Essays on Fair Operations

Shangzhou Xia

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2024

Abstract

Essays on Fair Operations

Shangzhou Xia

Fairness emerges as a vital concern to decision makers as crucial as efficiency, if not more important. Fair operations decisions are aimed at distributive justice in various scenarios. In this dissertation, we study two examples of distributively fair decision making in operations research, a dynamic fair allocation problem and a subpopulational robustness assessment problem for machine learning models.

We first study a dynamic allocation problem in which T sequentially arriving divisible resources are to be allocated to a number of agents with concave utilities. The joint utility functions of each resource to the agents are drawn stochastically from a known joint distribution, independently and identically across time, and the central planner makes immediate and irrevocable allocation decisions. Most works on dynamic resource allocation aim to maximize the utilitarian welfare, i.e., the efficiency of the allocation, which may result in unfair concentration of resources on certain high-utility agents while leaving others' demands under-fulfilled. In this work, aiming at balancing efficiency and fairness, we instead consider a broad collection of welfare metrics, the Hölder means, which includes the Nash social welfare and the egalitarian welfare. To this end, we first study a fluid-based policy derived from a deterministic surrogate to the underlying problem and show that for all smooth Hölder mean welfare metrics it attains an $O(1)$ regret over the time horizon length T against the hindsight optimum, i.e., the optimal welfare if all utilities were known in advance of deciding on allocations. However, when evaluated under the non-smooth

egalitarian welfare, the fluid-based policy attains a regret of order $\Theta(\sqrt{T})$. We then propose a new policy built thereupon, called Backward Infrequent Re-solving (BIR), which consists of re-solving the deterministic surrogate problem at most $O(\log T)$ times. We show under a mild regularity condition that it attains a regret against the hindsight optimal egalitarian welfare of order $O(1)$ when all agents have linear utilities and $O(\log T)$ otherwise. We further propose the Backward Infrequent Re-solving with Thresholding (BIRT) policy, which enhances the BIR policy by thresholding adjustments and performs similarly without any assumption whatsoever. More specifically, we prove the BIRT policy attains an $O(1)$ regret independently of the horizon length T when all agents have linear utilities and $O(\log^{2+\varepsilon} T)$ otherwise. We conclude by presenting numerical experiments to corroborate our theoretical claims and to illustrate the significant performance improvement against several benchmark policies.

The performance of ML models degrades when the training population is different from that seen under operation. Towards assessing distributional robustness, we study the worst-case performance of a model over *all* subpopulations of a given size, defined with respect to core attributes Z. This notion of robustness can consider arbitrary (continuous) attributes Z, and automatically accounts for complex intersectionality in disadvantaged groups. We develop a scalable yet principled two-stage estimation procedure that can evaluate the robustness of state-of-the-art models. We prove that our procedure enjoys several finite-sample convergence guarantees, including *dimension-free* convergence. Instead of overly conservative notions based on Rademacher complexities, our evaluation error depends on the dimension of Z only through the out-of-sample error in estimating the performance conditional on Z. On real datasets, we demonstrate that our method certifies the robustness of a model and prevents deployment of unreliable models.

Table of Contents

List of Figures

List of Tables

Acknowledgements

Without the slightest hesitation I decided five years ago to remain at Columbia to continue a doctoral journey at Decision, Risk and Operations, and yet the journey has still never failed to outbid my expectations. I have been intrigued by the research work developed and communicated here, enriched by the knowledge taught here, and impressed by the minds and hearts of the family residing here. Nonetheless, the journey was not without hurdles—there have been times of frustration, perplexity, anxiety and disconnection, but I was fortunate enough to pull through with helping hands from many. This dissertation, alongside my doctoral studies, would not have been possible without them.

I would like to express my deepest gratitude to my advisor, Santiago R. Balseiro, under whose guidance it has been a great honor and fortune for me to study and to research. He does not cease to impress me with his intellectual acuteness and scholarly talent. He thinks with the big picture in mind but remains attentive to details with utmost rigor. At times of bewilderedness he steps in with inspiring ideas and insights—he is always my most reliable recourse for the unsolvable or undecipherable. I also appreciate his very kind support in my job search and internships, in addition to his generosity in spending time meeting and discussing research ideas with me and continual encouragement in the difficult times. I am deeply grateful for his mentorship.

I am thankful for the rewarding and fruitful experience working with Hongseok Namkoong, who impresses me with strong work ethic and remarkable productivity, alongside his patience and support both in my research and beyond. I would like to thank Yash Kanoria, who provided me immense support in not only early-year studies but the very first research experience, especially in the challenging times in the height of the pandemic. I am also grateful for Ciamac Moallemi and Jing Dong, who helped me onboard the doctoral program and navigate the coursework plans. I am thankful for my dissertation committee members, Omar Besbes, Henry Lam, and Will Ma, for offering their precious perspectives in my research. It has been a pleasure and learning experience working as a teaching assistant for Evan Picoult, whose kindly shared interesting stories I truly enjoyed as well.

I am very grateful to Henry Lam, Dylan Possamaï and Clifford Stein, for sowing the seeds of operations research during my undergraduate years in the Industrial Engineering and Operations Research program. Each one of them provided a vastly different research perspective, offering me a comprehensive understanding of the field, but all were immensely generous in sharing their knowledge and experience and supporting my applications to doctoral programs.

I must acknowledge the instructional, financial and research support of Columbia Business School and the Decision, Risk and Operations division. I thank in particular the wonderful staff including Elizabeth Elam, Dan Spacher, Winnie Leung, Cristina Melo-Moya, Andrew Cassidy, Clara Magram, Kira Grant, Maria Micheles, Samantha Valente.

I benefited so much from the friendship with student and alumni of the DRO program as well as my friends in the past five years. I am thankful to Pengyu Qian, Yilun Chen, Jiaqi Lu, Zhe Liu, Yue Hu, Judy Gan, Yunbei Xu, Tony Zhang, Steven Yin and Jason Liang for their perspectives in studies and research. I enjoy the many discussions and fun I had with fellow students, including Mike Li, with whom I collaborated on a work now part of this dissertation, Omar Mouchtaki, Utkarsh Patange, Akshit Kumar, Wen Yun, Wenxin Zhang, Ari Boyarsky, Zhen Huang, Yibo Zeng and many more. I am also appreciative of my love Yixuan She, a great companion and supporter through happiness and hardships.

Lastly, I am profoundly grateful to my parents, whose unwavering and unconditional love supported me throughout this tortuous journey. They have taught me great values from their very own lives, guided me through my childhood and adolescence, and continued to serve as my

warmest and safest harbor. The pandemic separated us for more than three years in the midst of my doctoral program, but I never felt an absence of everlasting strength supplied by and transmitted from them. My profound indebtedness to them outweighs any expression.

Chapter 1: Dynamic Fair Allocation

1.1 Preliminaries

1.1.1 Introduction

Fair resource allocation has been widely studied in economics and operations research, with applications in school seat allocation [1, 2], work visa lotteries [3], affordable housing allocation [4, 5], machine load balancing [6, 7], etc. Involving multiple parties stipulates fairness as a crucial concern to the central planner. To this end, allocations are often evaluated on a balance between efficiency and fairness across the agents [8, 9, 10, 11, 12]. Common fairness criteria studied in the literature of fair allocation include Pareto efficiency [13], envy-freeness [2] along with its variants, and proportional fairness [14, 15].

Cardinal welfare metrics incorporate efficiency and fairness concerns of allocations by mapping agents' individual utilities to a real-valued numeric representation of collective welfare. These metrics are wieldy in operations research and computer science because they allow to cast allocation problems as single-objective optimization problems, and different allocations may be compared against one another, yielding a quantifiable welfare difference. One particular cardinal fairness metric is the egalitarian welfare, based on the theory of distributive justice by Rawls [16, 17], whose difference principle maximizes the welfare of the worst-off group of agents. The egalitarian welfare objective defines a max-min welfare optimization problem, whose indivisible variant is dubbed the Santa Claus problem [18, 19, 20]. Another well-studied cardinal example is the Nash social welfare, mathematically the geometric mean of the agents' utilities [21, 22, 23, 24, 25, 26]. Optimal allocations under the Nash social welfare are invariant to scaling of any single agent, for it is homogeneous on each agent's utility. In this work we consider a broad parameterized collection of *Hölder mean* welfare metrics [27], which subsumes the two examples above.

Allocation problems in real life, however, often concern dynamic and irrevocable decisions where either agents or resources arrive in a sequential manner [28, 29, 30]. Some literature has referred to this variant as the online rationing problem, with applications in allocating computational resources to cloud users [31, 32, 33] and most recently in medical equipment and vaccine rationing during the COVID-19 pandemic [34, 35, 36].

This work focuses on the specific variant of allocating resources arriving sequentially to heterogeneous agents, which is particularly relevant when resources are perishable or urgently needed and require immediate allocation decisions. Applications include allocating donated organs to hospitals [37], donated food to local charities [38, 39], and online advertisement slots to advertisers [40, 41, 42, 43, 44, 45], etc. Our objective is to design computationally efficient dynamic policies that maximize the overall welfare.

