Targeted Advertising Platforms: Data Sharing and Customer Poaching*

Klajdi Hoxha†

August 21, 2025

[Preliminary Draft: Please Do Not Circulate]

Abstract

E-commerce platforms are rolling out ambitious *targeted advertising* initiatives that rely on merchants *sharing customer data* with each other via the platform. Yet current platform designs fail to address participating merchants' concerns about customer *poaching*. This paper proposes a model of designing welfare improving targeted advertising platforms that account for merchants' willingness to share customer data in the presence of poaching concerns. Using a mechanism design framework, I characterize the optimal mechanism and show that it allocates targeted ads in bundles according to a merchant-specific scoring rule weighted by customer-specific click-through rates (CTR). Some degree of customer poaching may still occur, in which case negatively impacted merchants are compensated with monetary transfers. In large platforms, the optimal mechanism is to design three markets: i) a *selling market* where merchants can sell all their data at posted price p; ii) an *exchange market* where merchants share all the data in exchange for high CTR ads; iii) a *buying market* where high value merchants pay the full price to buy high CTR ads. The model is broad in scope with applications in other market design settings like the greenhouse gas credit markets and reallocating public resources (e.g., water/land rights, fishing quotas).

Keywords: Exchange market, CTR, customer data, bundling, mechanism design, platform, poaching, scoring rule, targeted advertising.

*I am grateful to Michael Ostrovsky, Andrzej Skrzypacz and Weijie Zhong for their guidance and support. I am thankful to Bing Liu for numerous discussions during the early stages of the paper, as well as Spencer Pantoja, Ilya Segal, Eric Tang, Frank Yang and seminar participants at Stanford University for helpful comments and suggestions. I owe special thanks to Varanya Chaubey for exceptional editing advice.

The author declares that he has no relevant or material financial interests that relate to the research described in this paper.

†Graduate School of Business, Stanford University. Email: hklajdi@stanford.edu.

1 Introduction

E-commerce platforms are rolling out ambitious targeted advertising initiatives that rely on merchants sharing customer data with each other via the platform. These initiatives are in response to recent privacy policies—such as California's CCPA, European Union's GDPR and Apple's ATT—limiting third-party cookie tracking, a cornerstone of targeted online advertising.[1] With Safari, Firefox and Edge having phased out third-party cookies, Chrome is now the only major browser still relying on them as it seeks alternative solutions.[2] One promising solution involves e-commerce platforms pooling customer data from merchants' own datasets. Doing so helps the platform build a larger and more targeted dataset, thereby delivering targeted ads to customers from their preferred merchants.

Yet current platform designs fail to address participating merchants' concerns about customer poaching. These designs give participating merchants the opportunity to run advertising campaigns on the platform that target every customer in the shared dataset. To determine which merchants get to target which customers, such platforms commonly allocate targeted ads by running ad auctions based on bid amounts and ad quality scores, where the latter measures the probability of a click. If competitors with ads at least of similar quality bid more aggressively, merchants' ad campaigns may reach fewer customers than before joining the platform.

Our running example that illustrates this setting is Shopify, an e-commerce platform that powers the (online) stores of millions of merchants. In 2022, Shopify pioneered these initiatives by launching Shopify Audiences, a tool designed to pool customer data from participating merchants to improve their targeted advertising. However, a frequently asked question on Shopify's official page reveals a growing concern: can merchants "exclude competitors from using their data"?[3] Shopify's ambiguous response—"...the data you contribute will be used to benefit other participating merchants"—only validates the customer poaching concerns.

This paper proposes a model of designing targeted advertising platforms and solves for the optimal platform design. The novelty in the model stems from the heterogeneity in customers' click-through rates (CTRs) for each merchant's product and the exclusivity of customer datasets initially held by individual merchants. This modeling choice is motivated by the existing inefficiencies of merchants' individual targeting efforts. Using a mechanism design framework, I solve for an optimal mechanism that improves upon individual targeting efforts while ensuring merchants' willingness to share their customer data. I show that the optimal mechanism allocates targeted ads based on a merchant-specific scoring rule that accounts for the characteristics of each merchants' initial customer dataset, while also factoring in bid amounts and ad quality scores. Under the optimal mechanism, targeted ads are allocated and priced in bundles to reduce information rents, and some degree of customer poaching may occur.

[1] 96% of iOS users have opted out of app tracking and 80% of advertisers rely on third-party cookies according to https://www.wordstream.com/blog/ws/2021/10/26/online-advertising-statistics and https://www.imore.com/96-iphone-users-have-opted-out-app-tracking-ios-145-launched

[2] https://privacysandbox.com/news/privacy-sandbox-update/

[3] https://help.shopify.com/en/manual/promoting-marketing/shopify-audiences/faq

Specifically, the model includes a platform designer designing targeted advertising campaigns to allocate targeted ads across a (finite) number of merchants, with the objective of maximizing a weighted sum of platform revenue and social value. Outside the platform, each merchant can run their own targeted ads using their exclusive customer dataset. On the platform, however, participating merchants compete for the ability to target any customer within the shared dataset. Since customers have different CTRs for each ad, the platform can improve upon current targeting efforts outside the platform by improving merchant-customer matchings. Merchants earn a private profit-margin per click.[4] To elicit this information truthfully, the designer assigns a participation fee—possibly negative—and credibly commits to some advertising mechanism such that all merchants are willing to participate.

I show that the optimal advertising (direct) mechanism allocates targeted ads in bundles according to a (non-decreasing) merchant-specific scoring rule that is a function of merchant's type. Specifically, each customer in the dataset receives the ad from the merchant whose score, weighted by the customer-specific CTRs, is highest. In the case of ties, the designer constructs a customer-specific tie-breaking rule that ensures participation constraints bind.

A single merchant-specific scoring rule is used to allocate all targeted ads on the platform to reduce merchants' information rents by designing bundling strategies. On the platform, merchants anticipate retaining advertising access to only their high CTR customers, while selling off the rest to the platform. At the same time, they anticipate buying access to high CTR customers previously part of competing merchants' customer datasets. As a result, countervailing incentives arise: merchants have incentives to over report when selling and under report when buying. To exploit these countervailing incentives, the designer bundles all the selling and buying transactions by constructing a single scoring rule based on merchants' initial customer datasets.

As a result, the designer implements a single, non-itemized participation fee—possibly a negative when the losses due to poaching exceed the gains of targeting new high CTR customers (if any). To alleviate poaching concerns, most adversely affected merchants are assigned a (weakly) higher score at each type. However, due to the differences in click probabilities and merchants' types, some poaching may still occur in the optimum. In the case when a merchant's net clicks fall below their pre-platform level, the designer provides sufficient compensation (i.e. negative participation fee) to guarantee their participation.

The multidimensionality aspect of the tie-breaking rules often makes characterization of the optimal mechanism untractable. In the second part of the paper I extend the model to a continuum of CTRs and show that ties occur with probability zero. I then provide full characterization of the scoring rule and optimal targeted advertising mechanism which reduce to solving a system of nonlinear equations. I then consider a stylized setting for large number of merchants and show that the optimal mechanism is ex-post implementable by optimally designing three markets: a selling, a buying and an exchange market. The later is a novel market compared to the standard bid/ask price mechanisms. Due to highly heterogeneous customer data, the designer finds it cheaper to compensate high value merchants with high CTR ads instead of involving monetary transfers.

The model is broad in scope and I show its applications in market design settings like greenhouse gas credit market and reallocating public resources. Essentially, any market

[4] I assume for simplicity that customers always convert (i.e. by the product) if they click on an ad.

design settings in which the agents own goods of heterogeneous quality such that initial allocation is inefficient with regards to the private value for the goods and the (observed) qualities.

Related Literature. Theoretically, the model generalizes the models of bilateral trading (Myerson and Satterthwaite (1983)) and partnership dissolution (Cramton et al. (1987), Loertscher and Wasser (2019)) where there trading of a number of heterogeneous goods. It also relates to the bundling literature since the optimal design trades data in bundles. Yang (2023) studies nested bundling in a framework with one-dimensional heterogeneity where preferences are non-additive. In my model I have one-dimensional heterogeneity and preferences are non-additive. In canonical multidimensional mechanism design problems with one-dimensional heterogeneity and additive preferences bundling strategies does not improve seller's objective. However, due to countervailing incentives for different goods, there's "pooling of incentives" which give rise to optimal bundling strategies.

On the practical side, the paper contributes to the vast literature on digital platforms. Bergemann and Bonatti (2024) and Bergemann et al. (2024) study digital platforms where advertisers can reach consumers on and off the platform with a focus on optimal ad pricing and efficient product matches. Some studies consider information design and pricing when the database is controlled by the platform (see Bergemann and Bonatti (2015), Elliott et al. (2022)). See also Bergemann et al. (2022), Galperti et al. (2024) and Galperti and Perego (2022) for settings when consumers control how much data to provide to these intermediaries and its implications. Lastly, see Farboodi and Veldkamp (2023) for a comprehensive review on data and markets.

Roadmap. The remainder of the paper proceeds as follows. Section 2 presents the model. Section 3 presents some illustrative examples. Section 4 presents the optimal mechanism in the finite case, sketches the main proof and characterizes the mechanism in a stylized setting. Section 4 extends the model in the continuum case and presents further results. Section 6 revisits model assumptions and presents some applications. Section 7 concludes. Omitted proofs are given in the appendices.

2 Model

Motivated by the inefficiencies of decentralized targeting, the novel feature in the model is that merchants can initially run off-platform campaigns targeting their exclusive customers who differ in their click-through rates (CTRs).

Specifically, there's a designer who designs a targeted advertising platform that allocates targeted ads to a finite set of merchants $i \in \mathcal{N} = \{1, \dots, N\}$ where $N \geq 2$. Each merchant $i \in \mathcal{N}$ supplies products at a profit margin (type) $\theta_i \in \Theta_i \subset \mathbb{R}_+$ per unit sold, which is private information. Designing target ads relies on having customer data—which participating merchants provide—about a *unit-mass* of customers in the market. A customer is represented by their CTRs $\omega = (\omega_i)_{i \in \mathcal{N}}$ where $\omega_i \in \Omega_i \subseteq [0, 1]$ denotes the probability that the customer intends to purchase merchant i's product. Product discovery is entirely driven by

advertising and customers have unit-demands. Merchant i's expected revenue from targeting a customer of CTR ω is given by $\theta_i\omega_i$.[5]

Each merchant $i \in \mathcal{N}$ owns exclusive access to target a subset of customers in the economy. Formally, let $\Omega = \times_{i\in\mathcal{N}}\Omega_i$ denote the set of feasible customer CTRs with $|\Omega| = K < \infty$, and $\alpha \in \Delta(\mathcal{N}\times\Omega)$ be a *discrete* distribution (probability measure) that represents the *aggregate dataset* in the economy. For every $i \in \mathcal{N}$ and $\omega \in \Omega$, α_i^ω gives the mass of customers with CTR ω on merchant i's individual dataset. Let $\alpha^\Omega \in \Delta(\mathcal{N})$ and $\alpha_\mathcal{N} \in \Delta(\Omega)$ be the marginal distributions of α where α_i^Ω gives the total mass of customer on the *individual dataset* of merchant $i \in \mathcal{N}$, and $\alpha_\mathcal{N}^\omega$ gives the total mass of customers in the economy that have CTR $\omega \in \Omega$.

Distributional Assumptions. Types θ_i lie in the compact set $\Theta_i = [\underline{\theta}, \overline{\theta}]$ and let $\theta = (\theta_i)_{i\in\mathcal{N}}$ denote a type profile that lies in $\Theta = \times\Theta_i$. Assume types θ_i are drawn from independent Borel probability measures $dF_i \in \Delta(\Theta_i)$ with cdf F_i that admits continuous and strictly positive densities f_i, and let the product measure $dF := \bigotimes_{i\in\mathcal{N}} dF_i$. In addition, assume that the cdf F_i is *regular*: the virtual value and virtual cost functions defined by $\phi_i^B(\theta_i) = \theta_i - (1 - F_i(\theta_i))/f_i(\theta_i)$ and $\phi_i^S(\theta_i) = \theta_i + F_i(\theta_i)/f_i(\theta_i)$, respectively, are strictly increasing. Lastly, the virtual type function $\phi_i(\theta_i, \theta_i')$ at the *critical type* θ_i' is defined by $\phi_i(\theta_i, \theta_i') = \phi_i^S(\theta_i)$ for all $\theta_i < \theta_i'$ and $\phi_i(\theta_i, \theta_i') = \phi_i^B(\theta_i)$ for all $\theta_i \geq \theta_i'$.

Informational Assumptions. The primitives F_i and α are common knowledge. The latter assumes that the designer and merchants agree on each CTR ω and each merchant-individual dataset.[6] I discuss in Section 6 the substance of the later assumption by connecting it to the literature on mechanism design with limited commitment. For the first I point the reader to the vast literature on 'detail-free' mechanism design, also known as the Wilson doctrine.

2.1 Platform Design

I describe how the platform operates by specifying the policy on data sharing and the design of targeted advertising campaigns.

Data Sharing. The designer requires participating merchants to give up all of their ownership rights to target own exclusive customers by sharing the corresponding merchant-specific dataset. Note that the designer can always allocate the same targeted ads that each merchant run off-platform. Assuming the designer can do so at a negligible cost, it follows readily that it is without loss of generality to restrict to settings where every merchant participates and share all their data.

[5] For a micro-foundation of the merchant's payoff function, let each merchant be having a private marginal cost of c_i of selling a unit of the product. An interested customer's demand function is given by willingness to pay $v_i \sim G_i$. At the optimal posted price $p^*(c_i)$ the sellers' private profit margin is then $\theta_i = (p^* - c_i)G_i(p^*)$.

[6] For example, using data about customers' past shopping behavior provided by the merchants, it is sufficient for the merchants and the designer to agree on some *predicted* click probabilities ω_i.

Targeted Advertising Campaigns. Let x_i^ω be the quantity of ads displaying merchant i's product to CTR ω customers, where the total quantity is subject to feasibility constraints

$$\sum_{i\in\mathcal{N}} x_i^\omega \leq \alpha_{\mathcal{N}}^\omega, \quad \forall \omega \in \Omega. \tag{Fsb}$$

Note that (Fsb) restricts the designer to displaying only one (winning) ad to every customer, which is without loss of generality given unit-demand customers (see Section 6 for model extensions). The designer assigns a transfer $t_i \in \mathbb{R}$ to each merchant i that accounts for merchant i's dataset contributions $(\alpha_i^\omega)_{\omega\in\Omega}$ *and* total targeted ads $x_i = (x_i^\omega)_{\omega\in\Omega}$. I refer to $(x,t) = (x_i,t_i)_{i\in\mathcal{N}}$ as a *targeted advertising campaign* (henceforth, campaign).

It is common practice for eCommerce platforms to run advertising campaigns for participating merchants by letting them choose an advertising budget (analogous to t_i) ahead of time and offer an (expected) quantity of targeted ads accordingly. Here, the model includes an additional layer of complexity: the designer needs to carefully choose which ad campaign to run so as to make sure each merchant is weakly better off than running individual targeted ads off-platform. As a result, transfers can also be negative when designer compensates merchants for lower on-platform targeted advertising performance.

2.2 Designer's Objective.

I consider a designer who designs campaigns with the objective of maximizing *total welfare*, defined as a weighted sum of platform revenue and social value. The latter is itself a weighted sum of customer engagement and merchants' surplus. Considering the interests of all stakeholders—the designer, customers and merchants—ensures long-term sustainability for the platform's ecosystem.

Platform Revenue. The platforms' revenue is the sum of transfers across all merchants' fees:

$$R(t) := \sum_{i\in\mathcal{N}} t_i. \tag{1}$$

Because building such platforms typically involves non-trivial costs, I examine optimal designs in which the designer may also generate profits. Because these costs are treated as sunk, I omit their analysis from the model.

