f(b_{r,i})$, direct differentiation of $h_{s,i} = f_s/D$ gives $\partial h_{s,i}/\partial b_{s,i} = f_s'(D-f_s)/D^2$ and, for $r\neq s$, $\partial h_{r,i}/\partial b_{s,i} = -f_r f_s'/D^2$. At the symmetric profile, where $f_r = f(t_i)$ for every r and $D = mf(t_i)+\delta$,

$$a(t_i) = \frac{f'(t_i)\big[(m-1)f(t_i)+\delta\big]}{\big(mf(t_i)+\delta\big)^2}, \quad d(t_i) = -\frac{f'(t_i)\,f(t_i)}{\big(mf(t_i)+\delta\big)^2}, \quad \psi(t_i) = \frac{\delta\, f'(t_i)}{\big(mf(t_i)+\delta\big)^2}. \tag{61}$$

The δ term cancels in $w = a - \psi$ and the $(m-1)f(t_i)$ term survives:

$$w(t_i) = \frac{(m-1)\, f(t_i)\, f'(t_i)}{\big(mf(t_i)+\delta\big)^2} \;>\; 0 \qquad \text{for every } t_i > 0,\ m \geqslant 2, \tag{62}$$

since f is increasing with $f(0) = 0$, so $f(t_i) > 0$ for $t_i > 0$, and $f'(t_i) > 0$.

Step 4 (Failure of welfare optimality). On the symmetric slice evaluated at $b_{s,i} = t_i$, W_1 reduces to $\sum_i M_i f(t_i)/(mf(t_i)+\delta) - \sum_i t_i$. Each summand composes the concave increasing map $x \mapsto x/(mx+\delta)$ with the concave map f, hence is concave, so the planner's restricted problem is a concave maximization subject to a linear constraint and KKT stationarity is necessary *and* sufficient for optimality.

At any vertex i with $t_i > 0$, complementary slackness gives $\mu_{s,i} = \hat{\mu}_i = 0$, so the two stationarity conditions of Step 2 reduce to

$$M_i a(t_i) = 1 + \lambda_s, \qquad M_i \psi(t_i) = 1 + \hat{\lambda}, \tag{63}$$

with $\lambda_s, \hat{\lambda} \geqslant 0$ shared across vertices. This holds regardless of whether the respective budget constraints bind, i.e. regardless of whether $\lambda_s, \hat{\lambda}$ are zero or strictly positive.

Case (i): heterogeneous centralities. Suppose $\mathbf{b}^\star = \{t_i^\star\}_{i\in\mathscr{N}}$ is active at two vertices i, j with $M_i \neq M_j$, i.e. $t_i^\star, t_j^\star > 0$. If $\mathbf{b}^\star$ also satisfied the planner's condition in (63), dividing the two equations at each vertex would give $a(t_i^\star)/\psi(t_i^\star) = (1+\lambda_s)/(1+\hat{\lambda}) = a(t_j^\star)/\psi(t_j^\star)$. By (61), $a(t_i)/\psi(t_i) = \big[(m-1)f(t_i)+\delta\big]/\delta$, which is strictly increasing in t_i because f is strictly increasing. Hence $t_i^\star = t_j^\star$, and substituting back into (63), using that a is strictly decreasing and therefore injective (see the proof of Lemma 4), gives $M_i = (1+\lambda_s)/a(t_i^\star) = (1+\lambda_s)/a(t_j^\star) = M_j$, contradicting $M_i \neq M_j$. So $\mathbf{b}^\star$ fails the planner's stationarity condition for every $\hat{\lambda} \geqslant 0$, and by the concavity noted above it does not maximize W_1 even on the symmetric slice. Such instances exist. Take $m = 2$, $f(b) = 1 - e^{-2b}$, $\delta = 1.2$ (so $\delta > m-1$ and the CSF is a Regular RCSF), $C = 0.5$, and the network of Table 2, whose centralities $M = (0.810, 1.210, 1.131, 0.958, 0.891)$ are pairwise distinct. The unique equilibrium is $\mathbf{t}^\star = (0.0468, 0.1235, 0.1095, 0.0769, 0.0636)$, active at all five consumers, with $W_1(\mathbf{b}^\star) = 0.0994$, whereas $\max_{\mathcal{B}} W_1 = 0.1053$, attained at $\mathbf{t}^{\text{opt}} = (0.0371, 0.0957, 0.0852, 0.0604, 0.0501)$. They are also the typical case, because equal centralities require a very special network. These instances alone therefore prove the theorem.

Case (ii): homogeneous centrality. Suppose instead that all vertices carry the same centrality $M_i = M$.

Lemma 10. *For every row-stochastic $\mathscr{E}$ and every $\alpha \in (0,1)$, $\sum_{i\in\mathscr{N}} M_i = n$ and $M_i \geqslant 1-\alpha$ for every $i \in \mathscr{N}$.*

Proof. Row-stochasticity gives $\mathscr{E}\mathbf{1} = \mathbf{1}$, hence $(I-\alpha\mathscr{E})^{-1}\mathbf{1} = \sum_{k\geqslant 0}\alpha^k\mathscr{E}^k\mathbf{1} = \big(\sum_{k\geqslant 0}\alpha^k\big)\mathbf{1} = (1-\alpha)^{-1}\mathbf{1}$. Therefore $\sum_i M_i = (1-\alpha)\,\mathbf{1}^\top(I-\alpha\mathscr{E})^{-1}\mathbf{1} = \mathbf{1}^\top\mathbf{1} = n$. For the lower bound, note that every term $\alpha^k\mathscr{E}^k$ of the Neumann series has non-negative entries, so $(I-\alpha\mathscr{E})^{-1} \geqslant I$ entrywise. Each column sum of $(I-\alpha\mathscr{E})^{-1}$ is therefore at least 1, and multiplying by $1-\alpha$ gives $M_i \geqslant 1-\alpha$. □

The above lemma forces $M_i = 1$ for every i (e.g. uniform complete graph). Then (63) does not by itself determine the relation between the firm and planner budget multipliers λ_s and $\hat{\lambda}$, so the wedge could in principle be absorbed by their difference. We rule this out by exhibiting an instance whose optimum is interior ($t_i^{\text{opt}} > 0$) and leaves the budget unexhausted ($\sum_i t_i^{\text{opt}} < C$), which forces the budget multipliers $\hat{\lambda} = 0$ and $\hat{\mu}_i = 0$. The planner's stationarity condition is therefore

$M_i\psi(t_i) = 1$, which reduces to $\psi(t) = 1$ as $M_i = 1$ and all vertices hence have the same budget allocation $t_i = t$.