Contribution

In this work, we consider the dynamic fair allocation problem, where T resources arriving sequentially are to be immediately and irrevocably allocated to n agents. The joint concave utility functions of each resource to the agents are drawn independently from a joint distribution that is known to the central planner, and are revealed only before allocation decisions are made. We are interested in developing non-anticipative policies with low regret (i.e., expected welfare loss) against the hindsight optimum, the optimal welfare if resources were known in advance. An optimal online policy could in principle be computed using dynamic programming but, because of the so-called *curse of dimensionality*, solving the problem to optimality is impractical when the horizon is large. We therefore seek to design policies that are computationally efficient and have provable performance guarantees.

We first consider a type-based fluid policy that is obtained by solving a fluid relaxation of the optimization problem in which all stochastic quantities are replaced by their mean values. Simple as it is, we show if the Hölder-mean welfare metric is strongly smooth, it attains a resounding $O(1)$ regret over any T against the hindsight optimum. The sole non-smooth exception is the egalitarian welfare objective, under which a fluid policy can attain a regret of order as large as $\Theta(\sqrt{T})$ and hence more complicated dynamic policies are warranted. Refer to Section 1.2 for rigorous definitions of the fluid relaxation problem, its optimum FLU, and fluid policies F. We summarize findings on the benchmarks (FLU and OPT) and the performance of the fluid policy F on the vertical axis in Figure 1.1 with respective suprema of the differences (i.e., worst-case arrival distribution) between them.

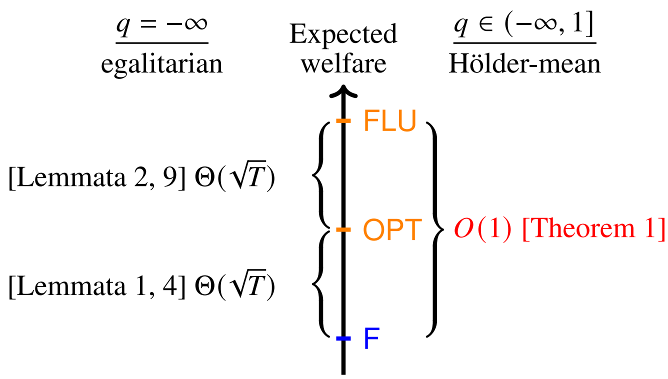


Figure 1.1: Summary of the fluid and hindsight optimum benchmark values and the fluid policy performance. Results relevant to the egalitarian welfare are shown on the left, and those relevant to all other Hölder-mean welfare metrics is shown on the right.

Under the egalitarian welfare, we propose the **Backward Infrequent Re-solving** (BIR) policy, a variant of the fluid policy that re-solves the updated fluid problem at certain epochs throughout the horizon. The BIR policy is inspired by Jasin and Kumar [46], who introduce an infrequent re-solving policy for revenue management problems with utilitarian objectives and additive rewards. The re-solving epochs $(t_k : k \in [K])$ of the BIR policy are recursively determined by a flexible contracting scheduling function f such that $T - t_k = f(T - t_{k-1})$; the larger f, the more frequently BIR re-solves the fluid problem. We show under a mild regularity condition on the initial fluid policy, BIR attains a low asymptotic regret of $O(1)$ when all utilities are linear and $O(\log T)$ otherwise. The proof of the performance guarantee is inspired by Balseiro *et al.* [47], who analyze their frequent re-solving policy allowing for non-linear utilities, yet we lift the ideas to more generality and flexibility in re-solving frequencies. The BIR policy is simple and intuitive, as it re-solves the fluid problem to incorporate past information into future decisions.

Our last main contribution is proposing the **Backward Infrequent Re-solving with Thresholding** (BIRT) policy for dynamic fair allocation under the egalitarian welfare by enhancing the re-solving policy with a thresholding adjustment. Our policy is inspired by Bumpensanti and Wang [48], who design similar policies for utilitarian objectives and linear utilities based on fluid policies, and lifts it to much more general re-solving schedules through the design of the scheduling function f and more general thresholding adjustments through the choice of threshold γ. We provide a novel analysis to show that against the hindsight optimum, the BIRT policy can achieve a low asymptotic regret of $O(1)$ when all utilities are linear and $O(\log^{2+\varepsilon} T)$ otherwise. The low regret guarantee implies that the BIRT policy performs comparably to a clairvoyant who foresees all arrivals in advance of acting.

A highlight of the BIRT policy is that its performance guarantee holds without any additional assumption on the problem primitives, including in particular the regularity condition needed for the BIR policy's performance guarantee. In other words, the $O(1)$ or $O(\log^{2+\varepsilon} T)$ regret guarantee of BIRT is insensitive to whether or not the initial fluid policy is regular (a.k.a. non-degenerate in the case of linear programs, when all agents have linear utilities) or nearly irregular. This is in drastic contrast to previous works on online stochastic optimization whose similar $O(1)$ or $\tilde{O}(\log T)$ loss results [47, 49] rely heavily on regularity or non-degeneracy conditions on the underlying fluid problem (see Section 1.4 for a numerical illustration).

We summarize in Figure 1.2 our performance guarantees under the egalitarian welfare for the BIR and BIRT policies parameterized by the scheduling function f that determines the re-solving frequency; the larger f, the more frequently re-solving occurs. Results are shown separately for linear and nonlinear utility functions. Note the guarantee for the BIR policy is dependent on the regularity condition on the initial fluid solution (Assumption 2).

All in all, we provide computationally efficient online policies that attain either uniformly bounded or logarithmically low regret for all Hölder-mean welfare metrics.

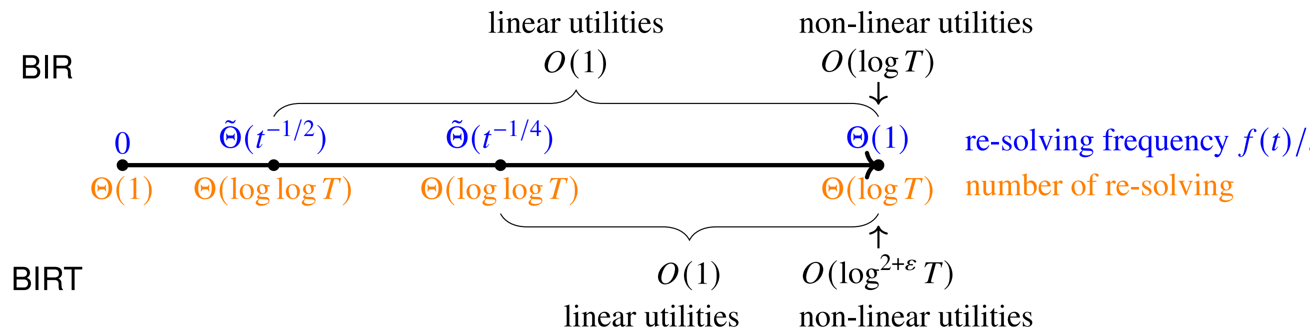


Figure 1.2: Summary of regret guarantees for the BIR and BIRT policies under the egalitarian welfare. The horizontal axis indicates the re-solving frequency determined by the scheduling function f, where the larger f, the more frequently re-solving occurs. Results are given separately for linear and non-linear utility functions. The guarantee on the BIR policy is based on the fluid regularity assumption 2.

1.1.2 Related work

Our work contributes to the literature on fair resource allocation and, more generally, to the literature on dynamic resource allocation problems.

Fair allocation There is a stream of literature studying the online version of the fair allocation problem in which utilities are linear but the arrival process in unknown to the central planner. Devanur *et al.* [34] consider the egalitarian objective when the agents' utilities are drawn independently from a distribution that is unknown to the central planner (the "unknown IID" model) and design an algorithm that attains $O(\sqrt{T})$ regret. Agrawal and Devanur [50] consider general online allocation problems with concave non-separable objectives, which subsume our problem with Hölder-mean welfare metrics, and propose an algorithm based on multiplicative weight updates method [51] that also attains $O(\sqrt{T})$ regret. Balseiro *et al.* [52] study similar general online allocation with hard resource constraints and propose fast algorithms based on mirror descent that do not require solving auxiliary optimization problems. More recently, Kawase and Sumita [53] consider the max-min fair allocation of indivisible resources arriving online to heterogeneous agents. They propose a deterministic policy that achieves an asymptotic approximation ratio of n in the adversarial arrival model, which they show is optimal. In the case of unknown IID arrivals, they give a policy that achieves $O(\sqrt{T})$ regret against the hindsight optimum benchmark, based on the multiplicative weights updates [51]. Most recently, Freund and Banerjee [54] studies a particular variant with a discrete action space where the final objective depends only on the final type-action counts, regardless of the order and timing of decisions; they propose an empirical re-solving policy based on thresholding type-action counts, which they show attains a remarkable $O(1)$ regret against the hindsight optimum under a Lipschitz continuity assumption on the offline problem (Definition 5).

In contrast to the previous line of work, we consider the case where the distribution of utilities is known to the central planner. When all agents have linear utilities, the $O(1)$ regret bound of the BIRT policy in the known IID model matches [54] and significantly improves upon the unknown IID model, in which the best previous known attainable regret was $O(\sqrt{T})$ [50, 34, 53, 52]. The drastic performance improvement shows the great use of distributional knowledge. This is similar in spirit to Banerjee *et al.* [55], who show in maximizing Nash social welfare, agents' utility predictions significantly improve the approximation ratio of online algorithms in adversarial settings to $O(\log(n \wedge T))$ from a trivial $\Theta(n)$. Other works on maximizing Nash social welfare include Freeman *et al.* [56] and Sinclair *et al.* [57].