Customer Engagement. Customer engagement is defined as the improved targeting accuracy—given by the net increase in product clicks across all participating merchants. Specifically, given target ads $x = (x_i)_{i\in\mathcal{N}}$, the net increase in the product clicks for every customer with CTR $\omega \in \Omega$ is given by the difference between expected on-platform clicks $\sum_{i\in\mathcal{N}} \omega_i x_i^\omega$ and off-platform clicks $\sum_{i\in\mathcal{N}} \omega_i \alpha_i^\omega$. Therefore, the total customer engagement is defined by

$$W^\alpha(x) := \sum_{\omega\in\Omega}\sum_{i\in\mathcal{N}} \omega_i(x_i^\omega - \alpha_i^\omega). \tag{2}$$

One can interpret $W(x,\alpha)$ as a measure of customer welfare. This is because the model does not consider product prices directly; instead, the welfare of a customer is represented by the net increase in engagement with ads provided by the platform. There are many reasons beyond the scope of the model that may incentivize the designer to care about customers—for

example, among others, poor targeting can lead to ad fatigue, whereas precise targeting can enhance user experience and encourage continued engagement with the platform.

Merchants' Surplus. Merchant surplus is defined as the difference in targeted ad performances across on and off-platform. Specifically, given targeted ads x, type θ_i's on-platform targeted ad performance is given by

$$v_i(\theta_i, x_i) = \theta_i \sum_{\omega \in \Omega} \omega_i x_i^\omega. \tag{3}$$

On the other hand, type θ_i's off-platform outside option involves running targeted ads independently by leveraging their own exclusive dataset. That is,

$$\underline{v}_i^\alpha(\theta_i) = \theta_i \sum_{\omega \in \Omega} \omega_i \alpha_i^\omega. \tag{4}$$

Therefore, type θ_i's (ex-post) *net* surplus from targeted ads x is given by

$$V_i^\alpha(\theta_i, x_i) := v_i(\theta_i, x_i) - \underline{v}_i^\alpha(\theta_i). \tag{5}$$

Denote by $V(\theta, x, \alpha)$ the *total* (ex-post) net surplus, i.e.

$$V^\alpha(\theta, x) := \sum_{i \in \mathcal{N}} V_i^\alpha(\theta_i, x_i). \tag{6}$$

The designer cares about merchants' surplus to incentivize participation and investing in the platform via data sharing, which benefits to platform sustainability and growth.

Now that the necessary terminology is in place, I formally state the designer's objective. Given dataset $\alpha \in \Delta(\Omega)$ and type profile $\theta \in \Theta$, the *ex-post* total welfare from campaign (x, t)—denoted by $\psi_\eta^\alpha(\theta, x, t)$—is given by

$$\psi_\eta^\alpha(\theta, x, t) := \eta_v V^\alpha(\theta, x) + \eta_w W^\alpha(x) + \eta_r R(t), \tag{Obj$_\theta$}$$

for some non-negative *welfare weight* $\eta = (\eta_v, \eta_w, \eta_r)$ such that $\eta_v + \eta_w + \eta_r = 1$ and $\eta_r > 0$. Note that, due to the imposed data sharing policy, the campaigns must meet a feasibility requirement: guaranteeing non-negative expected payoff to all merchants to ensure full participation. Adopting standard terminology, a campaign (x, t) is *individually rational* (IR) at type profile $\theta \in \Theta$ if it yields for every merchant nonnegative expected payoff $U_i^\alpha(\theta_i, x_i, t_i)$, defined by:

$$U_i^\alpha(\theta_i, x_i, t_i) := V_i^\alpha(\theta_i, x_i) - t_i \geq 0, \quad \forall i \in \mathcal{N}. \tag{IR$_\theta$}$$

Therefore, the designer's ideal objective is to design, at each type profile $\theta \in \Theta$, an IR campaign (x, t) that maximizes $\psi_\eta^\alpha(\theta, x, t)$. However, the designer faces *incentive compatibility* (IC) constraints: it requires that merchants truthfully report their private information (θ). I next employ tools from mechanism design theory to formally address these constraints.

2.3 Mechanism Design

To characterize the set of IR campaigns implementable under IC constraints, the designer's problem is reduced to choosing from the smaller space of IR and IC *direct revelation mechanisms.*

Direct Revelation Mechanisms. By the Revelation Principle, the designer can restrict, without loss of generality, to direct revelation mechanisms (henceforth, *mechanisms*) which are measurable maps given by *targeted ads rule* (henceforth, *ads rule*) $x = (x_i^\omega)_{i\in\mathcal{N},\omega\in\Omega}$ with $x_i^\omega(\cdot) : \Theta \to \mathbb{R}_+$ and *transfer rule* $t = (t_i)_{i\in\mathcal{N}}$ with $t(\cdot) : \Theta \to \mathbb{R}$ that satisfy *interim* IC and *interim* IR constraints

$$U_i^\alpha(\theta_i, X_i(\theta_i), T_i(\theta_i)) \geq U_i^\alpha(\theta_i, X_i(\tilde{\theta}_i), T_i(\tilde{\theta}_i)), \qquad \forall \theta_i, \tilde{\theta}_i \in \Theta_i, \tag{IC}$$

$$U_i^\alpha(\theta_i, X_i(\theta_i), T_i(\theta_i)) \geq 0, \qquad \forall \theta_i \in \Theta_i, \tag{IR}$$

and *ex-post* feasibility constraints

$$\sum_{i\in\mathcal{N}} x_i^\omega(\theta) \leq \alpha_{\mathcal{N}}^\omega, \quad \forall \omega \in \Omega, \forall \theta \in \Theta, \tag{Fsb$_\theta$}$$

where $X_i = (X_i^\omega)_{\omega\in\Omega}$ with $X_i^\omega(\theta_i) := \mathbb{E}_{\theta_{-i}}[x_i^\omega(\theta_i, \theta_{-i})]$ and $T_i(\theta_i) := \mathbb{E}_{\theta_{-i}}[t(\theta_i, \theta_{-i})]$ for every $\theta_i \in \Theta_i$ and $i \in \mathcal{N}$. I adopt the Bayes-Nash solution for the analysis of IC constraints, which is less restrictive and allows for richer and more flexible mechanisms. In many real-world mechanisms that platforms run, like standard auction formats, agents act strategically conditional on beliefs about others' private information. The more restrictive assumption is adopting interim IC's in the analysis: merchants are guaranteed a nonnegative expected payoff, but not necessarily ex-post.

Denote by $\mathcal{M}^\alpha(\Theta, dF)$ the space of feasible mechanisms. Note that $\mathcal{M}^\alpha(\Theta, dF) \neq \emptyset$ since the mechanism with $x_i^\omega(\cdot) \equiv \alpha_i^\omega$ for every $i \in \mathcal{N}$, $\omega \in \Omega$, and $t(\cdot) \equiv 0$ (henceforth, $(x(\cdot), t(\cdot)) \equiv (\alpha, 0)$) satisfies (IC), (IR) and (Fsb$_\theta$).

Designer's Problem. I consider a designer who maximizes the *expected* total welfare. Letting $\mathcal{W}^\alpha(x) := \mathbb{E}[W^\alpha(x(\theta))]$, $\mathcal{R}(t) := \mathbb{E}[R(t(\theta))]$ and $\mathcal{V}^\alpha(x) := \mathbb{E}[V^\alpha(\theta, x(\theta))]$, the expected value of objective function from mechanism (x, t)—denoted by $\Psi_\eta^\alpha(x, t)$—is given by

$$\Psi_\eta^\alpha(x, t) := \eta_v \mathcal{V}^\alpha(x) + \eta_w \mathcal{W}^\alpha(x) + \eta_r \mathcal{R}(t). \tag{Obj}$$

Therefore, the designer's problem (P) is to maximize the expected total welfare in (Obj) over the space of feasible mechanisms in $\mathcal{M}^\alpha(\Theta, dF)$:

$$\Pi(\alpha, \eta) := \sup_{(x,t)\in\mathcal{M}^\alpha(\Theta,dF)} \Psi_\eta^\alpha(x, t). \tag{P}$$

I refer to $\Pi(\alpha, \eta)$ as the *value* of designing a platform with welfare weight η and available dataset α. Note that the value $0 \leq \Pi(\alpha, \eta) < \infty$ since the elements of $\mathcal{M}^\alpha(\Theta, dF)$ are bounded measurable functions $(x, t \in \mathcal{L}^\infty(\Theta, dF))$ and $(\alpha, 0) \in \mathcal{M}^\alpha(\Theta, dF)$. Lastly, I assume the designer has *full commitment power*; see Section 6 for a discussion of the *limited* commitment case.

3 Illustrative Examples

I illustrate the optimal mechanism for two merchants and a revenue-maximizing (i.e. $\eta_r = 1$) platform in settings for which the optimal mechanism is known from the literature, such as *monopoly pricing*, *bilateral trading* and *partnership dissolution*. I conclude the section by presenting an novel example that departs from traditional models and will build intuition for the general characterization of the optimal mechanism in the next section.

Let $i = 1, 2$ and $\theta_i \overset{iid}{\sim} \text{Uniform}[0, 1]$, and welfare weight $\eta_r = 1$. Suppose there are only two types of customers: *a*) *exclusive customers* who have CTR $\omega_i = 1$ for some merchant $i = 1, 2$ and have no intent in purchasing from the other merchant $j \neq i$ (i.e. $\omega_j = 0$); *b*) *inclusive customers* who see both merchants as perfect substitutes and have CTRs $\omega = (1, 1)$.

Monopoly Pricing. Suppose that all customers are exclusive with CTR $\omega_1 = 1$ and all are initially part of merchant 2's dataset, i.e. $\alpha_2^{(1,0)} = 1$. Then clearly, the optimal mechanism is a classical monopoly pricing problem: the designer procures the dataset at no cost and sells the unit mass of goods (here, targeted ads) to merchant 1 according to a posted price $p^* = 1/2$—so there's trade if and only if $\theta_1 \geq 1/2$.

Bilateral Trade and Partnership Dissolution. Suppose that all customers are inclusive and consider first the case in which merchant 2 owns all the inclusive dataset, i.e. $\alpha_2^{(1,1)} = 1$. Then the problem becomes analogous to the profit-maximizing bilateral trading model via a mediator where merchant 1 is the *buyer* (B) and merchant 2 is the *seller* (S) as in Myerson and Satterthwaite (1983). They show (Theorem 4) that the optimal mechanism assigns scores equal to the virtual value $\phi_1^B(\theta_1) = 2\theta_1 - 1$ and virtual cost $\phi_2^S(\theta_2) = 2\theta_2$ and trade happens if and only if $\phi^B(\theta_1) \geq \phi^S(\theta_2)$—or $\theta_1 - \theta_2 \geq 1/2$.

Now consider the case in which merchant $i = 1, 2$ holds an ownership share of $r_i \in (0, 1)$ in the inclusive dataset, i.e. $r_i := \alpha_i^{(1,1)}$ such that $r_1 + r_2 = 1$. The problem becomes analogous to dissolving partnerships (Cramton et al. (1987)) at a profit via a mediator (Loertscher and Wasser (2019)). Similarly, the optimal mechanism assigns full ownership the the merchant with the highest score $g_i(\theta_i)$, where g_i is derived by ironing the virtual type $\phi_i(\theta_i, \theta_i')$ at some optimal critical *worst-off* type θ_i'—specifically, there exists some $z = (z_1(\theta_1'), z_2(\theta_2'))$ such that $g_i(\theta_i) = \phi_i^S(\theta_i)$ if $\phi_i^S(\theta_i) < z_i$, $g_i(\theta_i) = \phi_i^B(\theta_i)$ if $\phi_i^B(\theta_i) > z_i$ and $g_i(\theta_i) = z_i$ otherwise. For instance, when $r_i = 1/2$ then $z_i = 1/2$ implying trade will occur whenever $\theta_i > \theta_j$ unless $\theta_i \in [1/4, 3/4]$ (or $g_i(\theta_i) = 1/2$) for each $i = 1, 2$. In that case, ties are broken uniformly. Note that the trade volume (and revenues) increase compared to bilateral trade (i.e. $r_i = 1$). This is due to the presence of countervailing incentives that weaken incentive constraints: a high (low) type is more likely to buy (sell) and thus has an incentive to under(over)-report, whereas the types in the "middle" have the weakest incentives to misreport since they retain their shares in expectation by selling/buying "half" of the time.

Departing from Traditional Models. Suppose now that at least one merchant owns data about both types of customers. In particular, let $\alpha_{\mathcal{N}}^{(1,1)} \in (0, 1)$ be equally distributed across both merchants as above—$\alpha_1^{(1,1)} = \alpha_2^{(1,1)} = \alpha_{\mathcal{N}}^{(1,1)}/2$. Let the remaining $1 - \alpha_{\mathcal{N}}^{(1,1)}$ mass of customers be exclusive towards merchant 2 but which are initially owned by merchant 1, i.e. $\alpha_1^{(0,1)} = 1 - \alpha_{\mathcal{N}}^{(1,1)}$. Here the novel element is the heterogeneity of the goods to be traded. In particular, merchant 2 can be both a seller *and* a buyer: in the optimal mechanism merchant 2 sometimes sells their $\alpha_{\mathcal{N}}^{(1,1)}/2$ mass of inclusive customer but is always a potential buyer for the $1 - \alpha_{\mathcal{N}}^{(1,1)}$ mass of exclusive customers. A naive strategy for the designer is to design separate mechanisms to assign targeted ads for inclusive and exclusive customers above; see Figure 1 where $z^E = 0$ and $z^I = 1/2$ identify the tie-breaking regions for optimal ad design

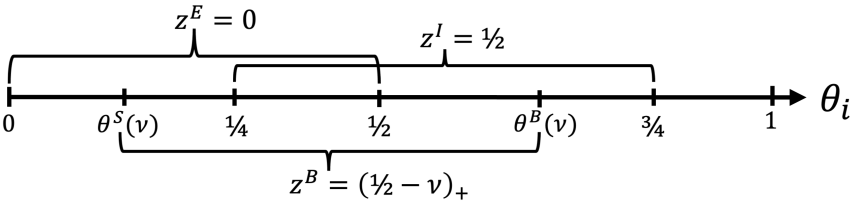


Figure 1: Optimal ad design separately (z^E and z^I) and in bundles (z^B).

of exclusive and inclusive datasets, respectively. Moreover, the optimal expected transfers T_i^E and T_i^I can be computed as following: $T_1^E \equiv 0$ and $T_2^E(\theta_2) = (1 - \alpha_{\mathcal{N}}^{(1,1)}) \times (1/2)$, and for each $i = 1, 2$

$$T_i^I = \begin{cases} \frac{\alpha_{\mathcal{N}}^{(1,1)}}{2}\left(\theta_i^2 - \frac{1}{4}\left(1 - \frac{1}{4}\right)\right), & \text{if } \theta_i \in \left[0, \frac{1}{4}\right] \\ \frac{\alpha_{\mathcal{N}}^{(1,1)}}{2}\left(\theta_i^2 - \frac{3}{4}\left(1 - \frac{3}{4}\right)\right), & \text{if } \theta_i \in \left[\frac{3}{4}, 1\right] \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

Now, I show that the designer can in fact do *strictly* better by designing targeted ads in *bundles*. Specifically, the optimal mechanism assigns scores g_i that irons the virtual type $\phi_i(\theta_i, \theta_i')$ with ironing parameter $z^B = (1/2 - \nu)_+$ where $\nu := (1 - \alpha_{\mathcal{N}}^{(1,1)})/\alpha_{\mathcal{N}}^{(1,1)}$ and $z^B > 0$ if and only if $\alpha_{\mathcal{N}}^{(1,1)} > 2/3$. The corresponding tie-breaking region is thus given by

$$\left[\theta^S(\nu), \theta^B(\nu)\right] = \left[\left(\frac{1}{4} - \frac{\nu}{2}\right)_+, \max\left(\frac{1}{2}, \frac{3}{4} - \frac{\nu}{2}\right)\right]. \tag{8}$$