Take the saturating exponential generator $f(t) = 1 - e^{-\beta t}$ and choose parameters with

$$m - 1 < \delta < \beta, \qquad \text{and} \qquad n\, t^{\text{opt}} < C, \tag{64}$$

where t^{opt} is the symmetric welfare optimal budget which is same for all vertices i (as defined later). The first chain of inequalities is consistent: $\delta > m - 1$ makes the CSF a Regular RCSF, and $\delta < \beta = f'(0)$ is compatible with it. Now $\psi(t) = \delta f'(t)/(mf(t) + \delta)^2$ is continuous and strictly positive, with $\psi(0) = f'(0)/\delta = \beta/\delta > 1$, and $\psi(t) \to 0$ as $t \to \infty$ (if f is bounded, concavity forces $f' \to 0$; if f is unbounded, $mf(t)+\delta \to \infty$). which implies $\psi(t) < 1$ for all sufficiently large t. By the intermediate value theorem, there is some interior critical point $t^{\text{opt}} \in (0, \infty)$ satisfying $\psi(t^{\text{opt}}) = 1$, and $t^{\text{opt}} \to 0$ as $\beta \downarrow \delta$, so the second condition in (64) can always be met by taking β close enough to δ. Since the welfare objective restricted to the symmetric slice is concave, the solution $\mathbf{b}^{\text{opt}} := (t^{\text{opt}}, \ldots, t^{\text{opt}})$ is the average welfare maximizer on that slice. By Step 1, symmetrization does not decrease average welfare, hence this symmetric planner optimum is also a maximizer of W_1 over the full feasible set $\mathcal{B}$.

We now show that this welfare optimum cannot simultaneously be a Nash equilibrium. Suppose, to the contrary, that $\mathbf{b}^\star = \mathbf{b}^{\text{opt}}$ were a Nash equilibrium. Since $\sum_i t_i^{\text{opt}} = nt^{\text{opt}} < C$, the firm's budget is slack, so $\lambda_s = 0$; and since $t^{\text{opt}} > 0$ the lower-bound multipliers vanish. Then (63) gives $a(t^{\text{opt}}) = 1$ and $\psi(t^{\text{opt}}) = 1$. Subtracting the two equations gives $a(t^{\text{opt}}) - \psi(t^{\text{opt}}) = w(t^{\text{opt}}) = 0$, contradicting Step 3. So $\mathbf{b}^{\text{opt}}$ is not a Nash equilibrium, and by uniqueness $\mathbf{b}^\star \neq \mathbf{b}^{\text{opt}}$. Since $\mathbf{b}^{\text{opt}}$ maximizes W_1 over $\mathcal{B}$, $\mathbf{b}^\star$ does not.

In either case $\mathbf{b}^\star$ fails to maximize W_1, which proves the theorem. □

C.5 Proof of Proposition 1

Let $(\lambda_s, \mu_{s,i})$ denote the multipliers of firm s's best-response problem $\max_{\mathbf{b}_s \in \mathcal{B}_s} u_s(\mathbf{b}_s, \mathbf{b}_{-s}^{\text{opt}})$ and $(\hat{\lambda}, \hat{\mu}_i)$ those of the planner's problem, both evaluated at the profile $\mathbf{b}^{\text{opt}}$. Because all constraints are linear, the KKT conditions are necessary at both the equilibrium and the optimum. (As established in the proof of Theorem 3 (Appendix C.4), the planner's budget multipliers are equal at a symmetric profile and the KKT conditions are the same as in Step 2).

(i) Budget exhaustion suffices. Substituting $\psi = ca$ into the planner's stationarity condition $M_i\psi(t_i^{\text{opt}}) - 1 = \hat{\lambda} - \hat{\mu}_i$ and adding and subtracting c on the left gives

$$c\big(M_i a(t_i^{\text{opt}}) - 1\big) + (c - 1) = \hat{\lambda} - \hat{\mu}_i,$$

so, dividing by $c > 0$,

$$M_i a(t_i^{\text{opt}}) - 1 \;=\; \underbrace{\frac{\hat{\lambda} + (1 - c)}{c}}_{=:\,\lambda_s} \;-\; \underbrace{\frac{\hat{\mu}_i}{c}}_{=:\,\mu_{s,i}}. \tag{65}$$

This is exactly firm s's stationarity condition $M_i a(t_i^{\text{opt}}) - 1 = \lambda_s - \mu_{s,i}$, with the multipliers relabelled. We check the remaining KKT requirements for the relabelled pair. Non-negativity holds: $\mu_{s,i} = \hat{\mu}_i/c \geqslant 0$ since $c > 0$, and $\lambda_s = (\hat{\lambda} + 1 - c)/c > 0$ since $\hat{\lambda} \geqslant 0$ and $c < 1$. Complementary slackness on the lower bound is inherited, since $\mu_{s,i} t_i^{\text{opt}} = \hat{\mu}_i t_i^{\text{opt}}/c = 0$. The only condition left is complementary slackness on the budget, $\lambda_s\big(C - \sum_i t_i^{\text{opt}}\big) = 0$. Because $c < 1$ forces $\lambda_s > 0$ strictly, this holds precisely because the budget is exhausted by hypothesis, i.e. $\sum_i t_i^{\text{opt}} = C$. So $\mathbf{b}^{\text{opt}}$ satisfies every firm's KKT system, and concavity of u_s in $\mathbf{b}_s$ makes these conditions sufficient, so $\mathbf{b}^{\text{opt}}$ is a best response for every firm, i.e. a Nash equilibrium.