Under the egalitarian welfare, the particular case of identical agents has been studied for decades under the name of the "online machine covering" problem in the context of scheduling [58, 59, 60, 61]. Moreover, there is a line of work studying an offline version of the max-min fair allocation problem (the case where all utilities are known in advance), which is referred to as the Santa Claus problem [62, 63, 64, 65, 66, 67, 20, 68].

In this paper, we study an *online supply* version of the dynamic fair allocation problem in which resources arrive online and the population of agents are fixed. The *online demand* version of the problem in which resources are fixed and agents arrive online has also been studied in the literature [28, 69, 70, 39].

Finally, there is a long line of literature on online allocation focusing on other fairness notions such as envy [12, 71], envy-freeness (up to 1 item) [72, 38], Pareto efficiency [12] and more general concave returns [73]. We recommend to the reader the survey by Aleksandrov and Walsh [74] for an overview of the literature of variants of online fair allocation problems.

Dynamic resource allocation Our work is also closely related methodologically to the stream of literature on dynamic resource allocation, especially revenue management, that aims at maximizing revenue or efficiency of an allocation given limited resources. Our policy is based on re-solving the fluid problem, a specific form of certainty equivalent of dynamic stochastic optimization problems when the problem instances are independent and identically distributed according to a distribution known to the central planner. The main differences with this line of work are that the Hölder-mean welfare objective we consider is not time separable, and that we allow for non-linear concave utilities, calling for different theoretical analyses.

The seminal works by Gallego and Van Ryzin [75] and Talluri and Van Ryzin [76] show fluid policies based on solving the fluid problem only once at the beginning suffice to achieve an $O(\sqrt{T})$ regret against the hindsight optimum. Reiman and Wang [49] propose a policy based on re-solving the fluid policy exactly once only incurring an $o(\sqrt{T})$ asymptotic revenue loss. Jasin and Kumar [46] show re-solving multiple times or even every period can give an $O(1)$ loss under the assumption that the fluid problem admits a non-degenerate optimal solution, which inspires our BIR policy and analysis. Bumpensanti and Wang [48] propose the Infrequent Re-solving with Thresholding policy for revenue management problems with additive and linear objectives, which incurs an $O(1)$ loss, and dispenses with the non-degeneracy condition. Our BIRT policy is based thereupon. Most recently, Vera and Banerjee [77] and Vera *et al.* [78] propose a general Bayes selector policy that attains low regret for a number of dynamic allocation problems. A specific variant called the multi-secretary problem also witnessed policies incurring $O(1)$ [79] or $O(\log T)$ [80] regret. We recommend to the reader a survey by Balseiro *et al.* [47] for an overview of dynamic allocation problems and a unifying analysis of re-solving policies based on certainty equivalents that allow for non-linear rewards.

1.1.3 Notation

We denote the extended real line by $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$. For $m \in \mathbb{N}$, we denote $[m] := \{1, 2, \ldots, m\}$, the probability m-simplex by $\Delta_m := \{x \in \mathbb{R}_+^m : \|x\|_1 = 1\}$, a column m-vector of ones by $\mathbf{1}_m$ and a column m-vector of zeros by $\mathbf{0}_m$. For $w \in \mathbb{R}$, we denote by $\lfloor w \rfloor := \max\{m \in \mathbb{Z} : m \leq w\}$ the largest integer smaller than or equal to w, and by $\lceil w \rceil := \min\{m \in \mathbb{Z} : m \geq w\}$ the smallest integer larger than or equal to w. We use boldface alphabets (e.g., $\boldsymbol{x}$) to denote matrices. For any random vector v, denote $\Delta v := v - \mathbb{E}[v]$. For any two positive-valued functions $f, g : \mathbb{N} \to \mathbb{R}_+$, we denote $f(T) = o(g(T))$ if $\lim_{T\to\infty} f(T)/g(T) = 0$; we denote $f(T) = O(g(T))$ if $\limsup_{T\to\infty} f(T)/g(T) < \infty$; we denote $f(T) = \Omega(g(T))$ if $\liminf_{T\to\infty} f(T)/g(T) > 0$; we denote $f(T) = \Theta(g(T))$ if $f(T) = \Omega(g(T))$ and $f(T) = O(g(T))$. With a slight abuse of notation, we sometimes denote optimization problems, their optima and optimal solutions interchangeably.

1.1.4 Problem Formulation

In this work, we consider the problem of dynamically allocating resources to n heterogeneous agents with linear utilities that are stochastically distributed. More formally, suppose each agent $i \in [n]$ starts with zero initial utility $U_0^i = 0$ to accumulates per-period utility additively while receiving allocated resources over time. The central planner receives a sequence of $T > 1$ divisible resources, each equipped with a revealed vector u_t of utility functions $u_t^i : \mathbb{R}_+ \to \mathbb{R}_+$ indicating its joint utility functions to the n agents, and immediately decides on its proportional allocation $x_t \in \Delta_n$ to the agents. In the sequel, we refer to the t-th resource and the t-th period interchangeably, and to resources and arrivals interchangeably. By the end of the horizon, agent i would have a cumulative utility of $U_T^i = \sum_{t=1}^T u_t^i(x_t^i)$. At times in the sequel, we denote by $y_t^i := u_t^i(x_t^i)$ the utility that agent i obtains from resource t.

We evaluate allocation decisions on their balance between efficiency and fairness through a cardinal welfare metric $w : \mathbb{R}_+^n \to \mathbb{R}_+$ of the vector of agents' cumulative utilities $U_T \in \mathbb{R}_+^n$. We denote the welfare generated by a sequence of allocation decisions $\boldsymbol{x} \equiv (x_t : t \in [T])$ according to policy π by

$$\mathsf{ALG}(\pi) := w\left(\sum_{t=1}^{T} u_t^i(x_t^i)\right). \tag{1.1}$$

An *online policy* is non-anticipative, or in other words, decides on allocations only based on the history observed so far. More formally, define the natural filtration $\mathcal{F} = (\mathcal{F}_t : t \in [T])$ given by $\mathcal{F}_t := \sigma(u_1, \dots, u_t)$, the σ-algebra generated by the first t utility function vectors. An online policy allocates $x_t \in \mathcal{F}_t$ for all $t \in [T]$, namely the allocation process $(x_t : t \in [T])$ is adapted to the natural filtration $\mathcal{F}$. Notice we assume that online policies are deterministic for the purpose of evaluating the welfare they generated—this is without loss of optimality because deterministic online policies weakly dominate their randomized counterparts (Lemma 24).

A conventional benchmark used to evaluate online policies is the *hindsight optimum*, defined as the maximum welfare (1.1) if allocations $\boldsymbol{x}$ were allowed to be anticipative, i.e., all vectors of joint utility functions $(u_t^i : t \in [T])$ were known in advance. We denote by OPT the hindsight optimum, as well as any such optimal policy with a slight abuse of notation. Formally, the hindsight optimum is the optimal value of the maximization problem

$$\mathsf{OPT} := \max\left\{w\left(\sum_{t=1}^{T} u_t^i(x_t^i)\right) : \boldsymbol{x} \in \Delta_n^T\right\}. \tag{1.2}$$

Welfare metric In this work, we evaluate allocations using cardinal welfare metrics $w : \mathbb{R}_+^n \to \mathbb{R}_+$, i.e., real-valued functions of the agents' cumulative utilities $U \in \mathbb{R}_+^n$, which allow to balance between efficiency and fairness concerns. Hence, allocation decisions can be directly compared against one another, and their difference indicates a quantifiable welfare loss.

We next enumerate three cardinal welfare examples and briefly discuss the intuition behind them. The *utilitarian welfare* $\sum_{i\in[n]} U^i$ neglects any fairness concerns; optimizing it leads to efficient allocations that could and most often lead to highly unfair concentration of resources at certain agents or groups. The *egalitarian welfare* $\min_{i\in[n]} U^i$ is the other extreme—defined as the utility of the worst-off agent, this metric provides all agents with a uniform minimal utility guarantee. The *Nash social welfare* $(\prod_{i=1}^n U^i)^{1/n}$ is a more balanced metric in terms of the trade-off between efficiency and fairness; in addition, its geometric mean form satisfies positive homogeneity in *each agent's utility*, which insulates the central planner against agents' possible strategic reporting of utility functions.

Formally, we consider a collection of welfare metrics that subsumes and generalizes all the examples above, the *Hölder means* $w_q : \mathbb{R}_+^n \to \mathbb{R}_+$, a.k.a. *generalized power means*, parameterized by $q \in \overline{\mathbb{R}}$, defined as follows and illustrated in the fairness spectrum shown in Figure 1.3. In this problem, we stipulate $q \in [-\infty, 1]$ because w_q with $q \in (1, +\infty]$ is not concave and hence deemed "unfair" as per Lemma 1.