The optimal design for targeted ad bundles is as follows: merchant 2 is always assigned the exclusive targeted ads whenever $g_2(\theta_i) > 0$—which holds a.e. if $z^B > 0$. On the other hand, the inclusive targeted ads are assigned to the merchant with the highest score $g_i(\theta_i)$ where ties are broken in favor of merchant 1 according to $p_1^{(1,1)} = \min(1, 1/2 + \nu)$. Moreover, the optimal *non-itemized* expected transfers from targeted ad bundles T_i^B can be written as a sum of the following *hypothetical* itemized expected transfers: $T_i^B \equiv \tilde{T}_i^E + \tilde{T}_i^I$ where $\tilde{T}_1^E \equiv 0$ and $\tilde{T}_2^E(\theta_2) = (1 - \alpha_{\mathcal{N}}^{(1,1)})\theta^S(\nu)$ for $\theta_2 < \theta^S(\nu)$, $\tilde{T}_2^E(\theta_2) = (1 - \alpha_{\mathcal{N}}^{(1,1)})\theta^B(\nu) > 0$ for $\theta_2 > \theta^B(\nu)$ and zero otherwise. On the other hand, for each $i = 1, 2$

$$\tilde{T}_i^I = \begin{cases} \frac{\alpha_{\mathcal{N}}^{(1,1)}}{2}\Big(\theta_i^2 - (1 - \theta^S(\nu)(1 - \theta^S(\theta)))\Big), & \text{if } \theta_i \in [0, \theta^S(\nu)] \\ \frac{\alpha_{\mathcal{N}}^{(1,1)}}{2}\Big(\theta_i^2 - (1 - \theta^B(\nu)(1 - \theta^B(\theta)))\Big), & \text{if } \theta_i \in [\theta^B(\nu), 1] \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

Note that if $\theta_2 < \theta^S(\nu)$, although $\tilde{T}_2(\theta_2) < 0$ overall, the itemized transfers satisfy $T_2^I(\theta_2) < \tilde{T}_2^I(\theta_2) < 0$ and $\tilde{T}_2^E(\theta_2) = (1 - \alpha_{\mathcal{N}}^{(1,1)})\theta^S(\nu) > 0$—a strictly positive per-unit posted price $p^* = \theta^S(\nu)$ that is higher than merchant 2's willingness to pay θ_2. Intuitively, since merchant 2 has a low value, she expects to sell the inclusive dataset and buy the exclusive ones, where the incentives to over-report (from selling) dominates the incentives to under-report (from buying). The designer thus exploits such mixed incentives by offering her a bundle of ads at a non-itemized price. Thus, if $\theta_i < 1/4$ then $\tilde{T}_i(\theta_i) > T_i^E(\theta_i) + T_i^I(\theta_i) = T_i^I(\theta_i)$ since $\tilde{T}_i(\theta_i) = 0 > T_i^I(\theta_i)$ for any $\theta_i \in (\theta^S(\nu), 1/4)$. Hence, on $[0, 1/4]$ the designer has strictly

higher revenues. One can do the same analysis for $\theta_i > 1/2$ and find that for merchant $i = 1$ the designer still assigns zero transfers in $[1/4, 1/2]$, does better in $[1/2, \theta^B(\nu)]$ by selling inclusive ads but worse in $[\theta^B(\nu), 3/4]$ by selling them at a lower price. On the other hand, for merchant $i = 2$ the designer does worse in region $[1/4, 1/2]$ by not selling any exclusive ads, but does better in $[1/2, 3/4]$ altogether by selling ads in bundles rather than separately. Overall, it can be shown that the designer achieves higher revenues from assigning ads in bundles. I provide full characterization of the optimal mechanism for arbitrary distribution F, exclusive/inclusive datasets α and welfare weight η in Section 4.2 and generalize these insights.

4 Optimal Mechanism

I show the existence of an optimal mechanism in which targeted ads are allocated according to positive and non-decreasing *scoring rules* $g_i : \Theta_i \to \mathbb{R}_+$ for every $i \in \mathcal{N}$. The mechanism targets customers $\omega \in \Omega$ to merchant(s) $i \in \mathcal{N}$ with the highest *weighted* score $\omega_i g_i(\theta_i)$. Ties are broken according to some *CTR-specific* tie-breaking rules $p^\omega : \Theta \to \Delta(\mathcal{N}_0)$ such that participation constraints of worst-off types bind.[7] Such mechanisms are formally defined as follows:

Definition 1 (*Scoring Mechanism*). *A scoring mechanism $(g, p) = ((g)_{i\in\mathcal{N}}, (p^\omega)_{\omega\in\Omega})$ is characterized by positive and non-decreasing scoring rules $g_i : \Theta_i \to \mathbb{R}_+$ for every $i \in \mathcal{N}$ and CTR-specific tie-breaking rules $p^\omega : \Theta \to \Delta(\mathcal{N}_0)$ for every $\omega \in \Omega$ such that $\forall\theta \in \Theta, \forall\omega \in \Omega, \forall i \in \mathcal{N}$:*

$$x_i^\omega(\theta) = p_i^\omega(\theta)\alpha_{\mathcal{N}}^\omega \quad \text{and} \quad p_i^\omega(\theta) = 0 \text{ if } i \notin \mathcal{N}^\omega(\theta) := \operatorname*{argmax}_{j\in\mathcal{N}_0}\{\omega_j g_j(\theta)\}, \tag{10}$$

where for simplicity of exposition let $\omega_0 g_0 \equiv 0$ denote the score assigned to designer.

To state the main theorem, I first introduce some definitions and a lemma.

Definition 2. *For each $k \in \{B, S\}$ define the weighted virtual value ($k = B$) and cost ($k = S$) functions $\phi_{\eta,i}^k$ by $\phi_{\eta,i}^k(\theta_i) := \eta_\omega + \eta_v\theta_i + \eta_r\phi_i^k(\theta_i)$ for every $\theta_i \in \Theta_i$ and $i \in \mathcal{N}$. Then, let $\phi_{\eta,i}(\cdot, \theta_i')$ be the weighted virtual type function at the critical type $\theta_i' \in \Theta_i$ be defined by $\phi_{\eta,i}(\theta_i, \theta_i') = \phi_{\eta,i}^S(\theta_i)$ if $\theta_i < \theta_i'$ and $\phi_{\eta,i}(\theta_i, \theta_i') = \phi_{\eta,i}^B(\theta_i)$ if $\theta_i \geq \theta_i'$, for every $\theta_i \in \Theta_i$ and $i \in \mathcal{N}$. Lastly, define types $\theta_{\eta,i}^k(y) := (\phi_{\eta,i}^k)^{-1}(y)$ for every $y \in [\phi_{\eta,i}^k(0), \phi_{\eta,i}^k(1)]$.[8] See Figure 2 for a visual description.*

Definition 3 (*Worst-Off Types*). *Given mechanism (x, t), for every $i \in \mathcal{N}$ let $\hat{\Theta}_i^\alpha(x, t)$ denote the set of worst-off types, i.e.*

$$\hat{\Theta}_i^\alpha(x, t) := \operatorname*{argmin}_{\theta_i'\in\Theta_i} U_i^\alpha(\theta_i', X_i(\theta_i'), T_i(\theta_i')). \tag{11}$$

[7] Denote by $\mathcal{N}_0 = \mathcal{N} \cup \{0\}$ where p_0^ω represents the share of ω customers that are not targeted by any merchant. Note that one can show that there always exists a mechanism in which every customer is targeted by some merchants, i.e. $p^\omega(\theta) \in \Delta(\mathcal{N})$. However, in some applications (see, for example, Section 4.2) we focus on a practical class of ads rules for which some customers do not get targeted by any merchant.

[8] Note that $\theta_{\eta,i}^S(y) < \theta_{\eta,i}^B(y)$ for every $y \in [\phi_{\eta,i}^S(0), \phi_{\eta,i}^B(1)]$ follows from regularity of F and $\eta_r > 0$.

Let $\hat{\Theta}^\alpha(x,t) = \times_{i\in\mathcal{N}}\hat{\Theta}_i^\alpha(x,t)$ *denote the set of worst-off types in* Θ *induced by mechanism* (x,t).

Lemma 1 (*Characterizing Worst-Off Types à la Cramton et al. (1987)*)**.** *Given* (IC) *and* (IR) *mechanism* (x,t)*, the set of worst-off types is characterized in terms of ads rules* x *only, i.e.* $\hat{\Theta}_i^\alpha(x,t) = \hat{\Theta}_i^\alpha(x)$ *where*

$$\hat{\Theta}_i^\alpha(x) = \begin{cases} \{\theta_i \in \Theta_i : S_i(\theta_i) = a_i\}, & \text{if } \exists \theta_i : S_i(\theta_i) = a_i, \\ \{\theta_i \in \Theta_i : S_i(y) < a_i, \forall y < \theta_i \text{ and } S_i(y) > a_i, \forall y > \theta_i\}, & o/w, \end{cases} \tag{12}$$

where $S_i(\theta_i) := \sum_{\omega\in\Omega}\omega_i X_i^\omega(\theta_i)$ *and* $a_i := \sum_{\omega\in\Omega}\omega_i\alpha_i^\omega$ *denote the (interim) expected clicks (henceforth, clicks) on and off-platform, respectively.*

We are now ready to state our main result showing the existence and characterization of an optimal mechanism:

Theorem 1 (*Optimal Mechanism*)**.** *For any arbitrarily chosen welfare weight* η *and dataset* α*, there exists a scoring mechanism* (g,p) *that attains the value* $\Pi(\alpha,\eta)$ *in* (P)*. In particular, for every* $i \in \mathcal{N}$ *there exists* $z_i \in \left[\underline{z}_{\eta,i}, \overline{z}_{\eta,i}\right] = \left[\max\{0, \phi_{\eta,i}^B(0)\}, \phi_{\eta,i}^S(1)\right]$ *such that*

$$g_i(\theta_i) = \phi_{\eta,i}(\theta_i, z_i) := \begin{cases} \phi_{\eta,i}^S(\theta_i), & \text{if} \quad \theta_i < \theta_i^S(z_i), \\ z_i, & \text{if} \quad \theta_i^S(z_i) \le \theta_i \le \theta_i^B(z_i), \\ \phi_{\eta,i}^B(\theta_i), & \text{if} \quad \theta_i > \theta_i^B(z_i), \end{cases} \tag{13}$$

where the tie breaking rule p *is such that the critical type* $\hat{\theta}_i(z_i)$*—uniquely defined by* $\hat{\theta}_i = \mathbb{E}\left[\phi_{\eta,i}(\theta_i, z_i) - \eta_w - \eta_v\theta_i\right]/\eta_r$*—lies in* $\hat{\Theta}_i^\alpha(x)$*. The interim expected transfers satisfy*

$$T_i(\theta_i) = \theta_i(S_i(\theta_i) - a_i) - \int_{\hat{\theta}_i(z_i)}^{\theta_i}(S_i(y) - a_i)dy, \quad \forall \theta_i \in \Theta_i \text{ and } i \in \mathcal{N}. \tag{14}$$

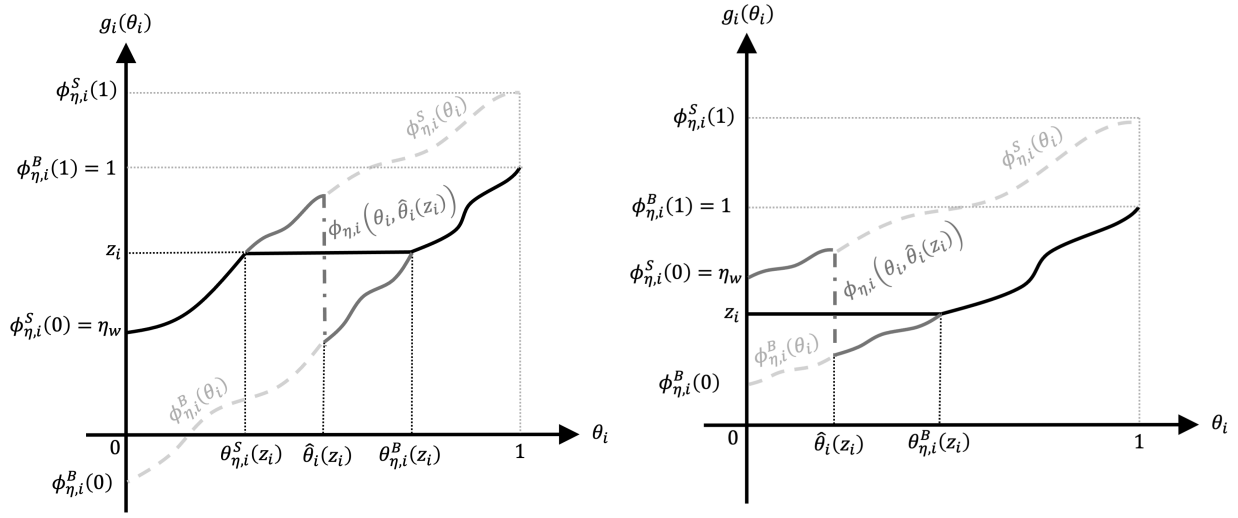
The main insights from the theorem is that a single scoring function g_i is used to allocated each targeted ad by taking the weighted score $\omega_i g_i$. This reduces information rents by bundling targeted ads and data sharing. Moreover, the mechanism is not efficient due to the trades at the ironing region $\left[\theta_i^S, \theta_i^B\right]$. Lastly, if the weight on customer engagement η_ω is sufficiently large (i.e. $\underline{z} > 0$) then all the data is reallocated to the merchants with highest CTR-weighted score.

Next, I present the main steps in proving Theorem 1 and leave the remaining details to Appendix B. We will see that although only one dimensional monotonicity of S_i's matter for (IC), the multidimensionality of tie-breaking rules p^ω matter for (IR).

4.1 Proof Sketch

The first steps consist of eliminating (IC) and (IR) constraints and the dependence on transfer rule t from (P). I then proceed to show the existence of an optimal mechanism and characterization of scoring mechanisms.

Eliminating (IC). Using standard Myersonian approach to characterize incentive compatibility, I reduce the designer's problem in (P) to optimizing over the space of ads rules



(a) $\phi^B_{\eta,i}(0) < 0$. (b) $\phi^B_{\eta,i}(0) \geq 0$.

Figure 2: Graphical representation of (optimal) scoring rules and (weighted) virtual types.

$x \in \mathcal{X}^\alpha(\Theta, dF)$ subject to suitable monotonicity constraints and transfers $t(\hat{\theta})$ at the worst-off types $\hat{\theta}$.

Specifically, let $U_i(\theta_i)$ denote the (interim) expected payoff of type $\theta_i \in \Theta_i$ that can be attained under some mechanism $(x,t) \in \mathcal{M}^\alpha(\Theta, dF)$, i.e.

$$U_i(\theta_i) := \theta_i(S_i(\theta_i) - a_i) - T_i(\theta_i). \tag{15}$$

Then, by standard characterization of (interim) incentive compatibility we know a mechanism (x,t) satisfies (IC) if and only if

$$U_i(\theta_i) = U_i(\theta'_i) + \int_{\theta'_i}^{\theta_i} (S_i(y) - a_i)dy, \quad \forall \theta_i, \theta'_i \in \Theta_i, \tag{ICFOC}$$

$$S_i(\cdot) \text{ is nondecreasing.} \tag{M}$$

Now, for every $i \in \mathcal{N}$ fix an arbitrary $\theta'_i \in \Theta_i$. Using (ICFOC), the (interim) expected transfers $T_i(\theta_i)$ at every $\theta_i \in \Theta_i$ are pinned down in terms of clicks rule S_i and payoff $U_i(\theta'_i)$:

$$T_i(\theta_i) = \theta_i(S_i(\theta_i) - a_i) - \int_{\theta'_i}^{\theta_i} (S_i(y) - a_i)dy - U_i(\theta'_i), \tag{16}$$

where note that $U_i(\theta'_i)$ also depends on transfers $T_i(\theta'_i)$.