(ii) Budget exhaustion is necessary when the optimum is nonzero. Suppose $\mathbf{b}^{\text{opt}} \neq \mathbf{0}$ is a Nash equilibrium and, for contradiction, that $\sum_i t_i^{\text{opt}} < C$. Complementary slackness on the budget then forces $\lambda_s = 0$ in every firm's system and $\hat{\lambda} = 0$ in the planner's. Since $\mathbf{b}^{\text{opt}} \neq \mathbf{0}$, there is a vertex j with $t_j^{\text{opt}} > 0$, and complementary slackness on the lower bound gives $\mu_{s,j} = \hat{\mu}_j = 0$. The two stationarity conditions at j therefore reduce to $M_j\, a(t_j^{\text{opt}}) = 1$ and $M_j\, \psi(t_j^{\text{opt}}) = 1$, implying $\psi(t_j^{\text{opt}}) = a(t_j^{\text{opt}})$. By (12) this forces $c = 1$, contradicting $c < 1$.

The hypothesis $\mathbf{b}^{\text{opt}} \neq \mathbf{0}$ is needed because the all-zero profile is not ruled out by the proportional-marginals condition alone: if $t_i^{\text{opt}} = 0$ at every vertex, the two stationarity conditions for the firm and planner become $M_i a(0) - 1 = -\mu_{s,i} \implies \mu_{s,i} = 1 - M_i a(0)$ and $M_i \psi(0) - 1 = -\hat{\mu}_i \implies \hat{\mu}_i = 1 - M_i c a(0)$ respectively. So whenever $M_i a(0) \leqslant 1$ for every i, both multipliers $\mu_{s,i}$ and $\hat{\mu}_i$ are non-negative and both complementary-slackness conditions $\mu_{s,i} t_i^{\text{opt}} = \hat{\mu}_i t_i^{\text{opt}} = 0$ hold trivially. The all-zero profile can then be a Nash equilibrium and a welfare optimum simultaneously with non-exhausted budgets, so no contradiction arises. □

C.6 Proof of Lemma 5

Dividing numerator and denominator of (15) by $f(B_i)$ gives

$$c(B_i) = \frac{m\, e(B_i)}{e(B_i) + m - 1}. \tag{66}$$

The map $x \mapsto mx/(x+m-1)$ is strictly increasing on $x > 0$, hence injective, so $c(B_i)$ is constant in B_i if and only if $e(B_i)$ is constant. For finding a function with constant elasticity η, i.e. $e(B_i) \equiv \eta$, solving the differential equation $f'(B_i)/f(B_i) = \eta/B_i$ gives the general solution $f(B_i) = \kappa B_i^{\eta}$; the requirements that f be increasing with $f(0) = 0$ force $\eta > 0$, and the requirement in Definition 7 that f be concave forces $\eta \leqslant 1$, since $f''(B_i) = \kappa\eta(\eta - 1)B_i^{\eta-2}$. Conversely $f(B_i) = \kappa B_i^{\eta}$ has $e(B_i) \equiv \eta$ identically. □

C.7 Proof of Proposition 2

For the Power RCSF, the utility of a firm s is concave in $\mathbf{b}_s$. This can be verified by the second derivative of firm s's utility with respect to $b_{s,i}$, which is $M_i \kappa(\eta - 1) B_i^{\eta-3}\big(2\sum_{r\neq s} b_{r,i} + \eta b_{s,i}\big) < 0$ for $\eta \in (0, 1)$ and $B_i > 0$.

Since $f(B_i) = \kappa B_i^{\eta}$ is iso-elastic, i.e. $e(B_i) \equiv \eta$, by Lemma 5, the Power RCSF has proportional marginals. Hence, (66) gives $c = m\eta/(\eta + m - 1)$, a constant. Since $\eta < 1$ we have $\eta(m-1) < m-1$, i.e. $m\eta < m + \eta - 1$, so $c < 1$; and $c > 0$ since $m \geqslant 2$ and $\eta > 0$. Explicitly, evaluating (14) at $B_i = mt_i$ gives $a(t_i) = \kappa m^{\eta-2}(m + \eta - 1)\, t_i^{\eta-1}$ and $\psi(t_i) = \kappa\eta\,(mt_i)^{\eta-1}$. □

C.8 Proof of Proposition 3

(i) For the power RCSF, $\sum_s h_{s,i} = \kappa B_i^{\eta}$, so utilitarian welfare W_1 depends on $\mathbf{b}$ only through the total budget $B_i = \sum_s b_{s,i}$:

$$W_1(\mathbf{b}) = \frac{1}{m}\sum_s u_s(\mathbf{b}) = \frac{1}{m}\Big[\kappa \sum_i M_i B_i^{\eta} - \sum_i B_i\Big],$$

and the attainable totals over $\mathcal{B}$ are exactly $\{\mathbf{B} \geqslant \mathbf{0} : \sum_i B_i \leqslant mC\}$ where any such $\mathbf{B}$ is realized by the symmetric split $t_i = B_i/m$. This objective is separable and strictly concave in $\mathbf{B}$, with unconstrained maximizer given by solving the planner's unconstrained first-order condition

$M_i\psi(t_i) = M_i\kappa\eta B_i^{\eta-1} = 1$, giving $B_i^{\text{opt}} = (\kappa\eta M_i)^\theta$. Hence at the welfare optimum, the full budget is exhausted if and only if the unconstrained maximizer requires at least the available aggregate budget, i.e., $\sum_i B_i^{\text{opt}} \geqslant mC \implies \sum_i (\kappa\eta M_i)^\theta \geqslant mC$ which is (19). In this case the constrained welfare optimum satisfies $\sum_i B_i = mC$, which at a symmetric allocation is equivalent to $\sum_i t_i = C$ for every firm s.

(ii) Substituting the largest admissible $\kappa = (mC)^{-\eta}$ into (19) gives $\eta^\theta (mC)^{-\eta\theta} \sum_i M_i^\theta \geqslant mC$. Since $1 + \eta\theta = 1 + \eta/(1-\eta) = \theta$, this rearranges to $\eta^\theta \sum_i M_i^\theta \geqslant (mC)^\theta$, and taking θ^{th} roots gives $\eta\|M\|_\theta \geqslant mC$. Any smaller κ makes the left side of (19) smaller, so this is necessary for every admissible κ.