- For $q = -\infty$, $w_q(U) = \min_i U^i$;
- for $q \in (-\infty, 0)$, $w_q(U) = (\frac{1}{n} \sum_{i \in [n]} (U^i)^q)^{1/q}$ if $B > 0$ and $w_q(U) = 0$ otherwise;
- for $q = 0$, $w_q(U) = (\prod_{i=1}^n U^i)^{1/n}$;
- for $q \in (0, +\infty)$, $w_q(U) = (\frac{1}{n} \sum_{i \in [n]} (U^i)^q)^{1/q}$;
- for $q = +\infty$, $w_q(U) = \max_i U^i$.

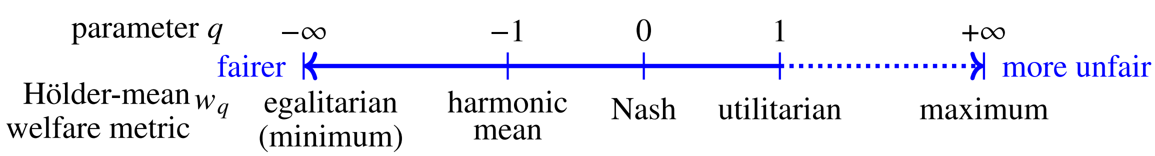


Figure 1.3: Fairness spectrum of the Hölder-mean welfare metrics w_q parameterized by $q \in \overline{\mathbb{R}}$. The fairness of the welfare metric decreases as q increases. Dashed line indicates the "unfair" regime ($q > 1$).

We consider the broad collection of Hölder mean welfare metrics w_q because they satisfy a number of axioms that are reasonable in common scenarios of fair allocation; moreover, classical results have shown that they are the only welfare metrics satisfying these axioms, up to a multiplicative constant, which we have disposed of through normalization of w_q in the definition above. Hence, $w_q(u\mathbf{1}) = u$ for any scalar $u \geq 0$, i,e., the welfare equals each agent's utility if all utilities are identical.

Lemma 1 ([81, 82]). *For any welfare metric* $w : \mathbb{R}_+^n \to \mathbb{R}$, $w \propto w_q$ *for some* $q \in [-\infty, 1]$ *if and only if* w *satisfies the following axioms.*

1. *Monotonicity:* $w(U) \geq w(\hat{U})$ *if* $U \geq \hat{U}$.

2. *Symmetry:* $w(U) = w(\varsigma(U))$ *for any* n*-permutation* $\varsigma \in \mathbb{S}_n$.

3. *Continuity:* w *is continuous.*

4. *Independence of unconcerned agents: if* $w(U^1, U^2, \ldots, U^n) \geq w(\hat{U}^1, U^2, \ldots, U^n)$, *then* $w(U^1, \hat{U}^2, \ldots, \hat{U}^n) \geq w(\hat{U}^1, \hat{U}^2, \ldots, \hat{U}^n)$.

5. *Homogeneity:* $w(\lambda U) = \lambda w(U)$ *for* $U \geq 0$ *and* $\lambda > 0$.

6. *Pigou–Dalton principle:* $w(U^1, U^2, U^3, \ldots) \leq w(U^1 + \varepsilon/2, U^2 - \varepsilon/2, U^3, \ldots)$ *for* $\varepsilon \in (0, U^2 - U^1)$ *if* $U^1 < U^2$.

Note the axioms serve as the foundation to fair allocation and are reasonably motivated. More specifically, monotonicity establishes that larger agents' utilities are preferable; symmetry guarantees equal treatment by making the welfare metric independent of the identities of the agents; independence of unconcerned agents determines that an agent's preferences do not change if we modify the utility of other agents; homogeneity guarantees that allocation is independents of the units used to measured the utility; the Pigou-Dalton principle stipulates that the welfare metric prefers fairness to unfairness, i.e., transferring utility from a high-utility to a low-utility individual should increase fairness. By imposing these axioms, we restrict our attention to a class of welfare metrics that are rich enough to incorporate flexibility in fairness concerns and structured enough to work with.

Utility functions Each agent assess utility independently and solely based on their own resources. More technically, each agent i starts with zero initial utility $U_0^i = 0$ and accumulates utility additively over periods. In each period t, the central planner allocates a portion x_t^i of the arriving resource to agent i based on all agents' utility functions $(u_t^j : j \in [n])$, resulting in a utility of

$u_t^i(x_t^i)$ for agent i. In the end, every agent i would accumulate a total utility of $U_T^i = \sum_{t\in[T]} u_t^i(x_t^i)$ after T resources have been allocated. Notice in this setup, utility can be nonlinear in allocated resources within each period, but is additive over time.

We specify certain properties on the agents' utility functions of the proportion of resource they receive in every period. They are defined on $\mathbb{R}_+$, but in practice, the domain of utility functions is effectively $[0, 1]$, as any agent can receive at most the entire resource in one period.

Definition 1. *Fix a universal constant $\kappa_u \geq 0$. A utility function $u : \mathbb{R}_+ \to \mathbb{R}_+$ is* admissible *if it is twice continuously differentiable and satisfies the following properties.*

1. $u(0) = 0$ *and* $u(x) > 0$ *if* $x > 0$.

2. *For any* $x \in \mathbb{R}_+$, $u'(x) \in (0, 1]$.

3. *For any* $x \in \mathbb{R}_+$, $u''(x) \in [-\kappa_u, 0]$.

Each of the properties is reasonably motivated. The first property states any agent obtains positive utility if and only if they received allocated resource; the second property states the utility is strictly monotone and bounds the marginal utility; the last property establishes the economic law of diminishing marginal utility and bounds the diminishment. The following utility function examples are admissible under appropriate constants κ_u, where $\beta \in (0, 1]$.

- $u(x) = \beta x$;

- $u(x) = \sqrt{1 + 2\beta x} - 1$;

- $u(x) = \beta \log(1 + x)$.

Note in particular if $\kappa_u = 0$, then a utility function is admissible if and only if it has the first form, i.e., a non-trivial linear utility.

In the sequel, we overload notation to denote an n-vector of the agents' utility functions by $u : [0, 1]^n \to [0, 1]^n$ given by concatenating individual utility functions $u(x) := (u^i(x^i) : i \in [n])$. We say a utility function vector u is *admissible*, and a resource equipped with utility function vector u is *admissible*, if all of its components u^i are admissible.

Arrival model We consider the IID arrival model. More formally, we assume the utility function vector of each arrival $u_t : [0,1]^n \to [0,1]^n$ is distributed according to a known distribution P, independently across time t. While our model assume that the joint utility functions are IID across time, they may be arbitrarily correlated across agents in a given period, allowing for heterogeneity in agents to the fullest extent.

To better describe the resources, we identify every admissible utility function vector with a *type* $\theta \in \Theta$, where Θ is the set of all types of admissible utility function vectors, so that we can categorize each resource t with admissible utility functions to some type $\theta_t \in \Theta$ and design allocation policies catered to it.

Definition 2. *An* admissible *probability distribution P is such that almost every resource is admissible. Alternatively, its support Θ_P consists only of admissible utility functions, i.e., $\Theta_P \subseteq \Theta$. The class of admissible distributions is denoted by $\mathcal{P}$.*

Note although the cardinality of Θ is infinite, the support Θ_P of an admissible distribution P can be finite; in fact, we mostly focus on distributions with finite support for expositional purposes. For example, the distribution P under which $u^i(x) = \beta^i x$ with $\beta \sim \text{Unif}\{(1, 1/2), (1/2, 1)\}$ has a finite support.

We briefly discuss and introduce some more notation for finite admissible distributions. A finite admissible distribution P has a finite support Θ_P and is given by $P\{\theta\} = p_\theta > 0$ for almost every type $\theta \in \Theta_P$, where $|\Theta_P| < \infty$ is the number of distinct arrival types, and $(p_\theta : \theta \in \Theta_P) \in \Delta_{|\Theta_P|}$ is a probability vector. We denote the minimum probability by $\underline{p} := \min_\theta p_\theta$ for finite distribution P. Since each type in the support has a positive probability mass, we can meaningfully denote the number of type-θ arrivals by $N_\theta := \sum_{t=1}^T \mathbf{1}\{\theta_t = \theta\}$ for $\theta \in \Theta_P$, which thus follows a multinomial distribution, i.e., $(N_\theta : \theta \in \Theta_P) \sim \mathsf{Multinomial}(T; (p_\theta : \theta \in \Theta_P))$ with T trials and success probabilities $(p_\theta : \theta \in \Theta)$.

Under the IID arrival model, we evaluate an online policy using its expected Hölder-mean welfare loss against the hindsight optimum, along with the relative counterpart. More rigorously, given an admissible distribution $P \in \mathcal{P}$, we define the *regret* and *relative regret* of a policy π by

$$\mathcal{R}_T(\pi) := \mathbb{E}_{\theta \overset{\text{i.i.d.}}{\sim} P}\left[\mathsf{OPT} - \mathsf{ALG}(\pi)\right] \quad \text{and} \quad \mathsf{RR}_T(\pi) := \frac{\mathbb{E}_{\theta \overset{\text{i.i.d.}}{\sim} P}\left[\mathsf{OPT} - \mathsf{ALG}(\pi)\right]}{\mathbb{E}_{\theta \overset{\text{i.i.d.}}{\sim} P}[\mathsf{OPT}]}. \tag{1.3}$$

The relative regret is well-defined as $\mathsf{OPT} > 0$ P-a.s. by Lemma 23. In the sequel, the underlying distribution P is omitted when implied in the context. More specifically, given admissible arrival distribution $P \in \mathcal{P}$, we are interested in analyzing the asymptotic regret $\mathcal{R}_T(\pi)$ of policy π as the horizon length T grows.