Next, fix $(x,t) \in \mathcal{M}^\alpha(\Theta, dF)$ that satisfies (IC)—hence (ICFOC) and (M). By standard integration by parts techniques (see Appendix B for details), using (16) simplifies (Obj) as follows:

$$\Psi^\alpha_\eta(x,t) = \Psi^\alpha_\eta(x,\theta') - \eta_r \sum_{i \in \mathcal{N}} U_i(\theta'_i), \quad \forall \theta' \in \Theta, \tag{17}$$

where

$$\Psi_\eta^\alpha(x,\theta') = \mathbb{E}\left[\sum_{i\in\mathcal{N}}\left(\sum_{\omega\in\Omega}\omega_i x_i^\omega(\theta) - a_i\right)\phi_{\eta,i}(\theta_i,\theta_i')\right]. \tag{18}$$

In standard mechanism design problems, the worst-off type $\hat{\theta}$ is typically predetermined—often taken to be either $\hat{\theta} = \underline{\theta}$ or $\hat{\theta} = \overline{\theta}$—*independently* of the optimal allocation rule x^*. By optimality, such types receive payoff $U_i(\hat{\theta}_i) = 0$ for every $i \in \mathcal{N}$. In those cases, one sets $\theta' = \hat{\theta}$, allowing (16) (hence (18) too) to be written solely in terms of the ads rule x. However, in settings with countervailing incentives—like the one considered here—the worst-off type(s) and allocation rule x are *interdependent*, which complicates the analysis. I next show how one can circumvent such complications.

Eliminating (IR) *and Existence.* I show that the optimization problem in (P) can be written as a *minimax* problem, which I then use to i) show existence of an optimal mechanism, and ii) solve explicitly for the ads rule x^* *independently* of the term $U_i'(\theta')$ in (17). The details are left to Appendix B.

Following the approach by Loertscher and Wasser (2019), the first step is characterizing the set of worst-off types by

$$\hat{\Theta}(x) = \operatorname*{argmin}_{\theta'\in\Theta} \Psi_\eta^\alpha(x,\theta'), \tag{19}$$

which then allows us to write the designer's problem as

$$\Pi(\alpha,\eta) = \max_{x\in\mathcal{X}^\alpha(\Theta,dF)} \min_{\theta'\in\Theta} \Psi_\eta^\alpha(x,\theta'), \tag{Max-Min}$$

where $\mathcal{X}^\alpha(\Theta,dF)$ denotes the space of measurable ads rules x that satisfy (Fsb_θ) and (M), which is nonempty since $\alpha \in \mathcal{X}^\alpha(\Theta,dF)$. Then, a solution to the (Max-Min) problem exists by a generalized version of von Neumann's minimax theorem (von Neumann (1928)). Moreover, the theorem implies that one can swap the order in (Max-Min), resulting in the following minimax problem:

$$\Pi(\alpha,\eta) = \min_{\theta'\in\Theta} \max_{x\in\mathcal{X}^\alpha(\Theta,dF)} \Psi_\eta^\alpha(x,\theta'), \tag{20}$$

ensuring that a saddle point $(x^*,\hat{\theta}) \in \mathcal{X}^\alpha(\Theta,dF)\times\Theta$ exists:

$$\max_{x\in\mathcal{X}^\alpha(\Theta,dF)} \Psi_\eta^\alpha(x,\hat{\theta}) = \Psi_\eta^\alpha(x^*,\hat{\theta}) = \min_{\theta'\in\Theta} \Psi_\eta^\alpha(x^*,\theta'), \tag{21}$$

where $\Pi(\alpha,\eta) = \Psi_\eta^\alpha(x^*,\hat{\theta})$. As a result, one can now characterize saddle points $(x^*,\hat{\theta})$ (hence the optimal mechanism) by solving each of the optimization problems in (21) separately, as I proceed to do next.

Ironing. Given saddle point $(x^*,\hat{\theta})$, from (12) we already have a characterization for the worst-off types in terms of the ads rule, i.e. $\hat{\theta} \in \hat{\Theta}(x^*)$. So, it remains to characterize the

optimal ads rules x^* that solve

$$\max_{x\in\mathcal{X}^\alpha(\Theta,dF)} \mathbb{E}\left[\sum_{i\in\mathcal{N}}\left(\sum_{\omega\in\Omega}\omega_i x_i^\omega(\theta)-a_i\right)\phi_{\eta,i}(\theta_i,\hat{\theta}_i)\right] \tag{22}$$

$$\text{subject to}\quad \hat{\theta}\in\hat{\Theta}(x). \tag{23}$$

Now, note that pointwise maximization of (22) does not necessarily result in non-decreasing $X_i^{*\omega}(\theta_i)$—hence non-decreasing $S_i^*(\theta_i)$—since $\phi_{\eta,i}(\cdot,\hat{\theta}_i)$ is non-monotone at $\theta_i=\hat{\theta}_i$. To circumvent this, the solution uses ironing techniques from Myerson (1981). Ignoring for now constraint (23), for each $i\in\mathcal{N}$ let $z_i(\hat{\theta}_i)$ be the unique ironing parameter that satisfies

$$\mathbb{E}[\phi_{\eta,i}(\theta_i;z_i)]=\mathbb{E}[\phi_{\eta,i}(\theta_i,\hat{\theta}_i)]=\eta_w+\eta_v\mathbb{E}[\theta_i]+\eta_r\hat{\theta}_i. \tag{24}$$

Then, we know that (22) is equivalent to solving

$$\max_{\substack{x\ \text{pointwise}\\ \text{s.t. (Fsb}_\theta\text{)}}} \mathbb{E}\left[\sum_{i\in\mathcal{N}}\left(\sum_{\omega\in\Omega}\omega_i x_i^\omega(\theta)-a_i\right)\phi_{\eta,i}(\theta_i,z_i(\hat{\theta}_i))\right]. \tag{25}$$

That is, since scoring rule $\phi_{\eta,i}(\cdot,z_i(\hat{\theta}_i))$ is non-decreasing, pointwise maximization of (25) yields non-decreasing $S_i^*(\theta_i)$—thus satisfying (M).

The subtle part is that often there does not exist a solution to (25) that satisfies (23) for an arbitrarily chosen $\hat{\theta}\in\hat{\Theta}(x^*)$. I show in Appendix B that there exists an *essentially unique* critical worst-off type $\theta^*\in\Theta(x^*)$ such that when ironing with respect to θ^*, (25) admits a solution that satisfies (23).

Scoring Mechanisms. I now proceed to describe pointwise maximizing ads rules x from (25) along with the corresponding tie-breaking rule.

Specifically, for every $\omega\in\Omega$ and $\theta\in\Theta$, the ads rule $x^{*\omega}(\theta)$ maximizes

$$x^{*\omega}(\theta)\in\operatorname*{argmax}_{x^\omega(\theta)}\sum_{i\in\mathcal{N}}x_i^\omega(\theta)\omega_i\phi_{\eta,i}(\theta_i,z_i(\hat{\theta}_i)),\quad\text{subject to}\ \sum_{i\in\mathcal{N}}x_i^\omega(\theta)\le\alpha_{\mathcal{N}}^\omega. \tag{26}$$

As it's readily seen, the solutions are given by (10). In case of ties, the designer chooses an CTR-specific tie-breaking rule p such that $\hat{\theta}\in\hat{\Theta}(x^*)$. The expected transfers follow by evaluating (16) at $\theta'=\hat{\theta}$.

4.2 A Stylized Setting: Exclusive & Inclusive Customers

I fully characterize the optimal scoring mechanism for $i=1,2$ and symmetric distributions $F_i\equiv F$ by restricting CTRs $\omega_i\in\{0,1\}$. I perform comparative statics on the model parameters and describe the insights it offers.

Specifically, a customer is said to be *exclusive* if they intend to purchase solely from some merchant $i=1,2$, and *inclusive* if they are indifferent among every merchant $i=1,2$. Formally, let $\Omega_1=\Omega_2=\{0,1\}$ where customers with CTR $\omega=(1,0)$ and $\omega=(0,1)$ are exclusive to merchant 1 and 2 respectively, whereas CTR $\omega=(1,1)$ represent inclusive customers to both merchants. For simplicity of exposition let $\omega_1:=(1,0)$ and $\omega_2:=(0,1)$.

Without loss of generality let $\alpha^{(0,0)} = 0$ and $\alpha_i^{\omega_i} = 0$ for each $i = 1, 2$.[9] Let $\mathcal{A}_{0,1} \subset \Delta(\{0,1\}^2)$ denote the set of all feasible inclusive and exclusive datasets.

Note that for any scoring rule with $z_i(\hat{\theta}_i) > 0$ pointwise maximization in (26) implies that all exclusive customers receives targeted ads from their preferred merchant, and in particular for type $\hat{\theta}_i$, i.e. $x^{*\omega_i}_i(\hat{\theta}_i, \theta_j) = \alpha_j^{\omega_i}$ for every $\theta_j \in \Theta_j$. However, if $z_i(\hat{\theta}_i) = 0$ then there can be multiplicity regarding the choice of the optimal tie-breaking rules between merchant i, merchant j (since $\omega_j = 0$) and the designer. In constructing the optimal mechanism, I restrict to scoring mechanisms that prioritize assigning exclusive targeted ads:

Definition 4 (*Exclusive-Priority Scoring Mechanisms*). *For given $\alpha \in \mathcal{A}_{0,1}$ a scoring mechanism (g, p) of the form $g_i \equiv \phi_{\eta,i}(\cdot, z_i(\hat{\theta}_i))$ for some $\hat{\theta} \in \Theta$ is said to be exclusive-priority (EP) if whenever there exists $i \in \mathcal{N}$ such that $z_i(\hat{\theta}_i) = 0$, ties are broken according to $p_i^{\omega_i}(\hat{\theta}_i, \theta_j) = \min\{\alpha_j^{\omega_i}, \alpha_i^{(1,1)}\}/\alpha_j^{\omega_i}$ for every $\theta_j \in \Theta_j$.*

EP scoring mechanisms prioritize targeting exclusive customers even when there's are remaining inclusive customers that can be targeted at no additional cost. I show that for an arbitrarily chosen dataset $\alpha \in \mathcal{A}_{0,1}$, welfare weight η and regular distribution F there exists an (essentially unique) EP scoring mechanism and provide conditions to compute (g, p) explicitly.

Before stating the result, I first introduce some additional terminologies. In what follows let $i \neq j \in \{1, 2\}$ and $\beta_i := \max\{0, (\alpha_i^{(1,1)} - \alpha_j^{\omega_i})/\alpha_{\mathcal{N}}^{(1,1)}\}$ denote merchant i's *adjusted* shares from $\alpha_{\mathcal{N}}^{(1,1)}$ mass of inclusive customers. Note that $\alpha_j^{\omega_i}$ mass of exclusive customers are initially 'dormant' as they never purchase from merchant j; the designer may then costlessly offer merchant i ads that target these customers. Therefore, of the total $\alpha_{\mathcal{N}}^{(1,1)}$ mass of inclusive customers, the designer retains residual ownership over a fraction $1 - \beta_1 - \beta_2$.[10] Next, define $P^S_{\eta,j}(z_i) := \mathbb{P}_F\left(\phi^S_{\eta,j}(\theta_j) \leq z_i\right)$ and $P^B_{\eta,j}(z_i) := \mathbb{P}_F\left(\phi^B_{\eta,j}(\theta_j) \leq z_i\right)$. If the optimal scoring rule is such that $z_i \leq z_j$ then the probability that the critical worst-off type $\hat{\theta}_i(z_i)$ ($\hat{\theta}_j(z_j)$) has a *strictly higher* score than merchant j (i) is given by $P^S_{\eta,j}(z_i)$ ($P^B_{\eta,i}(z_j)$). If $z_i = z_j$, then the probability that the critical worst-off type $\hat{\theta}_i(z_i)$ has equal score to merchant j is given by $P^B_{\eta,j}(z_i) - P^S_{\eta,j}(z_i)$.

Without loss of generality, I state the results for $\beta_1 \geq \beta_2$ since $\beta_1 \leq \beta_2$ is symmetric:

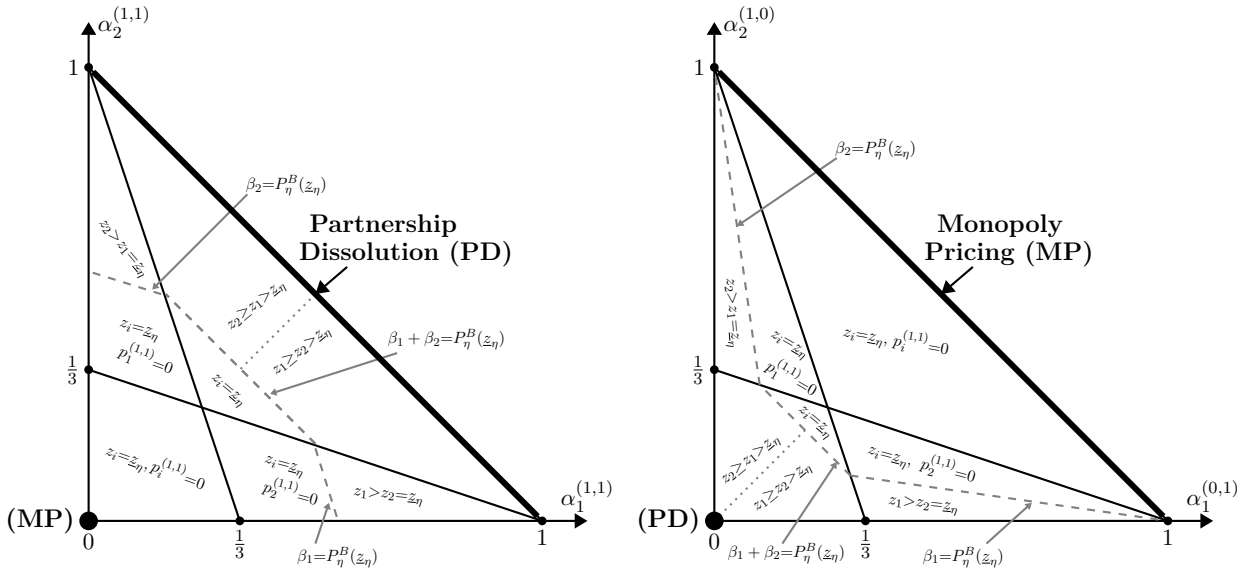
Proposition 1. *For any arbitrarily chosen $\alpha \in \mathcal{A}_{0,1}$ such that $\beta_1 \geq \beta_2$ and welfare weight η there exists (essentially unique) EP-scoring mechanism (g, p) such that:*

i) if $\beta_1, \beta_2 > 0$ and $\beta_1 + \beta_2 > P^B_\eta(\underline{z}_\eta)$, letting $\tilde{z} > \underline{z}_\eta$ be uniquely defined by $P^B_\eta(\tilde{z}) + P^S_\eta(\tilde{z}) = \beta_1 + \beta_2$, suppose first that $P^B_\eta(z) - P^S_\eta(z) \geq \beta_1 - \beta_2$. Then $z_1 = z_2 = \tilde{z}$ and the inclusive tie-breaking rule $p^{(1,1)}$ is uniquely given by

$$p_1^{(1,1)} = \frac{1}{2}\left(1 + \frac{\beta_1 - \beta_2}{P^B_\eta(z) - P^S_\eta(z)}\right). \tag{27}$$

[9] Customers $\omega = (0, 0)$ are irrelevant, and there's no gain from reallocating exclusive customer that are already part of their preferred merchants' dataset.

[10] We have $\beta_1 + \beta_2 \leq \left(\alpha_1^{(1,1)} + \alpha_2^{(1,1)}\right)/\alpha_{\mathcal{N}}^{(1,1)} = 1$.