(iii) For $\theta \geqslant 1$ the ℓ^θ norm is dominated by the ℓ^1 norm, so $\|M\|_\theta \leqslant \sum_i M_i = n$ by Lemma 10. With $\eta < 1$ this gives $\eta\|M\|_\theta < n \leqslant mC$, contradicting (ii). It remains to explain why this rules out alignment. By part (i), the symmetric welfare-maximizing profile then leaves every budget unexhausted. Moreover, since $M_i \geqslant 1 - \alpha > 0$ (Lemma 10) and the planner's marginal return $M_i\kappa\eta B_i^{\eta-1}$ grows without bound as $B_i \to 0$, moving a small amount of budget to any vertex with $B_i = 0$ strictly increases W_1, so $t_i^{\text{opt}} > 0$ on every vertex i, i.e. $\mathbf{b}^{\text{opt}} \neq 0$. Proposition 1(ii) therefore applies, and it shows that this profile cannot be a Nash equilibrium. □

C.9 Proof of Proposition 4

Under the taxed utilities, the firm's first-order condition (FOC) becomes

$$M_i a(t_i) - (1+\tau) = \lambda_s - \mu_{s,i}. \tag{67}$$

whereas the planner's FOC remains unchanged:

$$M_i\psi(t_i) - 1 = \hat{\lambda} - \hat{\mu}_i. \tag{68}$$

Let $\mathbf{b}^{\text{opt}}$ be a welfare-maximizing symmetric profile with common vertex values $b_{s,i}^{\text{opt}} = t_i^{\text{opt}}$ for every firm $s \in \mathscr{M}$. Substituting $\psi = ca$ in the planner's FOC (68) and dividing by c,

$$M_i a(t_i^{\text{opt}}) = \frac{1 + \hat{\lambda} - \hat{\mu}_i}{c} = \frac{1}{c} + \frac{\hat{\lambda}}{c} - \frac{\hat{\mu}_i}{c} = (1+\tau) + \lambda_s - \mu_{s,i}, \qquad \lambda_s := \frac{\hat{\lambda}}{c}, \quad \mu_{s,i} := \frac{\hat{\mu}_i}{c},$$

using $1 + \tau = 1/c$. We obtain exactly the firm's FOC (67) with the relabeled multipliers λ_s and $\mu_{s,i}$. These multipliers are non-negative because $c > 0$, and both complementary slackness conditions are inherited directly from the planner's, since $\mu_{s,i} t_i^{\text{opt}} = \hat{\mu}_i t_i^{\text{opt}}/c = 0$ and $\lambda_s(C - \sum_i t_i^{\text{opt}}) = \hat{\lambda}(C - \sum_i t_i^{\text{opt}})/c = 0$. Since $\widetilde{u}_s$ is concave in $\mathbf{b}_s$, KKT is necessary and sufficient, so $\mathbf{b}^{\text{opt}}$ is a best response for every firm. □

C.10 Proof of Lemma 6: uniqueness and symmetry for the Power RCSF

Part I: The equilibrium is always symmetric.
Throughout, firms face the common unit charge $1 + \tau \geqslant 1$, so utilities are $\widetilde{u}_s$ as in (22). For the Power RCSF,

$$\frac{\partial^2 h_{s,i}}{\partial b_{s,i}^2} = \kappa(\eta-1)B_i^{\eta-3}\big[2B_i + (\eta-2)b_{s,i}\big] = \kappa(\eta-1)B_i^{\eta-2}\big[2 + (\eta-2)\frac{b_{s,i}}{B_i}\big]$$

Since $\eta - 2 \in (-2,-1)$ and $\frac{b_{s,i}}{B_i} \leqslant 1$, the bracket is at least $2 + (\eta-2) = \eta > 0$, and $\kappa(\eta-1) < 0$, so the second derivative is strictly negative everywhere. Hence $\widetilde{u}_s$ is strictly concave in $\mathbf{b}_s$ for any

fixed $\mathbf{b}_{-s}$. Hence the game admits a pure strategy symmetric equilibrium because it is a symmetric game with compact and convex strategy spaces and continuous and strictly-concave utility function [13]. Now it remains to show that the game doesn't have any asymmetric equilibria.

This proof follows similar arguments as in the proof of Lemma 4 in Appendix C.3 establishing the uniqueness and symmetry of the equilibrium under the Regular RCSF. For the Power RCSF, we derive the same *marginal dominance (MD) inequality* as follows:

Fix a vertex i and firms p, q, where $b_{p,i} = x$, $b_{q,i} = y$ with $0 \leqslant y < x \leqslant C$ and $S = \sum_{r \neq p,q} b_{r,i}$, so $B_i = x + y + S$. By (17), $\partial h_{p,i}/\partial x = \kappa B_i^{\eta-2}[B_i - (1-\eta)x]$ and $\partial h_{q,i}/\partial y = \kappa B_i^{\eta-2}[B_i - (1-\eta)y]$. Subtracting the two, we get

$$\frac{\partial h_{p,i}}{\partial x} - \frac{\partial h_{q,i}}{\partial y} = \kappa B_i^{\eta-2}(1-\eta)(y-x) \;<\; 0,$$

since $\kappa B_i^{\eta-2} > 0$ and $1 - \eta > 0$. Thus,

$$\frac{\partial h_{p,i}}{\partial b_{p,i}} \;<\; \frac{\partial h_{q,i}}{\partial b_{q,i}}, \qquad \text{for } b_{p,i} > b_{q,i}.$$

which is exactly ($\star$) for the Power RCSF.

Since the taxed utility (22) is strictly concave in its own strategy, KKT conditions are necessary and sufficient. Since the rebate ρ_s does not depend on $\mathbf{b}_s$, the KKT conditions are:

$$\text{Stationarity:} \qquad M_i \frac{\partial h_{s,i}(\mathbf{b})}{\partial b_{s,i}} = 1 + \tau + \lambda_s - \mu_{s,i}, \tag{S}$$

$$\text{Dual feasibility:} \qquad \mu_{s,i} \geqslant 0, \quad \lambda_s \geqslant 0, \tag{D}$$

$$\text{Complementary slackness:} \qquad \mu_{s,i} b_{s,i} = 0, \quad \lambda_s \Big(C - \sum_i b_{s,i} \Big) = 0. \tag{CS}$$

Now we can write the same arguments as in Appendix C.3 and produce the same contradictions on the budget multipliers in order to rule out all asymmetric equilibria as those arguments use the *MD inequality* ($\star$) which has been derived here too.