Lastly, we formally define a particular collection of *type-based policies*, which will play an important role in our analysis. A type-based policy specifies an allocation vector $x_\theta \in \Delta_n$ over agents for each type $\theta \in \Theta$; in response to resource t, the policy allocates accordingly $x_t = x_{\theta_t}$. We also refer to the underlying matrix $\boldsymbol{x} \equiv (x_\theta : \theta \in \Theta) \in \Delta_n^{\Theta_P}$ as a type-based policy. Similarly, we denote by $y_\theta^i := u_\theta^i(x_\theta^i)$ agent i's utility obtained from a type-θ resource under the type-based policy $\boldsymbol{x}$. The technical intuition of a type-based policy is that it allocates incoming resources only depending on their types, i.e., it makes identical allocation decisions for resources of the same type. Under type-based policy $\boldsymbol{x}$, the cumulative utility of agent i is hence $U_T^i = \sum_{t\in[T]} u_{\theta_t}^i(x_{\theta_t}^i)$. It is worth noting that the concavity of admissible utility functions (Definition 1) implies any hindsight optimal policy is weakly dominated by a type-based policy (see Lemma 25).

1.2 Fluid Problem and Fluid Policies

Certainty equivalents have been widely used in general dynamic stochastic optimization problems to design simple heuristics with good performance guarantees [83]. Certainty equivalent policies seek to approximate the offline optimization by replacing stochastic quantities with fixed and known deterministic values. A particular choice is to replace all random quantities by their means. The resulting optimization problem is usually referred to as the *fluid problem*, or the deterministic program. The optimal solution to the fluid problem prescribes a simple type-based policy, which can be implemented in the original stochastic program.

We now formally define the fluid problem. Consider the IID arrival model with distribution $P \in \mathcal{P}$ known to the online central planner. In the fluid problem, the central planner substitutes each arrival utility function by the expected utility function. Note the fluid problem can be viewed a specific form of offline problem, so that Lemma 25 implies type-based policies are optimal for the fluid problem. The fluid problem is hereby defined as

$$\mathsf{FLU} := \text{maximize} \quad \left\{ w\left(T\,\mathbb{E}_\theta\left[u^i_\theta(x^i_\theta)\right]\right) : \boldsymbol{x} \in \Delta_n^{\Theta_P} \right\}. \tag{1.4}$$

In the fluid problem, the central planner chooses a type-based policy that maximizes the welfare induced by the agents' *expected* utilities, where that of agent i under the type-based policy $\boldsymbol{x}$ is $\mathbb{E}[U^i_T] = T\,\mathbb{E}_\theta\left[u^i_\theta(x^i_\theta)\right]$. With an abuse of notation, we also denote by FLU its optimum value, namely the *fluid benchmark*.

An optimal solution to the fluid problem (1.4) is called a *fluid policy*, denoted by $\mathbf{F}$. Fluid policies are classical heuristic policies that serve as a powerful tool for the online central planner who has no other knowledge than the underlying arrival distribution. The pseudocode for fluid policies is given in Algorithm 1. The positive homogeneity of the welfare metric w implies that fluid policies are in fact independent of the horizon length T and that the fluid benchmark FLU is simply a linear function of T.

The classical literature in revenue management have shown that for utilitarian welfare objec-

Algorithm 1 Fluid Policy (F)

Input: horizon length $T \in \mathbb{N}$, arrival distribution $P \in \mathcal{P}$.
Initialize: initial utilities $U_0^i \leftarrow 0$ for $i \in [n]$;
solve the fluid problem with initial utilities:

$$\mathsf{F} \in \arg\max \left\{ w \left(T\, \mathbb{E}_\theta \left[u_\theta^i(x_\theta^i) \right] \right) : \boldsymbol{x} \in \Delta_n^{\Theta_P} \right\}$$

for $t = 1, \ldots, T$ **do** `// implement the fluid policy`
  observe incoming resource type θ_t
  act $x_t \leftarrow \mathsf{F}_{\theta_t}$
  update $U_t^i \leftarrow U_{t-1}^i + u_t^i(x_t^i)$ for every agent $i \in [n]$

Output: allocations $(x_t : t = 1, 2, \ldots, T)$ and welfare $w(U_T)$.

tives, fluid policies incur a regret at most on the order of $\sqrt{T}$ [75, 76], which is tight in the worst case. Here we present similar results in our dynamic fair allocation problem for the fluid benchmark and fluid policies.

Proposition 1. *Fix $P \in \mathcal{P}$. Under Hölder-mean welfare w_q with $q \in [-\infty, 1]$, the regret of any fluid policy F is bounded by $\mathcal{R}_T(\mathsf{F}) = O(\sqrt{T})$.*

Proposition 1 is in fact a corollary of the two following lemmata. Lemma 2 establishes the fluid benchmark as an upper bound of the expected hindsight optimum. The result is standard and follows from Jensen's inequality and the concavity of the welfare metric (see proof in Section A.1.1). Hence, the regret of any policy (e.g., a fluid policy) can be bounded by its expected welfare loss from the fluid benchmark.

Lemma 2. *For any $P \in \mathcal{P}$, under Hölder-mean welfare w_q with $q \in [-\infty, 1]$, $\mathbb{E}[\mathsf{OPT}] \leq \mathsf{FLU}$.*

The next result, Lemma 3, establishes a stochasticity gap at most of order $\sqrt{T}$ between the welfare corresponding to the agents' expected utilities and the expected welfare under any type-based policy, following from a concentration of measure argument. The proof can be found in Section A.1.2. Proposition 1 follows because $\mathsf{FLU} = w(\mathbb{E}[U_T])$ and $\mathsf{ALG}(\mathsf{F}) = w(U_T)$ under the fluid policy F.

Lemma 3. *Fix $P \in \mathcal{P}$. Under Hölder-mean welfare w_q with $q \in [-\infty, 1]$, for based policy $\boldsymbol{x}$, $w_q(\mathbb{E}[U_T]) - \mathbb{E}[w_q(U_T)] = O(\sqrt{T})$.*

Whereas we have established an upper bound on the regret of fluid policies, we proceed to present the drastically different performances they have in distinct regimes of the fairness spectrum (Figure 1.3), $q \in (-\infty, 1]$ and $q = -\infty$. When $q \in (-\infty, 1]$, the upper bound is in fact very loose, as fluid policies are powerful enough to attain an $O(1)$ regret. In particular, the regret is exactly zero under the utilitarian welfare ($q = 1$). The proof can be found in Section A.1.3.

Theorem 1. *Fix $P \in \mathcal{P}$. Under Hölder-mean welfare w_q with $q \in (-\infty, 1]$, any fluid policy* F *attains a regret $\mathcal{R}_T(\mathsf{F}) = O(1)$ uniformly bounded over horizon length T and solely dependent on P.*

Theorem 1 shows the simple fluid type-based policy F can attain an impressive $O(1)$ regret. We next provide the reader with some intuition for this result. A key property of the Hölder-mean welfare w_q for $q \in (-\infty, 1]$ is its local strong smoothness around the expected utilities. More specifically, given any type-based policy $\boldsymbol{x}$, the time-average welfare $w_q(U_T/T)$ can be locally approximated with a quadratic form around the expected utilities $w_q(\mathbb{E}[U_T]/T)$, i.e.,

$$w_q\left(\frac{U_T}{T}\right) - w_q\left(\frac{\mathbb{E}[U_T]}{T}\right) \approx \nabla w_q\left(\frac{\mathbb{E}[U_T]}{T}\right)^{\top} \frac{U_T - \mathbb{E}[U_T]}{T} - O\left(\left\|\frac{U_T - \mathbb{E}[U_T]}{T}\right\|_2^2\right).$$

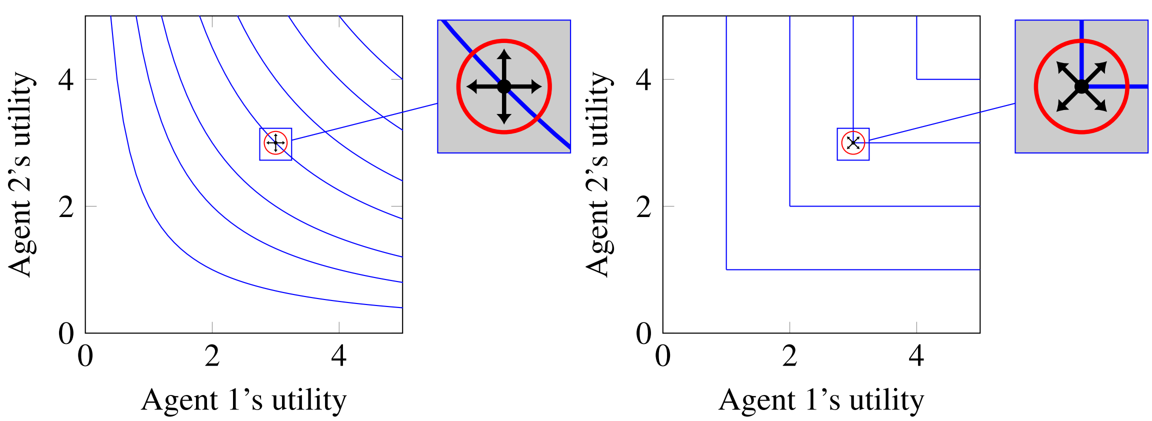
In the approximate quadratic form, the first-order term has zero mean by definition, and the second-order term is $O_p(T^{-1})$ by the Central Limit Theorem, as agent i's utility $U_T^i = \sum_{t\in[T]} y_{\theta_t}$ is a random walk. Note the second-order term is negative because the welfare metric w_q is concave. Because for some welfare metrics (such as the Nash social welfare) strong smoothness only holds locally, a technical step in the proof revolves around showing that utilities fall close to their expected values with a high probability. This implies $\mathbb{E}[w(U_T)] - w(\mathbb{E}[U_T]) = O(1)$, and the theorem statement follows as a result. See the proof in Section A.1.3 for a detailed discussion on local strong smoothness.