(a) Symmetric Exclusive Lists: $\alpha_1^{(0,1)}=\alpha_2^{(1,0)}$ (b) Symmetric Inclusive Lists: $\alpha_1^{(1,1)}=\alpha_2^{(1,1)}$

Figure 3: Exclusive & Inclusive Customers: Optimal EP Scoring Mechanisms

On the other hand, if $P_\eta^B(z) - P_\eta^S(z) < \beta_1 - \beta_2$, *then there exists* $z_1 \geq \tilde{z} \geq z_2 (> \underline{z}_\eta)$, *with at least one inequality being strict, that uniquely solve*

$$P_\eta^B(z_1) = \beta_1 \quad \textit{and} \quad P_\eta^S(z_2) = \beta_2. \tag{28}$$

Lastly, if $\beta_1 + \beta_2 \leq P_\eta^B(\underline{z}_\eta)$ *then* $z_1 = z_2 = \underline{z}_\eta (= 0)$ *and the unique inclusive tie-breaking rule* $p^{(1,1)}$ *satisfies*

$$p_1^{(1,1)} = \frac{\beta_1}{P_\eta^B(0)} \quad \textit{and} \quad p_2^{(1,1)} = \frac{\beta_2}{P_\eta^B(0)}. \tag{29}$$

ii) if $\beta_1 > \beta_2 = 0$ *then an optimal scoring rule* g_2 *with* $z_2 = \underline{z}$ *exists. Moreover, there exists a unique scoring rule* g_1 *such that if* $\beta_1 > P_{F,\eta}^B(\underline{z}_\eta)$ *then* $z_1 = (P_{F,\eta}^B)^{-1}(\beta_1) > \underline{z}_\eta$. *On the other hand, if* $\beta_1 \leq P_{F,\eta}^B(\underline{z}_\eta)$ *then* $z_1 = \underline{z}_\eta (= 0)$ *and the unique inclusive tie-breaking rule* $p^{(1,1)}$ *satisfies*

$$p_1^{(1,1)} = \frac{\beta_1}{P_{F,\eta}^B(\underline{z}_\eta)} \quad \textit{and} \quad p_2^{(1,1)} = 0. \tag{30}$$

iii) if $\beta_1 = \beta_2 = 0$ *then* $z_1 = z_2 = \underline{z}_\eta$. *Moreover, ties occur with positive probability if* $P_{F,\eta}^B(\underline{z}_\eta) < 0$, *which in that case* $p^{(1,1)} \equiv 0$.

See Figure 3 for an illustration of the optimal ironing parameters z and inclusive tie-breaking rule $p^{(1,1)}$ for any dataset $\alpha \in \mathcal{A}_{0,1}$ and welfare weight η. The takeaway from this exercise is that characterizing the optimal mechanism becomes highly untractable due to the multidimensionality of the tie-breaking rules. I therefore enrich the model by introducing continuum of CTRs and show how to surprisingly overcome the issue of multidimensionality.

5 Continuum CTRs

> *"...the raw data are distribution functions rather than real numbers... Thus one may convincingly argue that distributions are the numbers of the future."*
>
> —B. Schweizer and A. Sklar, *Probabilistic Metric Spaces (1983)*

Characterizing optimal scoring mechanisms with finite set Ω is challenging since ties may occur with positive probability. In this section, I enrich the model introduced in Section 2 by considering a continuum of CTRs. Under some regularity conditions, I show that there exists a *unique* optimal scoring mechanism and derive a system of (nonlinear) equations to solve for scoring functions g_i. I then consider a stylized setting with iid. datasets that allows us to solve for g_i analytically and perform comparative statics on the model parameters and present asymptotic results. I then show that in large markets the optimal mechanism in ex-post implementable by designing a selling, a buying and an exchange market.

Specifically, let $\Omega = [0,1]^N$ and the probability space $(\mathcal{N}\times\Omega, \mathcal{F}, \alpha)$ be endowed with sigma-algebra $\mathcal{F} := \mathcal{P}(\mathcal{N}) \bigotimes \mathcal{B}(\Omega)$ defined by the product of the power set $\mathcal{P}(\mathcal{N})$ and Borel σ-algebra $\mathcal{B}(\Omega)$. For each $i \in \mathcal{N}$ let the individual dataset α_i be defined by the tuple $\alpha_i := (\lambda(i), \mu_i)$ where $\lambda \in \Delta(\mathcal{N})$ is a strictly positive *discrete* measure that represents the *mass distribution* across merchants' datasets, and $\mu_i \in \Delta(\Omega)$ is a *continuous* $\mathcal{B}(\Omega)$-measure that represents the CTR distribution within merchant i's dataset. Then let the *aggregate* dataset α in the market be the $\mathcal{F}$-measure defined by

$$\alpha(S\times E) := \sum_{i\in S} \lambda(i)\mu_i(E), \quad \forall S \in \mathcal{P}(\mathcal{N}), \forall E \in \mathcal{B}(\Omega). \tag{31}$$

Lastly, let $\alpha_{\mathcal{N}}$ denote the marginal distribution of CTRs on Ω where

$$\alpha_{\mathcal{N}}(E) := \alpha(\mathcal{N}\times E), \quad \forall E \in \mathcal{B}(\Omega). \tag{32}$$

5.1 Mechanism Design

I define the space of direct mechanisms that accommodate a *continuum* of instruments (x_i^ω) and then define the designer's problem analogously to the finite case.

Direct Mechanism. Let $\mathcal{X}$ be the space of $\mathcal{B}(\Omega)$-measurable functions $x : \Omega \to \mathbb{R}_+$. A *direct revelation mechanism* (henceforth, *mechanism*) $(x,t) = ((x_i)_{i\in\mathcal{N}_0}, (t_i)_{i\in\mathcal{N}_0})$ is given by measurable ads rule $x_i(\cdot) : \Theta \to \mathcal{X}$—that is, $x_i^{(\cdot)}(\theta) : \Omega \to \mathbb{R}_+$ is a $\mathcal{B}(\Omega)$-measurable function—and measurable transfer rule $t_i(\cdot) : \Theta \to \mathbb{R}$.

Definition 5 (*Feasibility*)**.** *A mechanism (x,t) is feasible if for every $\theta \in \Theta$:*

i) x distributes a unit mass of customers, that is,

$$\mathbb{E}_\alpha[x(\theta)] := \mathbb{E}_{\alpha_{\mathcal{N}}}\left[\sum_{i\in\mathcal{N}_0} x_i^\omega(\theta)\right] = 1.^{11} \tag{Unit}$$

[11] Without loss of generality the designer $i = 0$ serves as the 'sink' for all customers (if any) that are not getting targeted by any merchant $i \in \mathcal{N}$.

ii) the distribution of CTRs induced by x agrees with the true distribution $\alpha_{\mathcal{N}}$, that is,

$$\mathbb{E}_{\alpha}[x\cdot\mathbb{1}_E] := \mathbb{E}_{\alpha_{\mathcal{N}}}\left[\sum_{i\in\mathcal{N}_0} x_i^{\omega}(\theta)\cdot\mathbb{1}_E(\omega)\right] = \alpha_{\mathcal{N}}^E, \quad \forall E\in\mathcal{B}(\Omega). \tag{Marg}$$

Observe that (Marg) implies (Unit) by taking $E=\Omega$.

Next, at each $\theta\in\Theta$ define analogously

$$s_i(\theta) := \mathbb{E}_{\alpha_{\mathcal{N}}}[\omega_i x_i^{\omega}(\theta)] \quad \text{and} \quad S_i(\theta_i) := \mathbb{E}_{(\alpha_{\mathcal{N}},F_{-i})}[\omega_i x_i^{\omega}(\theta_i,\theta_{-i})], \tag{33}$$

where S_i denotes the expected clicks from the mechanism x. The expected clicks off-platform are given by

$$a_i := \lambda(i)\mathbb{E}_{\mu_i}[\omega_i], \quad \forall i\in\mathcal{N}. \tag{34}$$

Then merchant i's (net) interim payoff from a mechanism (x,t) is given by

$$U_i(\theta_i,S_i(\theta_i),T_i(\theta_i)) = \theta_i(S_i(\theta_i)-a_i)-T_i(\theta_i). \tag{35}$$

Thus, define (IC) and (IR) analogously and let $\mathscr{M}^{\alpha}(\Theta,dF)$ denote the space of feasible, (IC) and (IR) mechanisms.

Example 1 (*Radon–Nikodym (RN) Mechanism*)**.** *Consider the ads rule $x^{RN}=(x_i)_{i\in\mathcal{N}_0}$ defined by $x_0\equiv 0$ and for every $i\in\mathcal{N}$ and $\theta\in\Theta$ define*

$$x_i^{\omega}(\theta) := \lambda(i)h_i(\omega), \quad \forall\omega\in\Omega, \tag{36}$$

where the measurable function $h_i:\Omega\to\mathbb{R}_+$ is given by

$$h_i \equiv \frac{d\mu_i}{d\alpha_{\mathcal{N}}}, \tag{37}$$

i.e. h is the RN derivative of the merchant-specific measure μ_i with respect to the aggregate measure $\alpha_{\mathcal{N}}$. From the definitions (31) and (32), μ_i is absolutely continuous with respect to $\alpha_{\mathcal{N}}$, thus the RN derivative exists (and is unique a.e.). Now, x^{RN} is feasible since it satisfies (Marg):

$$\mathbb{E}_{\alpha_{\mathcal{N}}}\left[\sum_{i\in\mathcal{N}_0} x_i^{\omega}\cdot\mathbb{1}_E(\omega)\right] = \sum_{i\in\mathcal{N}}\lambda(i)\mathbb{E}_{\alpha_{\mathcal{N}}}[h_i(\omega)\cdot\mathbb{1}_E(\omega)] \tag{38}$$

$$= \sum_{i\in\mathcal{N}}\lambda(i)\mathbb{E}_{\mu_i}[\mathbb{1}_E(\omega)] = \sum_{i\in\mathcal{N}}\lambda(i)\mu_i(E) = \alpha_{\mathcal{N}}^E. \tag{39}$$

Moreover, letting $t\equiv 0$, the RN mechanism, given by $\left(x^{RN},0\right)$, guarantees every merchant's outside option:

$$\mathbb{E}_{\alpha_{\mathcal{N}}}[\omega_i x_i^{\omega}] = \lambda(i)\mathbb{E}_{\alpha_{\mathcal{N}}}[\omega_i h_i(\omega)] = \lambda(i)\mathbb{E}_{\mu_i}[\omega_i], \quad \forall i\in\mathcal{N}. \tag{40}$$

Example 1 shows $\mathscr{M}^{\alpha}(\Theta,dF)\neq\emptyset$ and that the designer can always do weakly better.

Designer's Problem. The designer's problem is thus defined analogously as in (P).

5.2 Optimal Mechanism

I define a class of mechanisms that are crucial for characterizing the optimal mechanism.

Let partition P of the measurable space $(\Omega, \mathcal{B}(\Omega))$ be a collection of *countable* disjoint subsets $P = \{E_j\}_{j\in J}$ for some index set J such that $E_j \in \mathcal{B}(\Omega)$ for all $j \in J$ and $\cup_{j\in J} E_J = \Omega$. Let $\mathscr{P}$ be the set of all partitions P of Ω.

Definition 6 (*Targeted Ads Rules*). *An ads rule x is partitional if for every $\theta \in \Theta$ there exists a measurable partition $P(\theta) = \{E_i(\theta)\}_{i\in\mathcal{N}\cup\{0\}}$ of $(\Omega, \mathcal{B}(\Omega))$ such that*

$$x_i^\omega(\theta) = 1_{E_i(\theta)}(\omega), \quad \forall \omega \in \Omega.$$

Then, a partitional ads rule x is targeted if for every $\theta \in \Theta$ the partition $P^z(\theta) = \{E_i^z(\theta)\}_{i\in\mathcal{N}\cup\{0\}}$ is such that for every $i \in \mathcal{N}$

$$E_i^z(\theta) := \Big\{\omega \in \Omega : \omega_i \phi_{\eta,i}(\theta_i, z_i(\hat{\theta}_i)) > \omega_j \phi_{\eta,j}(\theta_j, z_j(\hat{\theta}_j)), \forall j \neq i, j \in \mathcal{N}\Big\},$$

and $E_0^z(\theta) = \Omega \setminus \cup_{i\in\mathcal{N}} E_i^z(\theta)$.

Now let $\hat{\Theta}(x)$ be defined accordingly.

Theorem 2 (*Optimal Mechanism*). *For any arbitrarily chosen welfare weight η and dataset α, there exists a unique scoring function $g_i(\theta_i) = \phi_{\eta,i}(\theta_i, z_i(\hat{\theta}_i)$ with critical worst-off type $\hat{\theta}_i \in \hat{\Theta}(x)$ and $z_i(\hat{\theta}_i) \geq 0$ for every $i \in \mathcal{N}$. Specifically, if $z_i(\hat{\theta}_i) = 0$ for every $i \in \mathcal{N}$ and $\theta \in \hat{\Theta}(x)$, then $(x(\theta), t(\theta)) = (x^{RN}, 0)$ for all θ_i such that $g_i(\theta_i) = z_i$. Otherwise, for every $i \in \mathcal{N}$, $z_i > 0$ and either $z_i = \phi_{\eta,i}^S(1)$ or it satisfies:*

$$\mathbb{E}_{\alpha_\mathcal{N}}\left[\omega_i \left(\prod_{j\neq i, j\in\mathcal{N}} \mathbb{P}_{F_j}\Big(\omega_i z_i > \omega_j \phi_{\eta,j}(\theta_j, z_j)\Big)\right)\right] = a_i. \tag{Opt-z}$$

Given optimal ironing parameter z, for each $\theta \in \Theta$ the optimal ads rule x is targeted—i.e. it is a partitional rule with $P^z(\theta) = \{E_i^z(\theta)\}_{i\in\mathcal{N}\cup\{0\}}$. Lastly, the expected (interim) transfers are given as in (14).

Proof Sketch. The first observation is that all the steps in the proof of Theorem 1 go through, thus allowing us to reduce the designer's problem as following:

$$\max_{\substack{x \text{ pointwise a.e.} \\ \text{s.t. (Marg)}}} \mathbb{E}_{(\alpha_\mathcal{N}, F)}\left[\sum_{i\in\mathcal{N}} \omega_i x_i^\omega(\theta) \phi_{\eta,i}(\theta_i, z_i(\hat{\theta}_i))\right] \tag{41}$$

$$\text{subject to} \quad \hat{\theta} \in \hat{\Theta}(x). \tag{42}$$

Therefore, it follows that there exists ironing parameters z such that the pointwise maximizer x of (41) given z yields $\hat{\theta}(z) \in \hat{\Theta}(x)$.