Suppose $\mathbf{b}^\star$ is an asymmetric PSNE, i.e. $\mathbf{b}_p^\star \neq \mathbf{b}_q^\star$ for some two firms $p, q \in \mathscr{M}$. Suppose $b_{p,i_0}^\star > b_{q,i_0}^\star \geqslant 0$ at some vertex i_0. Since $b_{p,i_0}^\star > 0$, (CS) gives $\mu_{p,i_0}^\star = 0$, and (S) gives

$$M_{i_0} \frac{\partial h_{p,i_0}}{\partial b_{p,i_0}}(\mathbf{b}^\star) = 1 + \tau + \lambda_p. \tag{69}$$

For firm q, (CS) gives $\mu_{q,i_0} \geqslant 0$, so (S) gives

$$M_{i_0} \frac{\partial h_{q,i_0}}{\partial b_{q,i_0}}(\mathbf{b}^\star) = 1 + \tau + \lambda_q - \mu_{q,i_0} \leqslant 1 + \tau + \lambda_q. \tag{70}$$

By the marginal dominance inequality ($\star$) at i_0 (the other $m-2$ firms do not matter),

$$\frac{\partial h_{p,i_0}}{\partial b_{p,i_0}}(\mathbf{b}^\star) < \frac{\partial h_{q,i_0}}{\partial b_{q,i_0}}(\mathbf{b}^\star). \tag{71}$$

Combining (69), (70), (71),

$$1 + \tau + \lambda_p < 1 + \tau + \lambda_q \implies \lambda_p < \lambda_q. \tag{72}$$

That is: whoever bids strictly more at some vertex has a strictly smaller budget multiplier.

Case A: There exist two vertices i, j with $b^\star_{p,i} > b^\star_{q,i}$ and $b^\star_{q,j} > b^\star_{p,j}$. This case arises whenever the two firms have equal total spend, i.e. $\sum_i b^\star_{p,i} = \sum_i b^\star_{q,i}$. Applying (72) at vertex i gives $\lambda_p < \lambda_q$; applying it at vertex j (roles reversed) gives $\lambda_q < \lambda_p$. Contradiction.

Case B: $b^\star_{p,i} \geqslant b^\star_{q,i}\ \forall i$, with strict inequality at some i_0. Then $\sum_i b^\star_{p,i} > \sum_i b^\star_{q,i}$ (unequal total spend). At i_0, (72) gives $\lambda_p < \lambda_q$.

- If $\sum_i b^\star_{p,i} = C$, then $\sum_i b^\star_{q,i} < C$, so (CS) gives $\lambda_q = 0$. Combined with $\lambda_p < \lambda_q = 0$, this forces $\lambda_p < 0$, violating (D).
- If $\sum_i b^\star_{p,i} < C$, then (CS) gives $\lambda_p = 0$. Then $\lambda_p < \lambda_q$ forces $\lambda_q > 0$, which by (CS) forces $\sum_i b^\star_{q,i} = C$. But this implies $\sum_i b^\star_{q,i} > \sum_i b^\star_{p,i}$, contradicting the case hypothesis.

Hence, in every case, no asymmetric NE is possible. Therefore every NE is symmetric.

Part II: uniqueness of the symmetric equilibrium.
At a symmetric profile with $B_i = mt_i$, (17) gives the closed form for the private marginal return

$$a(t_i) = \kappa\, m^{\eta-2}(m+\eta-1)\, t_i^{\eta-1} = K\, t_i^{\eta-1}, \tag{73}$$

where $K = \kappa\, m^{\eta-2}(m+\eta-1) > 0$. So $a'(t_i) = K(\eta-1)t_i^{\eta-2} < 0$ since $\eta < 1$ and hence $a(t_i)$ is strictly decreasing on $t_i \in (0,\infty)$ with $a(t_i) \to +\infty$ as $t_i \to 0^+$.

Rewriting the KKT conditions for a symmetric equilibrium profile $\mathbf{t}^\star = (t_i^\star)_{i\in\mathscr{N}}$ with common multipliers $\lambda \geqslant 0$ and $\mu_i \geqslant 0$ (shared across firms by symmetry) :

$$M_i\, a(t_i^\star) = 1 + \tau + \lambda - \mu_i$$

If $t_i^\star > 0$, then by complementary slackness, $\mu_i = 0$, the KKT becomes $M_i\, a(t_i^\star) = \lambda + 1 + \tau =: \Gamma$. If $t_i^\star = 0$, then $\mu_i \geqslant 0$ and we have $M_i\, a(0^+) = (\lambda+1+\tau) - \mu_i \leqslant \Gamma$. Note that since $a(t)$ is only defined for $t > 0$, we write $a(0^+) := \lim_{t_i^\star \downarrow 0} a(t_i^\star) = +\infty$. Hence for any vertex with $M_i > 0$, $M_i a(0^+) = +\infty$ and at the boundary solution $t_i^\star = 0$, the KKT condition which requires $M_i a(0^+) \leqslant \Gamma$ cannot hold for any finite Γ. Thus the symmetric solution $t_i^\star > 0$ always. This simplifies the argument relative to the RCSF case (Appendix C.3), where $a(0)$ is finite and corner solutions ($t_i^\star = 0$) are possible.

Since a is continuous and strictly decreasing, so is $M_i\, a(\cdot)$ for each fixed $M_i > 0$, and it is therefore invertible. Let $\theta := 1/(1-\eta)$. Inverting (73) to obtain the symmetric equilibrium solution $t_i^\star$,

$$t_i^\star = \beta_i(\Gamma) := a^{-1}\Big(\frac{\Gamma}{M_i}\Big) = \Big(\frac{M_i K}{\Gamma}\Big)^{\theta}, \qquad \Gamma > 0,$$

which is single-valued and strictly decreasing in Γ because $\theta > 0$. This proves that for each $\Gamma > 0$, the budget $t_i^\star$ is unique. Consequently the aggregate demand for budget, $S(\Gamma) = \sum_i \beta_i(\Gamma)$ is continuous and strictly decreasing on all of $\Gamma \in (0,\infty)$, with $S(\Gamma) \to 0$ as $\Gamma \to \infty$. Since the total budget constraint is $\sum_i t_i^\star \leqslant C$, two cases arise considering the smallest possible $\Gamma = 1+\tau$:

(1) $S(1+\tau) \leqslant C$: the unconstrained budget at the lowest possible marginal threshold $\Gamma = 1+\tau$ already satisfies the constraint, thus the budget multiplier is inactive, i.e. $\lambda = 0$. Thus there is a unique $\Gamma^\star = 1+\tau$, and $t_i^\star = \beta_i(1+\tau)$ is the unique solution.