However, under the egalitarian welfare $w_{-\infty}$, fluid policies do not perform as well. We show that a regret of order $\sqrt{T}$ given in Proposition 1 is essentially unimprovable for fluid policies under the egalitarian welfare $w_{-\infty}$ by providing a simple example in which the fluid policy does incur an

$\Omega(\sqrt{T})$ regret. See Section A.1.4 for the proof.

Lemma 4. *There exists $P \in \mathcal{P}$ such that $\mathcal{R}_T(\mathsf{F}) = \Omega(\sqrt{T})$ under the egalitarian welfare $w_{-\infty}$.*

Whereas the simple fluid policies attain an $O(1)$ optimality gap under Hölder-mean welfare w_q with $q \in (-\infty, 1]$, they fail to guarantee a bounded regret for the egalitarian welfare $w_{-\infty}$. In contrast to the discussion above for smooth Hölder-mean welfare metrics (i.e., $q \in (-\infty, 1]$), the egalitarian welfare $\min_i U_T^i$ is marked by its lack of smoothness at singular points where agents have equal cumulative utilities ($U_T^1 = \cdots = U_T^n$). Figure 1.4 illustrates the different behavior infinitesimal perturbations around expected utilities $\mathbb{E}[U_T]$ may have under smooth versus non-smooth welfare metrics. In the former, the welfare metric is locally linear, so first-order utility changes resulting from perturbations cancel out in expectation leaving only the second-order changes; in the latter, however, the welfare metric is not differentiable, so first-order utility changes negatively impact performance. Unfortunately, the expected utilities always lie on a singular point, i.e., the vertex of some iso-welfare hyper-surface, because the fluid policy always generate equal expected utilities across agents (see Lemma 6). The fluid policy is then doomed to a first-order welfare loss in expectation and hence a regret on the order of $\Theta(\sqrt{T})$.



(a) Nash social welfare w_0 (b) The egalitarian welfare $w_{-\infty}$

Figure 1.4: Contour plots and infinitesimal random walks around expected utilities, illustrating corresponding utility changes under Nash social welfare and the egalitarian welfare metrics

Therefore, the egalitarian welfare warrants a more sophisticated approach to narrowing the regret. In the next section, we focus on dynamic fair allocation under the egalitarian welfare and propose two computationally efficient policies based on fluid policies that boast low regrets.

1.3 Dynamic Allocation Under the Egalitarian Welfare

In this section, we concentrate on the dynamic fair allocation problem in the context of the egalitarian welfare. We first revisit and generalize the fluid problem, and apply the duality theory to understand its dual solutions and to obtain optimality conditions. We then propose an online policy that re-solves the fluid problem and show it attains $O(1) + \kappa_u O(\log T)$ under a classical regularity condition. Lastly, we improve this re-solving policy by integrating a thresholding technique to obtain $O(1) + \kappa_u O(\log^{2+\varepsilon} T)$ even in the absence of the regularity condition. In this section we focus on admissible arrival distributions $P \in \mathcal{P}$ with finite support Θ_P, namely $|\Theta_P| < \infty$.

1.3.1 Understanding the Fluid Problem and Its Dual

In the preceding section, we exposed the limitations of adhering strictly to single-shot fluid policies throughout the horizon, noting they can incur regret as large as $O(\sqrt{T})$. Fluid policies, despite fully exploiting the knowledge of the underlying distribution, suffer from their static nature. Specifically, fluid policies are computed at the outset of the horizon and ignorant of the entire arrival sequence, thereby neglecting any insights that could be extracted from the realizations. Although the central planner is challenged by the non-anticipativity constraint, past arrivals and allocations are accessible and can substantially guide subsequent allocation decisions. In light of this, we explore the potential of updating and re-solving the fluid problem throughout the decision-making process to allow for insights gained alongside incoming resources.

Before we start examining the fluid problem in greater detail for the purposing of re-solving it, we notice each agent has already accumulated a certain utility as some resources have arrived and allocation decisions have been made. Hence, it is no longer reasonable to assume a zero initial utility U_0^i for every agent i; rather, agents' utilities accumulated so far are updated and should be incorporated into the fluid problem formulation. We normalize the fluid problem by the time remaining and generalize it into the following epigraph form, parameterized by an admissible arrival distribution P as well as time-normalized initial utilities ρ (for example, initially the parameter is taken as $\rho_0 = U_0/T$). Recall the support of distribution P is denoted by $\Theta_P \subseteq \Theta$, which we assume to be finite in this section (i.e., $|\Theta_P| < \infty$).

$$
\begin{aligned}
J(\rho) = \underset{\Upsilon \in \mathbb{R}, \boldsymbol{x} \in \mathbb{R}^{|\Theta_P| \times n}}{\text{maximize}} \quad & \Upsilon \\
\text{subject to} \quad & \rho^i + \mathbb{E}_{\theta \sim P}[u^i_\theta(x^i_\theta)] - \Upsilon \geq 0, \quad && i \in [n] \\
& \sum_{i \in [n]} x^i_\theta = 1, && \theta \in \Theta_P \\
& x^i_\theta \geq 0, && i \in [n], \theta \in \Theta_P.
\end{aligned}
\tag{1.5}
$$

Notice in particular, $J(0) = \mathsf{FLU}/T$ is the time-normalized initial fluid benchmark. Since the objective coincides with the value of the variable Υ, we sometimes omit it when we refer to primal solutions in the sequel. For $\rho \in \mathbb{R}^n$, we denote by $\boldsymbol{x}_\rho$ an optimal solution to the primal fluid problem $J(\rho)$ (1.5).

To further motivate the parameterized fluid problem (1.5), the updated fluid problem that any re-solving policy re-solves immediately after resources t has been allocated can be formulated as a parameterized fluid problem $J(\rho)$, with the parameter ρ taken as $\rho_t = U_t/(T-t)$. Moreover, since time-normalized utilities have bounded random increments, we would focus mainly on local behavior of $J(\rho)$ to infer the performance of fluid policies and re-solving policies.

Dual problem To better understand the parameterized fluid problem (1.5), we consider the following optimization problem, which we soon show to be its dual problem.

$$
D_J(\rho) := \inf_{\nu \in \Delta_n} \left\{ \nu^\top \rho + \phi(\nu) \right\}, \tag{1.6}
$$

where we denote by $\phi(\nu) \equiv \mathbb{E}_\theta[\phi_\theta(\nu)]$ the expected value of $\phi_\theta(\nu) \equiv \sup_{y \in \mathcal{Y}_\theta} \nu^\top y$, which is the the support function over the feasible set of type-θ utilities $\mathcal{Y}_\theta := \{y \in \mathbb{R}^n : y^i = u^i_\theta(x^i) \text{ for some } x \in \Delta_n\}$. In particular, $D_J(0) = \inf\{\phi(\nu) : \nu \in \Delta_n\}$. In the following lemma, we show $D_J(\rho)$ is the dual problem to $J(\rho)$ and shares the same optimum by Lemma 5, a straightforward result from Slater's theorem. The proof is deferred to Section A.1.5.

Lemma 5. *$D_J(\rho)$ is the dual problem for $J(\rho)$, and strong duality holds, i,e., $J(\rho) = D_J(\rho)$.*

Lemma 5 reduces the dimension of the primal problem $J(\rho)$, which involves $n \times |\Theta_P| + 1$ variables, to n in the dual problem $D_J(\rho)$. We show in Lemma 6 that optimal primal solutions can be easily derived from optimal dual solutions.

Another implication of Lemma 5 involves the relationship between the $J(\rho)$ and $\phi(\nu)$ functions. By the lemma, $J(\rho) = D_J(\rho)$ is concave, being a pointwise infimum; the shape (or curvature) of $J(\rho)$ largely depends on the shape (or curvature) of the dual function $\phi(\nu)$ because the remaining part $\nu^\top \rho$ is linear. Indeed, we provide in Lemma 8 lower envelopes of $\phi(\nu)$ that increase with the distance $\|\nu - \nu_0\|$ from an optimal dual variable ν_0, parameterized by the curvature κ_u of the admissible utility functions and certain characteristics of optimal primal solutions (the fluid policies). Based on the shape of $\phi(\cdot)$, we will then present a result on the sensitivity of optimal dual variable ν_ρ against ρ, to argue about the shape of $J(\rho)$ and to characterize the performance of fluid policies and re-solving policies.