So, let z be an optimal ironing parameter with $z_i > 0$ for some $i \in \mathcal{N}$. I show in Appendix D that the optimal ads rule takes the form $x_i^\omega(\theta) = \mathbb{1}_{E_i^z(\theta)}(\omega)$ a.e. in Ω. But then, one can compute explicitly the interim expected clicks at $\hat{\theta}(z)$. Specifically, for each $i \in \mathcal{N}$, $S_i(\hat{\theta}_i(z_i))$ are given by

$$\begin{aligned}
&\mathbb{E}_{(\alpha_\mathcal{N}, F_{-i})}\left[\omega_i \mathbb{1}_{E_i^z\left(\hat{\theta}_i(z_i), \theta_{-i}\right)}\right] = \mathbb{E}_{\alpha_\mathcal{N}}\left[\omega_i \mathbb{E}_{F_{-i}}\left[\mathbb{1}_{E_i^z\left(\hat{\theta}_i(z_i), \theta_{-i}\right)}\right] \mid \omega\right] \\
&= \mathbb{E}_{\alpha_\mathcal{N}}\left[\omega_i \mathbb{P}_{F_{-i}}\Big(E_i^z\Big(\hat{\theta}_i(z_i), \theta_{-i}\Big)\Big)\right] = \mathbb{E}_{\alpha_\mathcal{N}}\left[\omega_i \left(\prod_{j\neq i, j\in\mathcal{N}} \mathbb{P}_{F_j}\Big(\omega_i z_i > \omega_j \phi_{\eta,j}(\theta_j, z_j)\Big)\right)\right]
\end{aligned} \tag{43}$$

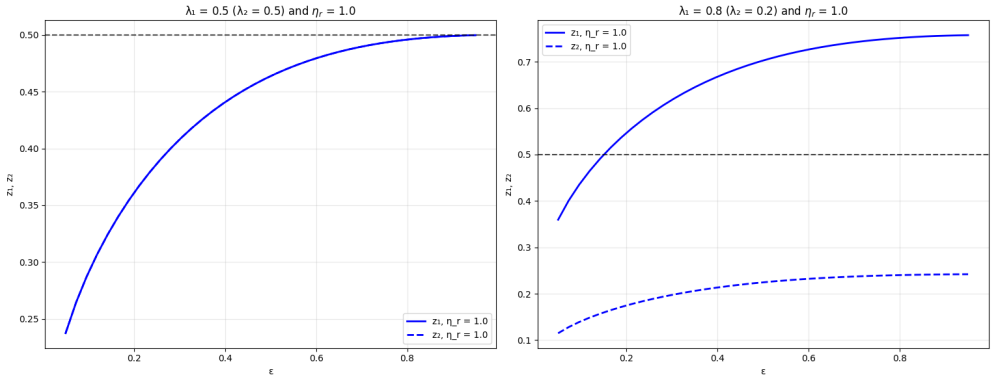


Figure 4: Optimal $(z_1(\varepsilon), z_2(\varepsilon))$ for $\eta_r = 1$, $\lambda(1) = 0.5$ (left) and $\lambda(1) = 0.8$ (right).

where the first equality is derived using the law of iterated expectations and the independence of the distributions F and $\alpha_{\mathcal{N}}$, and the last equality uses the independence of F_i's. Moreover, since $a_i > 0$, and given $z_i > 0$ for some $i \in \mathcal{N}$, from the last term in (43) we can see that it must also be the case that $z_i > 0$ for all $i \in \mathcal{N}$. Therefore, from definition of $\hat{\Theta}(x)$ it follows that if no z_i that makes (43) equal to a_i exists, then $z_i = \phi^S_{\eta,i}(1)$ and $S_i(\hat{\theta}_i(z_i)) < a_i$.

5.3 A Stylized Setting: IID Datasets

Let $\Omega = (\varepsilon, 1]^N$ for some $\varepsilon \geq 0$.[12] Suppose customer heterogeneity μ_i on each individual dataset is a product of iid. distributions, i.e. $\mu_i := \prod_{i \in \mathcal{N}} \alpha_i$ where $\alpha_i \in \Delta((\varepsilon, 1])$ for every $i \in \mathcal{N}$. Then $\alpha_{\mathcal{N}} \equiv \prod_{i \in \mathcal{N}} \alpha_i$ too and for every $i \in \mathcal{N}$, (43) can be written as

$$\mathbb{E}_{\alpha_i}\left[\omega_i \prod_{j \neq i, j \in \mathcal{N}} \mathbb{E}_{\alpha_j}\left[\mathbb{P}_{F_j}\Big(\omega_i z_i > \omega_j \phi_{\eta,j}(\theta_j, z_j)\Big)\Big|\omega_i\right]\right] = \lambda(i)\mathbb{E}_{\alpha_i}[\omega_i]. \tag{44}$$

Example 2. *Consider $N = 2$ with $\theta_i \sim$ Uniform$[0,1]$, $\omega_i \sim$ Uniform$[\varepsilon, 1]$, $\eta_w = 0$ and $\lambda(1) > \lambda(2)$. In Figure 4 I plot the optimal solutions for $\eta_r = 1$ and mass distributions $\lambda_1 = 0.5$ and $\lambda_2 = 0.8$ while varying data quality ε. Note that as $\varepsilon \to 1$ for $\eta_r = 1$ and $\lambda(1) = 0.5$, the optimal $z_i \to 1$, thus agreeing with the profit maximizing mechanism in the partnership dissolution model. Moreover, lower data quality implies that merchants except to often benefit from the platform and be net buyers (thus z small). Lastly, merchants with larger datasets have scoring rules with higher ironing regions.*

Next, consider the optimal design for symmetric λ for arbitrary $N \geq 2$, i.e. $\lambda(i) = 1/N$ for all $i \in \mathcal{N}$. Then for every $i, j \in \mathcal{N}$ we have $z_i = z_j = z \in (0, 1)$ where (44) is simplified to a single non-linear equation:

$$\mathbb{E}_{\alpha_1}\left[\omega_1\left(\mathbb{E}_{\alpha_2}\left[\mathbb{P}_{F_2}\Big(\omega_1 z > \omega_2 \phi_{\eta,2}(\theta_2, z)\Big)\Big|\omega_1\right]\right)^{N-1}\right] = \frac{\mathbb{E}_{\alpha_1}[\omega_1]}{N}. \tag{45}$$

[12] I interpret the parameter ε as data quality and heterogeneity where higher ε makes the data more homogeneous and likely have higher CTRs.

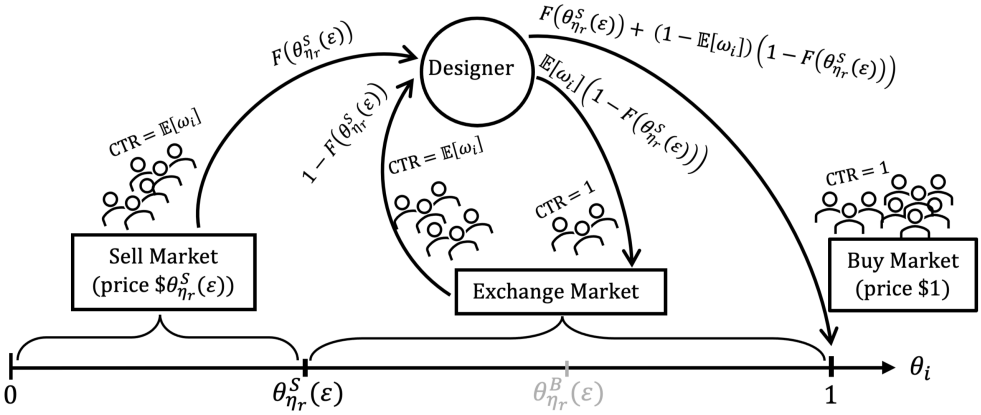


Figure 5: Optimal mechanism when $N \to \infty$.

Proposition 2 (*Optimal Mechanism In Large Platforms*)**.** *Let $z^\infty_\eta(\varepsilon)$ be the limit solution to (45). The optimal mechanism shares data and designs ads using three markets (see Figure 5):*

i) a selling market for which merchants can sell all their data at a posted the price $p^S = \theta^S_\eta(\varepsilon)$ per unit of average quality

ii) an exchange market in which merchants sell all their data and get in return ads with $CTR = 1$ that generate expected clicks a_i

iii) a selling market in which high types $\theta_i = 1$ buy ads with $CTR = 1$ at the full price $p^B = 1$

There are two main takeaways (among others) from Proposition 2: the first is that the optimal mechanism is ex-post implementable in large platforms by designing the three markets, and the second is that a new market—an *exchange market*—arises in the optimum contrasting with the Bid-Ask spread pricing mechanisms.

Example 3. *Suppose types θ_i and CTRs ω_i are drawn iid. from $\theta_i \sim Uniform[0,1]$ and $\omega_i \sim Uniform[\varepsilon, 1]$, respectively, $\eta_w = 0$ and $\lambda(i) = 1/N$ for every $i \in \mathcal{N}$. Then (45) reduces to*

$$(a-b)\left(\frac{N}{N+1}(a-b)+b\right)\left(\frac{a-b}{(1-\varepsilon)(1+\eta_r)}\right)^{N-1} - \frac{a^2(1-\varepsilon^2)}{2} \qquad \text{(LHS)}$$

$$= (a\varepsilon-b)\left(\frac{N}{N+1}(a\varepsilon-b)+b\right)\left(\frac{a\varepsilon-b}{(1-\varepsilon)(1+\eta_r)}\right)^{N-1} \qquad \text{(RHS)}$$

where $a = a(z,\varepsilon,\eta) = z\ln(1/\varepsilon) + \eta_r > 0$ and $b = b(\varepsilon,\eta) = \eta_r\varepsilon > 0$. Taking $N \to \infty$ the optimal z^∞ is given by

$$z^\infty(\varepsilon) = (1-\varepsilon)/\ln(1/\varepsilon), \tag{46}$$

where for this example $z^\infty(\varepsilon)$ doesn't depend on η_r. Then, from Figure 5 we compute the

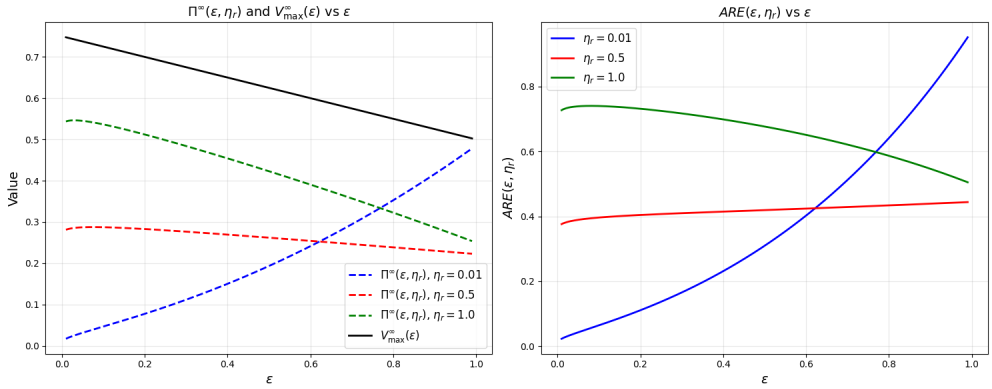


Figure 6: $\text{ARE}(\eta_r, \varepsilon)$ for $\eta_r \in \{0.01, 0.5, 1\}$.

revenue function:

$$Rev^{\infty}(\varepsilon,\eta_r) = \Big(1-\theta^S_{\eta_r}(\varepsilon)\mathbb{E}[\omega_1]\Big)F\Big(\theta^S_{\eta_r}(\varepsilon)\Big) + \Big(1-\mathbb{E}[\omega_1]\Big)\Big(1-F\Big(\theta^S_{\eta_r}(\varepsilon)\Big)\Big) \tag{47}$$

where $\mathbb{E}[\omega_1] = (1+\varepsilon)/2$, $F\Big(\theta^S_{\eta_r}(\varepsilon)\Big) = \theta^S_{\eta_r}(\varepsilon)$ *and* $\theta^S_{\eta_r}(\varepsilon) = z^{\infty}(\varepsilon)/(1+\eta_r) \downarrow$ *decreasing in* η_r. *For given* ε *and* η_r, *the asymptotic relative efficiency (ARE) of platform design is defined as*

$$ARE(\varepsilon,\eta_r) = \frac{\Pi^{\infty}(\varepsilon,\eta_r)}{V^{\infty}_{\max}(\varepsilon)} = \frac{(1-\eta_r)V^{\infty}(\varepsilon,\eta_r) + \eta_r Rev^{\infty}(\varepsilon,\eta_r)}{V^{\infty}_{\max}(\varepsilon)} \tag{48}$$

where $V^{\infty}(\varepsilon,\eta_r)$ *denotes the merchants' surplus from the mechanism and is given by*

$$V^{\infty}(\varepsilon,\eta_r) = \mathbb{E}[\omega_1]\mathbb{E}\Big[\Big(\theta^S_{\eta_r}(\varepsilon)-\theta_1\Big)_+\Big] = \frac{1+\varepsilon}{2}\frac{\Big(\theta^S_{\eta_r}(\varepsilon)\Big)^2}{2}$$

and $V^{\infty}_{\max}(\varepsilon)$ *denotes total attainable merchants' surplus and is given by*

$$V^{\infty}_{\max}(\varepsilon) = \lim_{N\to\infty}\left(\mathbb{E}_{(\alpha_{\mathcal{N}},F)}\Big[\max_{i\in\mathcal{N}}\omega_i\theta_i\Big] - \sum_{i\in\mathcal{N}}\lambda(i)\mathbb{E}_{(\mu_i,F_i)}[\omega_i\theta_i]\right) = \frac{3-\varepsilon}{4}.$$

In Figure 6 I plot $ARE(\eta_r,\varepsilon)$ *for* $\eta_r \in \{0.01, 0.5, 1\}$. *We see that ARE is higher (lower) for lower (higher) data quality (value of* ε*) when* η_r *is sufficiently high (low). That is, when merchants have high data quality then ARE is high when the price at the selling market is sufficiently high (or* η_r *is sufficiently small)—it reduces inefficient trading regions and most of the surplus is transferred to the merchants. On the other hand, lower data quality makes the exchange market profitable and hence most of the surplus is enjoyed by the designer when* η_r *is sufficiently high enough.*

6 Revisiting Model Assumptions and Applications

6.1 Revisiting Model Assumptions

Limited Commitment. One other frequently asked question—"Can I exclude specific customers from my audiences?"[13]—raises the question if the designer should make the choice to choose which data to share as part of the design. Formally, this speaks to the literature on mechanism design with limited commitment pioneered by Bester and Strausz (2001). Generally, analyzing such problems can be extremely difficult and it is left for future research. However, we know that full commitment generates highest values for the designer and is also what Shopify claims to be doing: "...you can't add or exclude customers to a specific audience."

Data Complementarities. Recall that customers have unit-demand and merchants don't compete in prices, so the designer chooses without loss of generality one "winner" to target a given customer. In practice, the advantages of designing such platforms can be significant if leveraging data insights from complementary products—"the Audiences algorithms might provide insights based on people who have purchased complementary products in the past."[14] For example, if you sell swimsuits, then the designer can provide you with targeted ads about customers who have purchased complementary products like sandals, sunglasses, sunscreen and sunhat. Analyzing targeted advertising platforms where customers have multi-unit demand and/or merchants can choose prices is left for future research.

6.2 Applications

The model introduced in Section 2 is broad in scope and is applicable to other economic settings that involve trading of goods among a set of agents.

Greenhouse gas (GHG) credit market. A GHG credit (e.g., carbon credit) is "a tradable instrument that conveys a claim to avoided GHG emissions or to the enhanced removal of GHG from the atmosphere."[15] The GHG credit market is a trading system in which firms buy and sell rights to emit specific GHGs. Each firm $i \in \mathcal{N}$ holds a portfolio of credits $\alpha_i = \{\alpha_i^\omega\}$ across GHG $\omega \in \Omega$. Let ω_i represent the intensity of using GHG ω at the industry where i operates. Firms need to comply with emissions targets set by regulatory agencies and have private compliance cost of θ_i—for example, outdated, emissions-intensive tech and fossil-heavy energy have a higher compliance cost θ_i than modern and cleaner tech, thus willing to pay more for a unit of GHG credit.

Article 6 of the 2015 Paris Agreement on climate change enables parties to cooperate in reducing their greenhouse gas emissions—this means that emission reductions can be transferred between countries and counted towards nationally determined contributions. The

[13] https://help.shopify.com/en/manual/promoting-marketing/shopify-audiences/faq#can-i-exclude-specific-customers-from-my-audiences

[14] https://www.shopify.com/partners/blog/shopify-audiences

[15] https://en.wikipedia.org/wiki/Carbon_offsets_and_credits

Paris Agreement offers ways in which parties may cooperate on a voluntary basis including market mechanisms like the carbon credit market as well as non-market-based approaches. This paper offers a market mechanism to trade all types of GHG to help meet the desired emission levels.

Reallocating public resources. Public resources like water/land rights and fishing quotas are often distributed based on outdated criteria or arbitrarily. For example, California's Water Allocation during the 2012–2016 Drought distributed water rights based on a seniority system dating back to the 19th century.

The paper proposes a market mechanism in which locals ($\mathcal{N}$) trade resources ω (like water/land rights and fishing quotas) where ω_i represents how accessible/ease of use the resource ω is to local i, and θ_i represents personal costs (e.g., time, money, effort) to capitalize on gains from any resource.