(2) $S(1+\tau) > C$: the budget constraint is active, $\lambda > 0$. Since $\lim_{\Gamma\to\infty} S(\Gamma) = 0$, there exists a finite $\Gamma_1 > 1+\tau$ such that $S(\Gamma_1) < C$. Since S is continuous on $(0,\infty)$, it is continuous on the closed interval $[1+\tau, \Gamma_1]$ with $S(1+\tau) > C > S(\Gamma_1)$. The Intermediate Value Theorem then implies that there exists a $\Gamma^\star > 1+\tau$ with $S(\Gamma^\star) = C$, and strict monotonicity implies that the $\Gamma^\star$ is unique as a strictly monotone continuous function crosses any given level exactly once. Thus $t_i^\star = \beta_i(\Gamma^\star)$.

The two cases are mutually exclusive and exhaustive, determined by the sign of $S(1+\tau)-C$, and each yields at most one candidate. In both, $S(\Gamma^\star)=\min\{S(1+\tau),C\}$. Combined with *Part I*, the equilibrium is unique and symmetric. □

C.11 Proof of Proposition 5

Fix i and let $\mathbf{b}^{(B_i)}\in\mathcal{B}$ place $b_{s,i}=B_i/m$ at i for every firm s, distributing each firm's remaining budget $C-B_i/m$ over the other $n-1\geqslant 1$ vertices; this is feasible for every $B_i\in(0,mC]$. Differentiating (16) twice gives

$$\frac{\partial^2 h_{s,i}}{\partial b_{s,i}^2}=\kappa(\eta-1)B_i^{\eta-3}\big[2B_i+(\eta-2)b_{s,i}\big],\qquad \frac{\partial^2 h_{s,i}}{\partial b_{r,i}\,\partial b_{s,i}}=\kappa(\eta-1)B_i^{\eta-3}\big[B_i+(\eta-2)b_{s,i}\big]. \tag{74}$$

At $\mathbf{b}^{(B_i)}$, where $b_{s,i}=B_i/m$, the diagonal entry of block H_i has magnitude:

$$|(H_i)_{ss}|=M_i\kappa(1-\eta)\frac{2m+\eta-2}{m}B_i^{\eta-2}$$

and $2m+\eta-2>0$ for $m\geqslant 2$ and $\eta\in(0,1)$. Since $\|H_i\|_{\mathrm{op}}\geqslant|\mathbf{e}_s^\top H_i\mathbf{e}_s|=|(H_i)_{ss}|$ for the sth standard basis vector $\mathbf{e}_s$, we get

$$\|H_i\|_{\mathrm{op}}\geqslant M_i\kappa(1-\eta)\,\tfrac{2m+\eta-2}{m}\,B_i^{\eta-2},$$

and this operator norm diverges as $B_i\to 0^+$ because $\eta-2<0$. By the reduction of SB property to vertex-wise blocks H_i in Appendix C.1, the unboundedness of $\|H_i\|_{\mathrm{op}}$ implies that the bounded-operator-norm condition in Definition 4 cannot hold for any finite L. □

C.12 Proof of Lemma 7

Let

$$\bar\varepsilon_1:=\Big(\frac{K\min_i M_i}{1+\tau}\Big)^\theta,\qquad \bar\varepsilon_2:=C\,\frac{\min_i M_i^\theta}{\sum_j M_j^\theta},$$

so that $\bar\varepsilon=\min\{\bar\varepsilon_1,\bar\varepsilon_2\}$ by definition. Recall from Lemma 6 that $t_i^\star=\beta_i(\Gamma^\star)$, where $\beta_i(\Gamma):=(M_iK/\Gamma)^\theta$, and $\Gamma^\star$ is the unique value with $S(\Gamma^\star)=\min\{S(1+\tau),C\}$, $S(\Gamma):=\sum_i\beta_i(\Gamma)$. We show $\min_i t_i^\star=\bar\varepsilon$ by identifying, in each of the two cases of $S(\Gamma)$ in the proof of Lemma 6 (Appendix C.10), which of $\bar\varepsilon_1,\bar\varepsilon_2$ equals $\min_i t_i^\star$ exactly, and then verifying that this value is always the smaller of the two, so that taking the minimum of the two constant expressions recovers the correct one in either case.

Since $\beta_i(\Gamma)=(K/\Gamma)^\theta M_i^\theta$ for every i, the factor $(K/\Gamma)^\theta$ cancels in the ratio $\beta_i(\Gamma)/S(\Gamma)=M_i^\theta/\sum_j M_j^\theta$, which is therefore independent of Γ; multiplying through by $S(\Gamma)$ and taking the minimum over i gives, for every $\Gamma>0$,

$$\min_i\beta_i(\Gamma)\;=\;S(\Gamma)\,\frac{\min_i M_i^\theta}{\sum_j M_j^\theta}. \tag{75}$$

Case (i): $S(1+\tau)\leqslant C$. Here $\Gamma^\star=1+\tau$, so applying (75) with $S(\Gamma)=(K/\Gamma)^\theta\sum_j M_j^\theta$ at $\Gamma=1+\tau$,

$$\min_i t_i^\star\;=\;\min_i\beta_i(1+\tau)\;=\;\Big(\frac{K}{1+\tau}\Big)^\theta\min_i M_i^\theta\;=\;\Big(\frac{K\min_i M_i}{1+\tau}\Big)^\theta\;=\;\bar\varepsilon_1.$$

Because $\theta = 1/(1-\eta) > 0$, the map $x \mapsto x^\theta$ is increasing on $x \geqslant 0$ which implies $\min_i M_i^\theta = (\min_i M_i)^\theta$.