Optimal dual solution and optimality conditions The parameterized dual problem (1.6) always has a minimizer by Weierstrass' theorem, as Δ_n is compact and the objective is continuous. Furthermore, we show in Lemma 7 that the optimal dual solution is unique if there exists a optimal primal solution that is regular (Assumption 2).

For $\rho \in \mathbb{R}^n$, we denote an optimal dual solution to the dual problem (1.6) by

$$\nu_\rho \in \arg\min_{\nu \in \Delta_n} \{\nu^\top \rho + \phi(\nu)\}.$$

In particular, we denote an optimal dual solution to $D_J(0)$ by $\nu_0 \in \arg\min_{\nu\in\Delta_n} \phi(\nu)$.

In fact, the complete dual problem to (1.5) involves more dual variables than merely ν: the complete set of optimal dual variables $(\nu_\rho, \lambda_\rho, \mu_\rho) \in \arg\min\{q_\rho(\nu, \lambda, \boldsymbol{\mu}) : \nu \in \Delta_n, \lambda \in \mathbb{R}^{|\Theta_P|}, \boldsymbol{\mu} \in \mathbb{R}_+^{n\times|\Theta_P|}\}$, where λ and $\boldsymbol{\mu}$ denote dual variables corresponding to the second and third set of constraints in (1.5). However, we can omit the nuisance dual variables $(\lambda_\rho, \boldsymbol{\mu}_\rho)$ and only refer to ν_ρ as the optimal dual solution, because ν_ρ suffices to identify the complete set of $(\nu_\rho, \lambda_\rho, \mu_\rho)$. See Section A.2 for a more detailed discussion as well as the Karush-Kuhn-Tucker optimality conditions.

We present a rather classical result on the general relationship between the optimal primal and dual solutions $(\boldsymbol{x}_\rho, \nu_\rho)$ and describe them in more detail at $\rho = 0$. The proof can be found in Section A.1.6.

Lemma 6. *Fix $P \in \mathcal{P}$ with finite support $|\Theta_P| < \infty$. For any $\rho \in \mathbb{R}^n$, if $\boldsymbol{x}_\rho$ is an optimal primal solution to $J(\rho)$, then for almost every type $\theta \in \Theta_P$,*

$$(x_\rho)_\theta \in \arg\max \left\{ \nu_\rho^\top u_\theta(x_\theta) : x_\theta \in \Delta_n \right\}.$$

In particular, for $\rho = 0$, if $(\boldsymbol{x}_0, \nu_0)$ is an optimal primal-dual solution pair to $(J(0), D_J(0))$, then

a) $J(0) = \mathbb{E}_\theta[u^i_\theta((x_0)^i_\theta)]$ for every agent i; and

b) ν_0 is interior, i.e., $\nu^i_0 > 0$ for every agent i.

1.3.2 Re-solving the Fluid Problem

In this section, we materialize the idea of developing policies based on updating and re-solving the fluid problem, to propose the Backward Infrequent Re-solving (BIR) policy, detailed in Algorithm 2. The BIR policy is inspired by fluid-based re-solving policies proposed in previous literature, especially by Jasin and Kumar [46] and Balseiro *et al.* [47]. The BIR policy is characterized by re-solving the updated fluid problem at certain epochs determined by a scheduling function $f : \mathbb{N} \to \mathbb{N}$ and by adopting the updated fluid policy between consecutive re-solving epochs. We show that under a mild and classical regularity condition on the initial fluid solution, BIR attains an asymptotic regret bounded by $O(1) + \kappa_u O(\log T)$ under the egalitarian welfare $w_{-\infty}$, marking a significant improvement from the fluid policy.

To clarify some notation used in this section, recall the definition of the natural filtration $\mathcal{F} = (\mathcal{F}_0, \mathcal{F}_1, \ldots, \mathcal{F}_T)$ given by $\mathcal{F}_t = \sigma(\theta_s : s \in [t])$. For ease of notation, for any random vector $v \in \mathbb{R}_n$, denote $\overline{v} := \sum_{i\in[n]} v^i/n$ and $\tilde{v} := v - \overline{v}\mathbf{1}$. For any random vector-valued process $(v_t)_t$ adapted to the natural filtration $\mathcal{F}$, denote the operator Δ such that $\Delta v_t := v_t - \mathbb{E}[v_t|\mathcal{F}_{t-1}]$.

The updated fluid problem At the end of period $t \in [T]$ (i.e., after resource t has arrived and been allocated to agents), to incorporate past information, the central planner could update previously forecast utilities with realizations observed so far and react to the stochastic deviations from the expected path initially predicted by the fluid problem. More formally, by time $t \in [T]$, each agent i has accumulated a utility of U_t^i, and the expected terminal utility is $\mathbb{E}[U_T^i | \mathcal{F}_t] = U_t^i + (T-t)\mathbb{E}_\theta[u_\theta^i(x_\theta^i)] = (T-t)(\rho_t^i + \mathbb{E}_\theta[u_\theta^i(x_\theta^i)])$ under type-based policy $\boldsymbol{x}$ conditional on the history up to time t, where we define

$$\rho_t = \frac{U_t}{T-t} \in \mathbb{R}_+^n$$

as the agents' time-normalized utilities, adapted to the natural filtration (i.e., $\rho_t \in \mathcal{F}_t$). In particular, $\rho_0 = U_0/T = 0$ at the beginning. The agents' utilities are normalized by $T-t$, the time remaining until the end of the horizon, because the closer to the end, the larger influence existing utilities have on optimizing future allocation decisions. The updated fluid problem at the end of period t is therefore to

$$\text{maximize}\left\{\min_i \left(\rho_t^i + \mathbb{E}_\theta\left[u_\theta^i\left(x_\theta^i\right)\right]\right) : \boldsymbol{x} \in \Delta_n^{\Theta_P}\right\}, \tag{1.7}$$

after omitting the coefficient of $T-t$. The updated fluid problem (1.7) takes the form of a parameterized fluid problem (1.5), with the parameter ρ taken as the time-normalized utilities ρ_t, and hence its optimum equals $J(\rho_t)$ in Equation (1.5).

Infrequent re-solving policy An infrequent re-solving policy is defined by a re-solving schedule, given by the set of re-solving epochs $\mathcal{T} := \{t_0, t_1, \ldots, t_K\}$, where $0 = t_0 < t_1 < t_2 < \cdots < t_K = T-1$. For every t, denote the last re-solving epoch $L(t) := \sup\{s \in \mathcal{T} : s < t\}$ and the next re-solving epoch $N(t) := \inf\{s \in \mathcal{T} : t \leq s\}$, such that $L(\cdot)$ and $N(\cdot)$ are both non-decreasing, and $L(t) < t \leq N(t)$. In particular, $L(t_k) = t_{k-1}$ and $N(t_k) = t_k$ for any $k \in [K]$. Alternatively, denote the last re-solving index $k(t) := \sup\{\ell \geq 0 : t_\ell < t\}$ so that $t_{k(t)} = L(t)$.

An infrequent re-solving policy, by definition, re-solves the updated fluid problem after receiving and allocating the t_k-th resource for each re-solving epoch t_k, obtains an optimal type-based fluid policy F_{t_k}, and acts accordingly until the next re-solving epoch t_{k+1}, i.e., if we denote the current fluid policy in use immediately after the t-th resource by F_t, then $\mathsf{F}_{t_k} = \mathsf{F}_{t_k+1} = \cdots = \mathsf{F}_{t_{k+1}-1}$. For example, the initial fluid policy is denoted by $\mathsf{F} \equiv \mathsf{F}_0$ as $t_0 = 0$. Hence, faced with resource t, the central planner acts according to the fluid policy F_{t-1} from the end of period $t-1$; F_{t-1} is in turn computed in the last re-solving epoch by $t-1$, namely $\mathsf{F}_{t-1} = \mathsf{F}_{L(t)} \in \mathcal{F}_{L(t)}$. The allocation decision for resource t is thus $x_t = (\mathsf{F}_{t-1})_{\theta_t} = (\mathsf{F}_{L(t)})_{\theta_t} \in \Delta_n$.

In particular, we consider the class of **Backward Infrequent Re-solving** (**BIR**) policies, whose re-solving epochs $(t_0, t_1, \ldots, t_K)$ are determined recursively based on their time remaining until the end of the horizon. More specifically, the re-solving epochs are defined recursively by $t_0 = 0$ and $T - t_k = f(T - t_{k-1})$ for an admissible scheduling function $f : \mathbb{N} \to \mathbb{N}$ (see Definition 3), until the penultimate period $t_K = T - 1$ first shows up in the sequence of re-solving epochs. We show in Lemma 27 that this prototypical definition is in fact equivalent to that of an infrequent re-solving policy. Allowing for maximum flexibility because the function f is arbitrary and the central planner knows the horizon length, it is solely for the purposes of assisting with policy design and performance analysis. We thenceforth parameterize BIR policies by their scheduling functions f and denote BIR_f.