7 Conclusions

I design targeted advertising platforms that host merchants who share their data to improve targeted advertising though it raises poaching concerns. I show that the optimal design that alleviates poaching concerns by proving merchants with either higher CTR ads or monetary compensations. In a large market the optimal design simplifies nicely to the design of three markets: a selling, a buying and an exchange market. I also discuss model extension regarding data complementarities and the limited commitment case, and their analysis is left for future research. Lastly, I conclude by emphasizing that the model speaks to other market design settings like the GHG credit markets and reallocation of public resources.

References

Bergemann, D. and A. Bonatti (2015): "Selling cookies," *American Economic Journal: Microeconomics*, 7, 259–294.

——— (2024): "Data, competition, and digital platforms," *American Economic Review*, 114, 2553–2595.

Bergemann, D., A. Bonatti, and T. Gan (2022): "The economics of social data," *The RAND Journal of Economics*, 53, 263–296.

Bergemann, D., A. Bonatti, and N. Wu (2024): "How Do Digital Advertising Auctions Impact Product Prices?" *Review of Economic Studies*, rdae087.

Bester, H. and R. Strausz (2001): "Contracting with imperfect commitment and the revelation principle: the single agent case," *Econometrica*, 69, 1077–1098.

Cramton, P., R. Gibbons, and P. Klemperer (1987): "Dissolving a partnership efficiently," *Econometrica: Journal of the Econometric Society*, 615–632.

Elliott, M., A. Galeotti, A. Koh, and W. Li (2022): "Matching and information design in marketplaces," *Available at SSRN 4283968*.

Farboodi, M. and L. Veldkamp (2023): "Data and markets," *Annual Review of Economics*, 15, 23–40.

Galperti, S., A. Levkun, and J. Perego (2024): "The value of data records," *Review of Economic Studies*, 91, 1007–1038.

Galperti, S. and J. Perego (2022): "Competitive markets for personal data," *Available at SSRN 4309966*.

Loertscher, S. and C. Wasser (2019): "Optimal structure and dissolution of partnerships," *Theoretical Economics*, 14, 1063–1114.

Myerson, R. B. (1981): "Optimal auction design," *Mathematics of operations research*, 6, 58–73.

Myerson, R. B. and M. A. Satterthwaite (1983): "Efficient mechanisms for bilateral trading," *Journal of economic theory*, 29, 265–281.

von Neumann, J. (1928): "Zur theorie der gesellschaftsspiele," *Mathematische annalen*, 100, 295–320.

Yang, F. (2023): "Nested bundling," Tech. rep., Working Paper.

Appendix

A Proof of Lemma 1

By envelope theorem, $U_i'(\theta_i) = S_i^*(\theta_i) - a_i$ almost everywhere, implying that $U_i(\theta_i)$ is convex in θ_i by (M) and absolute continuity of $U_i(\cdot)$.[16] Then, similar to Cramton et al. (1987), the set $\hat{\Theta}_i^\alpha(x,t)$ can be characterized in terms of ads rule x^* alone as in (12).

B Proof of Theorem 1

Eliminating (IR). Fix an IC mechanism (x,t). By definition, it is sufficient that (IR) holds at every $\hat{\theta} \in \hat{\Theta}^\alpha(x,t)$, which from optimality has $U_i(\hat{\theta}_i) = 0$. Now note that the RHS in (17) is constant for all $\theta' \in \Theta$. By definition, for any two arbitrary $\hat{\theta} \in \hat{\Theta}^\alpha(x,t)$ and $\theta' \in \Theta \setminus \hat{\Theta}^\alpha(x,t)$ we have $U_i(\theta') \geq U_i(\hat{\theta}_i)$ for all $i \in \mathcal{N}$, with at least one inequality being strict. It then follows that

$$\Psi_\eta^\alpha(x,\theta') > \Psi_\eta^\alpha(x,\hat{\theta}). \tag{49}$$

Since Θ is compact, the worst-off types admit the following characterization as a function of the ads rule x alone:

$$\hat{\Theta}(x) = \operatorname*{argmin}_{\theta' \in \Theta} \Psi_\eta^\alpha(x,\theta'). \tag{50}$$

It then follows that for any $\theta' = \hat{\theta}(x) \in \hat{\Theta}(x)$, using (17) we get

$$\Psi(x,t) = \Psi_\eta^\alpha(x,\hat{\theta}(x)) = \min_{\theta' \in \Theta} \Psi_\eta^\alpha(x,\theta'). \tag{51}$$

Therefore, the designer's problem is equivalently given by

$$\Pi(\alpha,\eta) = \sup_{x \in \mathcal{X}^\alpha(\Theta,dF)} \min_{\theta' \in \Theta} \Psi_\eta^\alpha(x,\theta'). \tag{Sup-Min}$$

Lastly, note that $\mathcal{X}^\alpha(\Theta,dF) \subseteq \mathcal{L}(\Theta,dF)$ where ads rules $x \equiv x'$ dF-a.e. are identified and $\mathcal{L}(\Theta,dF)$ is endowed with the weak*-topology. Since constraints on x are linear constraints, $\mathcal{X}^\alpha(\Theta,dF)$ is a convex set and by Alaoglu's theorem $\mathcal{X}^\alpha(\Theta,dF)$ is a compact set. Therefore, one can replace the sup in (Sup-Min) with max.

Existence. Observe first that $\Psi_\eta^\alpha(\cdot,\theta') : \mathcal{X} \to \mathbb{R}$ is a linear functional for every $\theta' \in \Theta$, hence continuous. On the other hand, $\Psi_\eta^\alpha(x,\cdot) : \Theta \to \mathbb{R}$ is continuous at every $x \in \mathcal{X}^\alpha(\Theta,dF)$ since we can write $\Psi(x,\theta')$ as

$$\Psi(x,\theta') = \sum_{i \in \mathcal{N}} \left(\int_0^{\theta_i'} (S_i(\theta_i) - a_i) \phi_{\eta,i}^S(\theta_i) dF_i(\theta_i) + \int_{\theta_i'}^1 (S_i(\theta_i) - a_i) \phi_{\eta,i}^B(\theta_i) dF_i(\theta_i) \right). \tag{52}$$

[16] If a function h is absolutely continuous and h' is increasing *almost everywhere* then h is convex.

Since S_i, $\phi^S_{\eta,i}$ and $\phi^B_{\eta,i}$ are measurable functions, it follows that (see Royden, 4th Ed., Chapter 5, Theorem 14) $\Psi(x,\cdot)$ is absolutely continuous and

$$\frac{\partial \Psi(x,\theta')}{\partial \theta'_i} = S_i(\theta'_i) - a_i \quad \text{and} \quad \frac{\partial^2 \Psi(x,\theta')}{\partial \theta'_j \partial \theta'_i} = 0 \quad dF\text{-a.e.} \tag{53}$$

Then, since $\partial\Psi(x,\cdot)/\partial\theta'$ increasing dF-a.e. by (M), $\Psi(x,\cdot)$ is convex at every $x \in \mathcal{X}^\alpha(\Theta, dF)$. Lastly, $\Psi(\cdot,\theta')$ is continuous at each $\theta' \in \Theta$. Hence, applying the generalized version of von Neumann's minimax theorem we get the existence of a saddle point (i.e. an optimal mechanism) and swap the order of optimization.

Ironing. I show in online appendix that there exists an *essentially unique* critical worst-off type $\theta^* \in \Theta(x^*)$ such that when ironing with respect to θ^*, (25) admits a solution that also satisfies (23).

C Proof of Proposition 1

I construct EP scoring mechanisms by finding (essentially unique) ironing parameter z and inclusive tie-breaking rule $p^{(1,1)}$ such that the solution x^* to (26) satisfies $\hat{\theta}_i(z_i) \in \hat{\Theta}_i(x^*)$ for every $i = 1, 2$.[17]

Observe first that symmetry on F implies $\phi_{\eta,i}(s, z_i) \geq \phi_{\eta,j}(s, z_j)$ for every $s \in [\underline{\theta}, \overline{\theta}]$ if and only if $z_i \geq z_j$. Then, the optimal clicks $S_i^{(1,1)}(\hat{\theta}_i(z_i)) < S_j^{(1,1)}(\hat{\theta}_j(z_j))$ whenever $(\underline{z}_\eta \leq) z_i < z_j$ and $\alpha_{\mathcal{N}}^{(1,1)} > 0$, regardless of the inclusive tie-breaking rule $p^{(1,1)}$. Therefore, if merchant i has a higher adjusted share, i.e. $\beta_i \geq \beta_j$, then the optimal ironing parameter is such that $z_i \geq z_i$ to ensure (in expectation) the critical worst-off type $\hat{\theta}_i(z_i)$ its outside option.

i) $\alpha \in \mathcal{R}_{+,+}$ and $\beta_1 \geq \beta_2$. Consider first the case $\phi^B_{\eta,i}(0) < 0$.[18] I show that there exists $z_1 \geq z_2$ and $p^{(1,1)}$ such that $S_i(\hat{\theta}_i(z_i)) = a_i$.

Since for every $i = 1, 2$ the adjusted share $\beta_i > 0$ then $\alpha_i^{(1,1)} > \alpha_j^{\omega_i}$. By definition of EP scoring mechanism and optimality, the critical worst-off type $\hat{\theta}_i(z_i)$ is assigned all of the exclusive targeted ads $\alpha_j^{\omega_i}$, regardless of the value of z_i.

Then, suppose first that there exists $\tilde{z} \geq 0$ such that $z_1 = z_2 = \tilde{z}$. From (26) the condition $S_i(\hat{\theta}_i(\tilde{z})) = a_i$ for each $i = 1, 2$ reduces to solving the following system of (nonlinear) equations:

$$\begin{cases} \alpha_{\mathcal{N}}^{(1,1)} \left(P^S_{\eta,2}(\tilde{z}) + p_1^{(1,1)} \left(P^B_{\eta,2}(\tilde{z}) - P^S_{\eta,2}(\tilde{z}) \right) \right) + \alpha_2^{(1,0)} = \alpha_1^{(1,1)} \\ \alpha_{\mathcal{N}}^{(1,1)} \left(P^S_{\eta,1}(\tilde{z}) + p_2^{(1,1)} \left(P^B_{\eta,1}(\tilde{z}) - P^S_{\eta,1}(\tilde{z}) \right) \right) + \alpha_1^{(0,1)} = \alpha_2^{(1,1)} \end{cases} \tag{54}$$

for some inclusive tie-breaking rule $p^{(1,1)}$ such that $p_1^{(1,1)} + p_2^{(1,1)} = 1$ whenever $\tilde{z} > 0$. For every $k = B, S$, since $P^k_{\eta,1} \equiv P^k_{\eta,2}$ by symmetry on F, define $P^k_\eta := P^k_{\eta,i}$ for every $i = 1, 2$ and

[17] Note that $p^{(1,1)}(z)$ matters only when $z_1 = z_2$ and $\phi_{\eta,i}(\theta_i, z_i) = z_i$ occurs with strictly positive probability.

[18] It follows that $\underline{z}_i := \max\{0, \phi^B_{\eta,i}(0)\} = 0$, $P^B_{\eta,i}(0) > 0$ and $P^S_{\eta,i}(z_i) = 0$ for every $z_i \in [0, \phi^S_{\eta,i}(0)]$, where $\phi^S_{\eta,i}(0) = \eta_w \geq 0$ for every $i = 1, 2$.

rearrange (54) to get

$$p_i^{(1,1)} P_\eta^B(\tilde{z}) + \left(1 - p_i^{(1,1)}\right) P_\eta^S(\tilde{z}) = \beta_i, \quad \forall i = 1, 2. \tag{55}$$

Now, adding both sides of the equations in (55) gives

$$\left(p_1^{(1,1)} + p_2^{(1,1)}\right) P_\eta^B(\tilde{z}) + \left(2 - p_1^{(1,1)} - p_2^{(1,1)}\right) P_\eta^S(\tilde{z}) = \beta_1 + \beta_2. \tag{56}$$

Then, since $P_\eta^S(\cdot)$ is non-decreasing and $P_\eta^B(\cdot)$ is strictly increasing on $[0,1]$, $\beta_1 + \beta_2 \leq 1$ and $p_1^{(1,1)} + p_2^{(1,1)} = 1$ whenever $\tilde{z} > 0$, from (56) it follows readily that $\tilde{z} = \underline{z}_i (= 0)$ if and only if $\beta_1 + \beta_2 \leq P_\eta^B(\underline{z}_i)$. In that case, using (55) it follows that there exists unique inclusive tie-breaking rules $p_i^{(1,1)} > 0$ such that $p_i^{(1,1)} = \beta_i / P_\eta^B(\underline{z}_i)$ for every $i = 1, 2$.

So, suppose instead that $\beta_1 + \beta_2 > P_\eta^B(\underline{z}_i)$. Then, $\tilde{z} > 0$ and $p_1^{(1,1)} + p_2^{(1,1)} = 1$, and the identity in (56) reduces to

$$P_\eta^B(\tilde{z}) + P_\eta^S(\tilde{z}) = \beta_1 + \beta_2. \tag{57}$$

Since $P_\eta^B(1) = 1$, it follows readily that there exists a unique $\tilde{z} \in [0,1]$ satisfying (57). Then, using (55), solving for $p_1^{(1,1)}$ yields

$$p_1^{(1,1)} = \frac{1}{2}\left(1 + \frac{\beta_1 - \beta_2}{P_\eta^B(\tilde{z}) - P_\eta^S(\tilde{z})}\right), \tag{58}$$

where $P_\eta^B(\tilde{z}) - P_\eta^S(\tilde{z}) > 0$ since F is regular. Hence, a unique inclusive tie-breaking rule $p^{(1,1)}$ exists with $z_1 = z_2 = \tilde{z} > 0$ if and only if (58) is well-defined, i.e. $p_1^{(1,1)} \leq 1$ or $P_\eta^B(\tilde{z}) - P_\eta^S(\tilde{z}) \geq \beta_1 - \beta_2$.

Thus, suppose now that $\beta_1 + \beta_2 > P_\eta^B(\underline{z}_i)$ and $P_\eta^B(\tilde{z}) - P_\eta^S(\tilde{z}) < \beta_1 - \beta_2$, where $\tilde{z}$ is the solution to (57). I show that there exists unique $z_1 \geq \tilde{z} \geq z_2 > \eta_w$, with at least one of the inequalities being strict.

To see this, note that the analog of the nonlinear system in (54) that the designer solves when $z_1 > z_2$ is given by

$$\begin{cases} \alpha_{\mathcal{N}}^{(1,1)} P_\eta^B(z_1) + \alpha_2^{(1,0)} = \alpha_1^{(1,1)} \\ \alpha_{\mathcal{N}}^{(1,1)} P_\eta^S(z_2) + \alpha_1^{(0,1)} = \alpha_2^{(1,1)} \end{cases} \tag{59}$$

Rearranging (59) yields

$$P_\eta^B(z_1) = \beta_1 \text{ and } P_\eta^S(z_2) = \beta_2. \tag{60}$$

Then, since $\beta_2 > 0$ and $P_\eta^S(z_2) = 0$ for all $z_2 \leq \eta_w$, the ironing parameter z_2 must satisfy $z_2 > \phi_\eta^S(0)$. Moreover, since $P_\eta^B(\tilde{z}) + P_\eta^S(\tilde{z}) = \beta_1 + \beta_2$ and $P_\eta^B(\tilde{z}) - P_\eta^S(\tilde{z}) < \beta_1 - \beta_2$, it follows readily that $\beta_2 \leq P_\eta^S(\tilde{z}) < P_\eta^B(\tilde{z}) \leq \beta_1$. Then, since $P_\eta^S(\cdot)$ is strictly increasing on $[\eta_w, \tilde{z}]$ and $P_\eta^B(\cdot)$ is strictly increasing on $[\tilde{z}, 1]$ with $P_\eta^S(\eta_w) = 0$ and $P_\eta^B(1) = 1$, there exists unique $\tilde{z} \downarrow z_2$ and $\tilde{z} \uparrow z_1$ that solve (60).