We check that $\bar{\varepsilon}_1 \leqslant \bar{\varepsilon}_2$ in this regime. Writing $\bar{\varepsilon}_1 = (K/(1+\tau))^\theta \min_i M_i^\theta$ and using $S(1+\tau) = (K/(1+\tau))^\theta \sum_j M_j^\theta \leqslant C$, we get $(K/(1+\tau))^\theta \leqslant C/\sum_j M_j^\theta$. Multiplying both sides by $\min_i M_i^\theta \geqslant 0$ preserves the inequality:

$$\bar{\varepsilon}_1 \;=\; \Big(\frac{K}{1+\tau}\Big)^\theta \min_i M_i^\theta \;\leqslant\; C\,\frac{\min_i M_i^\theta}{\sum_j M_j^\theta} \;=\; \bar{\varepsilon}_2.$$

Hence $\min\{\bar{\varepsilon}_1, \bar{\varepsilon}_2\} = \bar{\varepsilon}_1 = \min_i t_i^\star$ throughout this regime.

Case (ii): $S(1+\tau) > C$. Here $\Gamma^\star$ is the value with $S(\Gamma^\star) = C$, so applying (75) at $\Gamma = \Gamma^\star$,

$$\min_i t_i^\star \;=\; \min_i \beta_i(\Gamma^\star) \;=\; S(\Gamma^\star)\,\frac{\min_i M_i^\theta}{\sum_j M_j^\theta} \;=\; C\,\frac{\min_i M_i^\theta}{\sum_j M_j^\theta} \;=\; \bar{\varepsilon}_2.$$

We check that $\bar{\varepsilon}_2 < \bar{\varepsilon}_1$ in this regime. From $S(1+\tau) > C$, i.e. $(K/(1+\tau))^\theta \sum_j M_j^\theta > C$, we get $(K/(1+\tau))^\theta > C/\sum_j M_j^\theta$. Multiplying both sides of this inequality by $\min_i M_i^\theta > 0$ (since every vertex i has $M_i \geqslant 1-\alpha > 0$ (Lemma 10)) gives

$$\bar{\varepsilon}_1 \;=\; \Big(\frac{K}{1+\tau}\Big)^\theta \min_i M_i^\theta \;>\; C\,\frac{\min_i M_i^\theta}{\sum_j M_j^\theta} \;=\; \bar{\varepsilon}_2.$$

Hence $\min\{\bar{\varepsilon}_1, \bar{\varepsilon}_2\} = \bar{\varepsilon}_2 = \min_i t_i^\star$ throughout this regime as well.

Since Cases (i) and (ii) are exhaustive and in each of them $\min_i t_i^\star$ coincides with the smaller of $\bar{\varepsilon}_1, \bar{\varepsilon}_2$, we conclude that, regardless of which case actually holds,

$$\min_i t_i^\star \;=\; \min\{\bar{\varepsilon}_1, \bar{\varepsilon}_2\} \;=\; \bar{\varepsilon}.$$

Moreover $\bar{\varepsilon}_1, \bar{\varepsilon}_2 > 0$, since $K, \min_i M_i, C > 0$, so $\bar{\varepsilon} = \min\{\bar{\varepsilon}_1, \bar{\varepsilon}_2\} > 0$.

Now let ε satisfy $0 < \varepsilon < \bar{\varepsilon}$. Since $\bar{\varepsilon} = \min_i t_i^\star$, every coordinate satisfies $t_i^\star \geqslant \bar{\varepsilon} > \varepsilon$, so the floor is inactive at the equilibrium for every such ε. Finally, $\min_i M_i^\theta / \sum_j M_j^\theta \leqslant 1/n$ (from Lemma 10, the smallest of n non-negative numbers never exceeds their average, so $\min_i M_i^\theta \leqslant \frac{1}{n}\sum_j M_j^\theta$), so $\bar{\varepsilon} \leqslant C/n$ and $n\varepsilon < n\bar{\varepsilon} \leqslant C$. Hence $\mathcal{B}^\varepsilon$ is non-empty. □

C.13 Why is a participation floor needed?

Proposition 5 is a statement about the Power RCSF family (9) and the SB-based convergence argument used here, rather than a general impossibility result for two reasons: the SB property (Definition 4) is sufficient but not necessary for BRD convergence, and Lemma 5 characterizes the alignment condition only within the SRCSF class (13).

In practice, there are three ways to address the unbounded curvature as $B_i \to 0$. First, one could modify the CSF near zero: since the equilibrium and welfare optimum lie strictly above $\bar{\varepsilon}$ (Lemma 7), one could choose $B_0 < m\bar{\varepsilon}$, retain the iso-elastic form on $[B_0, mC]$, and replace $f(B_i)$ on $[0, B_0)$ by a function with bounded second derivative, as is common for smoothing a contest success function near zero Elkind et al. [17], Ghosh [24]. Second, one could retain the Power RCSF and replace the SB-based convergence argument with a method that tolerates unbounded curvature. Third, one could restrict the strategy space so that B_i is bounded away from zero, which is what the participation floor does.

Within the SRCSF class and the SB-based BRD convergence framework, exact iso-elasticity on the whole feasible set and globally bounded curvature cannot hold together, so one of the three

must be chosen. We choose the floor because it leaves the Power RCSF unchanged and, for $\varepsilon < \bar{\varepsilon}$, it is inactive at the solution. Because the corrective charge aligns the firms' and the planner's KKT systems whether or not budgets are exhausted (Proposition 4), the choice of ε need not depend on which case of (25) applies. The floor's only role is to restore the SB property, and Lemma 7 confirms that it does so without disturbing the solution.

C.14 Proof of Lemma 8: the Power RCSF satisfies SB on the floored set

By Appendix C.1 it suffices to establish (48) for each vertex block H_i, and the linear cost term (with or without the charge) does not affect H_i.

Structure. Direct differentiation of the Power RCSF (16) gives (from (74)),

$$(H_i)_{ss} = M_i\kappa(\eta-1)B_i^{\eta-3}\,[2B_i + (\eta-2)b_{s,i}] = c_0\,(v_s + B_i)$$

$$(H_i)_{sr} = M_i\kappa(\eta-1)B_i^{\eta-3}\,[B_i + (\eta-2)b_{s,i}] = c_0 v_s \;\text{ for } r \neq s$$

where $c_0 := M_i\kappa(\eta-1)B_i^{\eta-3} < 0$ and $v_s := B_i + (\eta-2)b_{s,i}$. Hence the vertex-wise Hessian block can be written in the matrix form as $H_i = c_0(\mathbf{v}\mathbf{1}^\top + B_i I_m)$.

Strong concavity. $\frac{1}{2}(H_i + H_i^\top) = c_0\mathsf{M}$ with $\mathsf{M} := B_i I_m + \frac{1}{2}(\mathbf{v}\mathbf{1}^\top + \mathbf{1}\mathbf{v}^\top)$. Since $c_0 < 0$, it suffices to show $\mathsf{M} \succeq \chi(m,\eta)B_i\, I_m$. The symmetric matrix $\frac{1}{2}(\mathbf{v}\mathbf{1}^\top + \mathbf{1}\mathbf{v}^\top)$ has rank at most two (since $\mathrm{Rank}(\mathbf{v}\mathbf{1}^\top) = \mathrm{Rank}(\mathbf{1}\mathbf{v}^\top) = 1$), so it vanishes on the orthogonal complement of $\mathrm{span}\{\mathbf{1},\mathbf{v}\}$. Hence it suffices to analyze its restriction to $\mathrm{span}\{\mathbf{1},\mathbf{v}\}$ on which it has eigenvalues $\mu_\pm = \frac{1}{2}\big(\mathbf{1}^\top\mathbf{v} \pm \sqrt{m}\,\|\mathbf{v}\|_2\big)$ and its remaining eigenvalues are zero, with multiplicity at least $m-2$. By Cauchy–Schwarz inequality, $\mathbf{1}^\top\mathbf{v} \leqslant \sqrt{m}\;\|\mathbf{v}\|$, which implies $\mu_- \leqslant 0$, so the minimum eigenvalue of matrix M is $\lambda_{\min}(\mathsf{M}) = B_i + \mu_-$. Next we have to find a lower bound on μ_- so that we can bound $\lambda_{\min}(\mathsf{M})$.

$$\mu_- = \tfrac{1}{2}\big(\mathbf{1}^\top\mathbf{v} - \sqrt{m}\,\|\mathbf{v}\|_2\big)$$

Now $\mathbf{1}^\top\mathbf{v} = \sum_s \big[B_i + (\eta-2)b_{s,i}\big] = (m+\eta-2)B_i$, and, using $\sum_s b_{s,i}^2 \leqslant B_i^2$, which holds for any nonnegative vector summing to B_i,

$$m\,\|\mathbf{v}\|_2^2 = m\Big[(m+2\eta-4)\,B_i^2 + (2-\eta)^2\sum_s b_{s,i}^2\Big] \;\leqslant\; m\,(m-2\eta+\eta^2)\,B_i^2.$$

Substituting in the minimum eigenvalue of M:

$$\lambda_{\min}(\mathsf{M}) = B_i + \tfrac{1}{2}\Big[(m+\eta-2)B_i - \sqrt{m}\,\|\mathbf{v}\|_2\Big] \;\geqslant\; \frac{m+\eta-\sqrt{m(m-2\eta+\eta^2)}}{2}\,B_i \;=\; \chi(m,\eta)\,B_i.$$

By rationalizing the denominator of $\chi(m,\eta)$, we get

$$\chi(m,\eta) = \frac{\eta\,\big(4m-(m-1)\eta\big)}{2\,\big[\,m+\eta+\sqrt{m(m-2\eta+\eta^2)}\,\big]} \;>\; 0,$$

since $(m+\eta)^2 - m(m-2\eta+\eta^2) = \eta\big(4m-(m-1)\eta\big) > 0$ for $\eta \in (0,1)$ and $m \geqslant 2$ (as $(m-1)\eta < m-1 < 4m$). Hence, $\mathsf{M} \succeq \chi(m,\eta)B_i\, I_m$ and multiplying by $c_0 = -M_i\kappa(1-\eta)B_i^{\eta-3} < 0$ and using $B_i \leqslant mC \Rightarrow B_i^{\eta-2} \geqslant (mC)^{\eta-2}$ (since $\eta-2<0$ reverses the inequality), gives $\frac{1}{2}(H_i + H_i^\top) \preceq -\widetilde{\lambda} I_m$ for every i and hence in (47), $\frac{1}{2}\big(\widetilde{G}(\mathbf{b}) + \widetilde{G}(\mathbf{b})^\top\big) + \widetilde{\lambda} I_{mn} \preceq 0$ where

$$\widetilde{\lambda} := \kappa(1-\eta)\chi(m,\eta)(mC)^{\eta-2}\min_i M_i \;>\; 0$$

Since the strong-concavity bound uses only the upper bound $B_i \leqslant mC$ and holds on the whole unrestricted set, so the participation floor doesn't play any role here.

Bounded operator norm. For $0 \leqslant b_{s,i} \leqslant B_i$ and $\eta - 2 \in (-2, -1)$ we have $|2B_i + (\eta-2)b_{s,i}| \leqslant 2B_i$ and $|B_i + (\eta-2)b_{s,i}| \leqslant B_i$, so every entry of H_i has magnitude at most $M_i\kappa(1-\eta)B_i^{\eta-3} \cdot 2B_i = 2M_i\kappa(1-\eta)B_i^{\eta-2} \leqslant 2M_i\kappa(1-\eta)(m\varepsilon)^{\eta-2}$, using $B_i \geqslant m\varepsilon$. Hence $\|H_i\|_{\mathrm{op}} \leqslant \|H_i\|_F \leqslant m \cdot 2M_i\kappa(1-\eta)(m\varepsilon)^{\eta-2} \leqslant 2m\,\kappa\,(1-\eta)\,(m\varepsilon)^{\eta-2}\,\max_i M_i := \widetilde{L}$, and in (47), $\|\widetilde{G}\|_{\mathrm{op}} \leqslant \widetilde{L}$. This bound uses the lower bound $B_i \geqslant m\varepsilon$ and is therefore the step at which the participation floor is essential. □