Lastly, consider the case $\phi_{\eta,i}^B(0) \geq 0$.[19] Here, we always have $\beta_1 + \beta_2 > P_\eta^B(\underline{z}_i) = 0$. The rest is identical to the preceding steps. □

ii) $\alpha \in \mathcal{R}_{+,0}$. Consider first the case $\phi_{\eta,i}^B(0) < 0$. I show that there exists $z_1 \geq z_2$ and $p^{(1,1)}$ such that $S_i(\hat{\theta}_i(z_i)) = a_i$.

[19] Now $\underline{z}_i := \phi_{\eta,i}^B(0) = \eta_w - \eta_r / f(0) \geq 0$, $P_{\eta,i}^B(\underline{z}_i) = 0$ and $P_{\eta,i}^S(z_i) = 0$ for every $z_i \in [\underline{z}_i, \eta_w]$ and $i = 1, 2$.

Fix some optimal z_1. Since $\beta_2 = 0$, suppose first that $\alpha_2^{(1,1)} < \alpha_1^{(0,1)}$. Then, to have $S_2(\hat{\theta}_2(z_2)) = \alpha_2^{(1,1)}$ it must be that $z_2 = 0$, as otherwise by optimality merchant 2 would target all $\alpha_1^{(0,1)}$ mass of customers if $z_2 > 0$. Therefore, by definition of EP-scoring mechanism the designer sets $p_2^{(0,1)} = \alpha_2^{(1,1)}/\alpha_1^{(0,1)}$. On the other hand, if $\alpha_2^{(1,1)} = \alpha_1^{(0,1)}$, then any $0 \leq z_2 \leq \min\{z_1, \eta_w\}$ can be optimal. Suppose to the contrary that $(z_1 \geq) z_2 > \eta$. But then merchant 2 wins the right to target inclusive customers with probability $P_\eta^S(z_2) > 0$, yielding in total $S_2(\hat{\theta}_2(z_2)) > \alpha_2^{(1,1)}$. For simplicity of exposition, I let $z_2 = \underline{z}_\eta = 0$.

To solve for z_1, suppose first that $\beta_1 \leq P_\eta^B(\underline{z}_\eta)$. Then setting $z_1 = \underline{z}_\eta$ and inclusive tie-breaking rule $p^{(1,1)}$ such that $p_1^{(1,1)} = \beta_1/P_\eta^B(0)$ and $p_2^{(1,1)} = 0$ yields $S_i(\hat{\theta}_i(z_i)) = \alpha_i^{(1,1)}$ for every $i = 1, 2$. On the other hand, if $\beta_1 > P_\eta^B(0)$ then similar to part i) it follows readily that there exists unique $z_1 = (P_\eta^B)^{-1}(\beta_1) > 0$. If $z_2 = z_1$, then break ties in favor of merchant 1, i.e. $p_1^{(1,1)} = 1$.

Next, consider the case $\phi_{\eta,i}^B(0) \geq 0$. I show that there exists $z_1 \geq z_2$ and $p^{(1,1)}$ such that $S_i(\hat{\theta}_i(z_i)) = a_i$, unless $\phi_{\eta,i}^B(0) > 0$ *and* $\alpha_2^{(1,1)} < \alpha_1^{(0,1)}$, in which case $S_i(\hat{\theta}_i(z_i)) > a_i$.

Thus, suppose first $\phi_{\eta,i}^B(0) = 0$ or $\alpha_2^{(1,1)} = \alpha_1^{(0,1)}$. Solving for z_2 is similar to the preceding steps: any $\underline{z}_\eta \leq z_2 \leq \min\{z_2, \eta_w\}$ is optimal. On the other hand, since $\beta_1 > 0 = P_\eta^B(\underline{z}_\eta)$, it follows readily that $z_1 = (P_\eta^B)^{-1}(\beta_1) > \underline{z}_\eta$. Lastly, if $z_2 = z_1$, then similarly break ties in favor of merchant 1, i.e. $p_1^{(1,1)} = 1$.

Now suppose that $\phi_{\eta,i}^B(0) > 0$ and $\alpha_2^{(1,1)} < \alpha_1^{(0,1)}$. Since $\underline{z}_\eta = \phi_{\eta,i}^B(0) > 0$, then by optimality $p_2^{(0,1)} = 1$. Thus, any $\underline{z}_\eta \leq z_2 \leq \min\{z_2, \eta_w\}$ is optimal but gives $S_2(\hat{\theta}_2(z_2)) > a_2$. Solving for z_1 we have again $z_1 = (P_\eta^B)^{-1}(\beta_1) > \underline{z}_\eta$, where again ties are broken in favor of merchant 1. □

iii) $\alpha \in \mathcal{R}_{0,0}$. Clearly, $z_i = \underline{z}_i$ for each $i = 1, 2$. Now, if $\phi_{\eta,i}^B(\underline{z}_\eta) \leq 0$ then by definition of EP scoring mechanism the designer assigns $p_i^{\omega_i} = \alpha_i^{(1,1)}/\alpha_j^{\omega_i}$ and $p_i^{(1,1)} = 0$, yielding $S_i(\hat{\theta}_i(\underline{z}_i)) = a_i$ for each $i = 1, 2$.[20] Lastly, consider $\phi_{\eta,i}^B(\underline{z}_\eta) > 0$. Then by optimality the designer assigns $p_i^{\omega_i} = 1$, yielding $S_i(\hat{\theta}_i(\underline{z}_i)) = a_i$ for each $i = 1, 2$.[21] □

D Proof of Theorem 2

Fix some $\theta \in \Theta$. Suppose first that $\phi_{\eta,i}\Big(\theta_i, z_i(\hat{\theta}_i)\Big) > 0$ for some $i \in \mathcal{N}$. It then follows that for every $\omega \in \Omega$, there exists $i_\omega \in \mathcal{N}$ such that $\omega \in E_{i_\omega}^z(\theta)$. Then, given x maximizes (41) pointwise, for every $i \in \mathcal{N}_0$ and $\omega \in \Omega$ it must be that $x_i^\omega(\theta) = 0$ for every $i \in \mathcal{N}$ such that $i \neq i_\omega$.

Now, consider $i \in \mathcal{N}$ such that $E_i^z(\theta)$ and has positive measure. Then, to determine the (pointwise) optimal ads function $x_i^\omega(\theta)$ on the set of CTRs $E_i^z(\theta)$, using the fact that x must

[20] Note the multiplicity in optimal tie-breaking rules and how the restriction to EP scoring mechanisms makes the exposition simpler.

[21] For example, in this case ties ($z_1 = z_2$) occur with probability 0, so any $p^{(1,1)}$ works.

satisfy (Marg) and the preceding analysis, one has

$$\mathbb{E}_{\alpha_{\mathcal{N}}}[x_i^{\omega}(\theta)\cdot\mathbb{1}_E(\omega)] = \alpha_{\mathcal{N}}^E, \quad \forall E \subseteq E_i^z(\theta) \text{ s.t. } E \in \mathcal{B}(\Omega). \tag{61}$$

Thus, since the function

$$\mu_{\theta,i}(E) := \mathbb{E}_{\alpha_{\mathcal{N}}}[x_i^{\omega}(\theta)\cdot\mathbb{1}_E(\omega)], \quad \forall E \in \mathcal{B}(E_i^z(\theta)) \tag{62}$$

defines a σ-finite measure on $\mathcal{B}(E_i^z(\theta))$, and $\mu_{\theta,i} \equiv \alpha_{\mathcal{N}}|_{E_i^z(\theta)}$ then by RN theorem the ads function $x_i(\theta)$ is the RN derivative of the two measures, i.e.

$$x_i(\theta) = \frac{d\mu_{i,\theta}}{d\alpha_{\mathcal{N}}|_{E_i^z(\theta)}} \equiv 1 \text{ a.e. on } E_i^z(\theta). \tag{63}$$

Thus, one concludes that

$$x_i^{\omega}(\theta) = \mathbb{1}_{E_i^z(\theta)}(\omega), \quad \forall \omega \in \Omega. \tag{64}$$

Lastly, suppose instead that $\phi_{\eta,i}\Big(\theta_i, z_i(\hat{\theta}_i)\Big) = 0$ for every $i \in \mathcal{N}$, which happens with positive probability if and only if $\phi_{\eta,i}^B(\theta_i) < 0$ *and* $z_i(\hat{\theta}_i) = 0$ for every $i \in \mathcal{N}$. In that case, the designer chooses any x subject to $\hat{\theta} \in \hat{\Theta}(x)$. In particular, setting $x \equiv x^{RN}$ and $t \equiv 0$ satisfies the requirement.

E Proofs: Example 2

Suppose $z_i > 0$. Then

$$\mathbb{P}_{F_j}\left(\phi_{\eta,j}(\theta_j, z_j) < z_i\frac{\omega_i}{\omega_j}\right) = \mathbb{P}\left(\phi_{\eta,j}^S(\theta_j) < \max\left(\eta_w, z_i\frac{\omega_i}{\omega_j}\right)\right)\mathbb{1}_{\left\{\frac{\omega_j}{\omega_i} > \frac{z_i}{z_j}\right\}} + \mathbb{P}\left(\phi_{\eta,j}^B(\theta_j) < z_i\frac{\omega_i}{\omega_j}\right)\mathbb{1}_{\left\{\frac{\omega_j}{\omega_i} < \frac{z_i}{z_j}\right\}}.$$

Given $\theta_i \sim \text{Uniform}[0,1]$, we can then write

$$\mathbb{P}_{F_j}\left(\phi_{\eta,j}(\theta_j, z_j) < \frac{\omega_i}{\omega_j}z_i\right) = \frac{\max\left(\eta_w, z_i\frac{\omega_i}{\omega_j}\right) - \eta_w}{\eta_v + 2\eta_r}\mathbb{1}_{\left\{\frac{\omega_j}{\omega_i} > \frac{z_i}{z_j}\right\}} + \frac{z_i\frac{\omega_i}{\omega_j} - \eta_w + \eta_r}{\eta_v + 2\eta_r}\mathbb{1}_{\left\{\frac{\omega_j}{\omega_i} < \frac{z_i}{z_j}\right\}}$$

Suppose in addition that $\eta_w = 0$. Then, solving for the LHS in (Opt-z) yields

$$\mathbb{E}_{\alpha_{\mathcal{N}}}\left[\omega_i\mathbb{P}_{F_j}\left(\phi_{\eta,j}(\theta_j, z_j) < \frac{\omega_i}{\omega_j}z_i\right)\right] = \mathbb{E}_{\alpha_{\mathcal{N}}}\left[\omega_i\frac{z_i\frac{\omega_i}{\omega_j}}{\eta_v + 2\eta_r} + \omega_i\frac{\eta_r}{\eta_v + 2\eta_r}\mathbb{1}_{\left\{z_j \le z_i\frac{\omega_i}{\omega_j}\right\}}\right] \tag{65}$$

$$= \left(z_i\mathbb{E}_{\omega_i}\left[(\omega_i)^2\right]\mathbb{E}_{\omega_j}\left[1/\omega_j\right] + \eta_r\mathbb{E}_{\alpha_{\mathcal{N}}}\left[\omega_i\mathbb{1}_{\{\omega_j z_j \le z_i\omega_i\}}\right]\right)/\left(\eta_v + 2\eta_r\right). \tag{66}$$

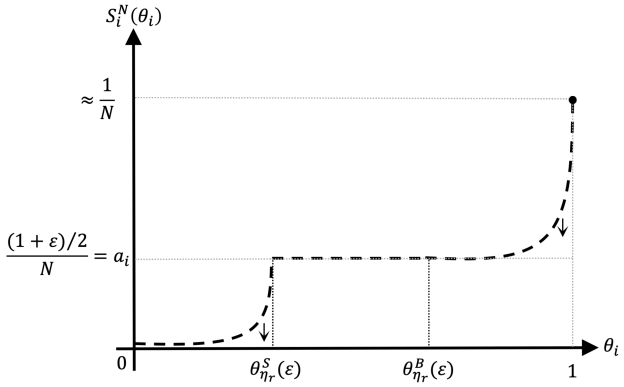


Figure 7: Expected Clicks as $N \to \infty$.

F Proof Sketch for Proposition 2

For sufficiently large N the expected clicks S_i are given by (see Figure 7)

$$S_i^N(\theta_i) \approx \begin{cases} 0 & \text{if } \theta_i \in [0, \theta_{\eta_r}^S) \\ a_i & \text{if } \theta_i \in [\theta_{\eta_r}^S, 1) \\ \frac{1}{N} & \text{if } \theta_i = 1. \end{cases}$$

Then computing transfers $T_i(\theta_i)$ yields approximately

$$T_i^N(\theta_i) \approx \begin{cases} -\theta_{\eta_r}^S \times \frac{\mathbb{E}[\omega_i]}{N} & \text{if } \theta_i \in [0, \theta_{\eta_r}^S) \\ 0 & \text{if } \theta_i \in [\theta_{\eta_r}^S, 1) \\ 1 \times \left(\frac{1}{N} - \frac{\mathbb{E}[\omega_i]}{N}\right) & \text{if } \theta_i = 1. \end{cases}$$

Thus taking the limit the expected revenues are given by (informal)

$$\lim_{N\to\infty} \mathbb{E}\left[\sum_{i=1}^N T_i(\theta_i)\right] = -\theta_{\eta_r}^S(\varepsilon)\mathbb{E}[\omega_i]F\left(\theta_{\eta_r}^S(\varepsilon)\right) + 1 \times \left(F\left(\theta_{\eta_r}^S(\varepsilon)\right) + (1 - \mathbb{E}[\omega_i])\left(1 - F\left(\theta_{\eta_r}^S(\varepsilon)\right)\right)\right).$$

G Proofs: Example 3

For $\theta_i \sim \text{Uniform}[0,1]$, using derivations in (65) and $\eta_w = 0$, equation (45) reduces to:

$$\mathbb{E}_{\alpha_1}\left[\omega_1\left(\mathbb{E}_{\alpha_2}\left[z\frac{\omega_1}{\omega_2} + \eta_r \mathbb{1}_{\{\omega_2 \le \omega_1\}}\middle|\omega_1\right]\right)^{N-1}\right] = (1+\eta_r)^{N-1}\frac{\mathbb{E}_{\alpha_1}[\omega_1]}{N} \tag{67}$$

$$\Leftrightarrow \quad \mathbb{E}_{\alpha_1}\left[\omega_1\left(\omega_1\left(z\ln(1/\varepsilon) + \eta_r\right) - \eta_r\varepsilon\right)^{N-1}\right] = ((1-\varepsilon)(1+\eta_r))^{N-1}\frac{\mathbb{E}_{\alpha_1}[\omega_1]}{N} \tag{68}$$

Now, after some algebraic manipulations

$$E[\omega_1(a\omega_1 - b)^{N-1}] = \frac{1}{a}\mathbb{E}[a\omega_1(a\omega_1 - b)^{N-1}] = \frac{1}{a}\left(\mathbb{E}[(a\omega_1 - b)^N] + b\mathbb{E}[(a\omega_1 - b)^{N-1}]\right) \tag{69}$$

where $a(z,\varepsilon,\eta) = z\ln(1/\varepsilon) + \eta_r > 0$ and $b(\varepsilon,\eta) = \eta_r\varepsilon > 0$. Thus we need to solve

$$\frac{1}{a}\Big(\mathbb{E}[(a\omega_1 - b)^N] + b\mathbb{E}[(a\omega_1 - b)^{N-1}]\Big) = ((1-\varepsilon)(1+\eta_r))^{N-1}\left(\frac{1+\varepsilon}{2N}\right) \tag{70}$$

where since $\omega_1 \sim \text{Uniform}[\varepsilon, 1]$ the random variable $a\omega_1 - b$ is distributed according to $\text{Uniform}(a\varepsilon - b, a - b)$ with $a\varepsilon - b > 0$. Thus it's n^{th} moment is given by

$$\mathbb{E}[(a\omega_1 - b)^n] = \frac{(a-b)^{n+1} - (a\varepsilon - b)^{n+1}}{(n+1)(1-\varepsilon)a}. \tag{71}$$

Thus, after some algebraic simplifications one gets (LHS) and (RHS).