Definition 3. *An* admissible scheduling function *is a non-decreasing function* $f : \mathbb{N} \to \mathbb{N}$ *which is strictly contracting except at* 1*, i.e.,* $f(t) < t$ *for* $t \geq 2$.

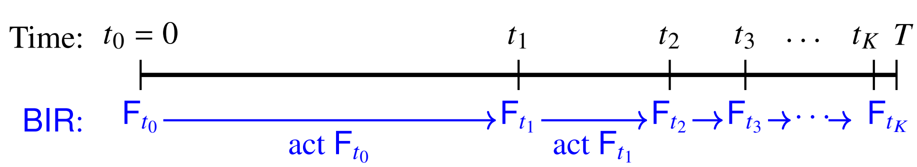


Figure 1.5: Re-solving schedule of the BIR policy $(t_0 = 0, t_1, \ldots, t_K = T - 1)$

Below are some running examples of re-solving schedules formulated in the backward recursion form. We will revisit them with specific parameterizations to instantiate the performance guarantee.

- The arithmetic BIR policy takes $f(t) = \max(t-h, 1)$ for some positive integral constant $h \in \mathbb{N}$ and re-solves $K = \Theta(T)$ times. In particular, the policy is also known as the frequent re-solving (FR) policy if $h = 1$.

- The geometric BIR policy takes $f(t) = \max(\lfloor \alpha t \rfloor, 1)$ for some constant $\alpha \in (0, 1)$ and re-solves $K = \Theta(\log T)$ times. In particular, the policy is also known as the midpoint re-solving policy if $\alpha = 1/2$.

- The exponential-geometric BIR policy takes $f(t) = \lfloor t^{1/\eta} \rfloor$ for some constant $\eta > 1$ and re-solves $K = \Theta(\log \log T)$ times.

The prototypical BIR policy is detailed in Algorithm 2 for a general welfare metric w, although we are primarily interested in its performance under the egalitarian welfare $w_{-\infty}$, and the re-solving schedule is summarized in Figure 1.5.

Algorithm 2 Backward Infrequent Re-solving (BIR)

Input: horizon length $T \in \mathbb{N}$, admissible scheduling function $f : \mathbb{N} \to \mathbb{N}$,
welfare metric $w : \mathbb{R}_+^n \to \mathbb{R}$, admissible arrival distribution $P \in \mathcal{P}$ supported on Θ_P.
Initialize: set $U_0^i \leftarrow 0$ for $i \in [n]$
set $t_0 \leftarrow 0$ and $t_k \leftarrow T - f(T - t_{k-1})$ for $k = 1, \ldots$ until $t_K \geq T - 1$
for $k = 0, 1, \ldots, K$ **do**
  solve the updated fluid problem with updated utilities

$$\mathsf{F} \in \arg\max \left\{ w\left(U_{t_k} + (T - t_k)\mathbb{E}_{\theta \sim P}\left[u_\theta(x_\theta)\right]\right) : \boldsymbol{x} \in \Delta_n^{\Theta_P} \right\};$$

  for $t = t_k + 1, \ldots, t_{k+1}$ **do** `// implement the fluid policy`
    observe incoming resource type $\theta_t \sim P$
    allocate $x_t \leftarrow \mathsf{F}_{\theta_t}$
    update utilities $U_t^i \leftarrow U_{t-1}^i + u_t^i(x_t^i)$ for every agent $i \in [n]$
Output: allocations $(x_t : t \in [T])$ and final welfare $w(U_T^i)$.

Performance

The performance of fluid-based policies naturally depends on the fluid problem (see Section 1.3.1); likewise, that of re-solving policies depends on the parameterized updated fluid problem $J(\cdot)$ (1.5). Before presenting the main performance guarantee of the BIR policy, we present as its foundation the following assumption on the local behavior of $J(\cdot)$ around zero. The assumption is centered around the de-meaned $\tilde{\rho}$ because $J(\cdot)$ is invariant to constant shifts, namely $J(\rho) = \overline{\rho} + J(\tilde{\rho})$, and the $(\tilde{\rho}_t : t \in [T])$ process is martingale-like (see Lemma 14). We defer a more detailed discussion on when the assumption holds to the next subsection, where we give a simple sufficient condition (Proposition 3) based on the initial fluid solution.

Assumption 1. *There exists some constants $\delta > 0$ and $\kappa_J \geq 0$ such that if $\|\tilde{\rho}\|_\infty < \delta$, then*

1. *for any optimal solution $\boldsymbol{x}_\rho$ to the fluid problem $J(\rho)$, the expected utilities are equal across agents, i.e.,*

$$J(\rho) = \rho^i + \mathbb{E}_\theta\left[u^i_\theta((x_\rho)^i_\theta)\right] \text{ for any agent } i;$$

2. *$J(\cdot)$ is differentiable at 0 and admits a lower quadratic envelope around zero, i.e.,*

$$J(\rho) \geq J(0) + \nabla J(0)^\top \rho - \frac{\kappa_J}{2}\|\tilde{\rho}\|_2^2 .$$

Assumption 1 is inspired by a similar assumption in Balseiro *et al.* [47] used to show performance guarantees for the frequent re-solving (FR) policy. It assume two properties on the $J(\cdot)$ function when $\tilde{\rho} = \rho - \overline{\rho}\mathbf{1}$ is close to zero, i.e., when the agents' time-normalized utilities ρ^i are close to one another. The first states that agents have equal expected utilities under optimal fluid policies. This occurs when the gaps in agents' existing utilities are not too large to neutralize with the fluid solution; the property establishes a connection between agents' expected per-period utilities with time-normalized existing utilities, and allows for a closer examination in the evolution dynamics of the $(\tilde{\rho}_t : t \in [T])$ process (Lemma 14). The second is a local strong smoothness property on the concave function $J(\cdot)$, which will be useful in developing lower bounds on the difference between $J(\rho)$ and $J(0)$. Since the fluid benchmark is an upper bound of the expected hindsight optimum (recall Lemma 2), this translates to a bound on the regret of the BIR policy.

We present the following performance guarantee on the BIR_f policy parameterized by the admissible scheduling function f. The proof can be found in Section A.1.7.

Theorem 2. *Fix $P \in \mathcal{P}$ and consider the egalitarian welfare $w_{-\infty}$. Let Assumption 1 hold. The* BIR_f *policy with scheduling function $f(t) = \Omega(\sqrt{t \log^{2+\varepsilon} t})$ for any positive $\varepsilon > 0$ incurs a regret bounded by*

$$\mathcal{R}_T(\mathsf{BIR}_f) \le O(1) + \frac{\kappa_J}{2} \sum_{t=1}^{T} \frac{1}{f(t)}.$$

Theorem 2 shows the BIR_f policy is asymptotically optimal because $\mathcal{R}_T(\mathsf{BIR}_f)/T \le O(T^{-1}) + \kappa_J \tilde{O}(T^{-1/2})$ for $f(t) = \Omega(\sqrt{t \log^{2+\varepsilon} t})$. It also exhibits a clear distinction in the performance of the BIR policy driven by the problem primitive κ_J: the regret is of order $O(1)$ when $\kappa_J = 0$ and of order $O(\log T)$ at the tightest when $\kappa_J > 0$, as any admissible scheduling function f must satisfy $f(t) \le t$ for $t \in \mathbb{N}$. In particular, it gives the following performance guarantees to the three running examples of BIR policies; when $\kappa_J > 0$, the geometric BIR policy is able to attain the best-order $O(\log T)$ regret with $\Theta(\log T)$ re-solving epochs, whereas when $\kappa_J = 0$, the exponential-geometric BIR policy attains uniformly bounded $O(1)$ regret with $\Theta(\log \log T)$ re-solving epochs. Remarkably, re-solving only $O(\log \log T)$ times is sufficient for a similar performance to that of the frequent re-solving (FR) policy (see the experiment in Section 1.4.2).

- The arithmetic BIR policy takes $f(t) = \max(t - h, 1) = \Omega(t)$ for some positive integral constant $h \in \mathbb{N}$, re-solves $K = \Theta(T)$ times, and incurs a regret of $O(1) + \kappa_J O(\log T)$.
- The geometric BIR policy takes $f(t) = \max(\lfloor \alpha t \rfloor, 1) = \Omega(t)$ for some constant $\alpha \in (0, 1)$, re-solves $K = \Theta(\log T)$ times, and incurs a regret of $O(1) + \kappa_J O(\log T)$.
- The exponential-geometric BIR policy takes $f(t) = \lfloor t^{1/\eta} \rfloor$ for some constant $\eta \in (1, 2)$, re-solves $K = \Theta(\log \log T)$ times, and incurs a regret of $O(1) + \kappa_J O(T^{1-1/\eta})$.

The implication of Theorem 2 is twofold. With the regret of non-anticipative policies against the hindsight optimum attributable to either an intrinsic lack of clairvoyance or sub-optimal policy design, the performance guarantee demonstrates that BIR policies can be near-optimal in both dimensions with appropriate re-solving schedules. On the one hand, their low regret gives credit to the policy design, in addition to the computational efficiency. On the other hand, with the non-anticipative BIR attaining performance comparable to a clairvoyant, Theorem 2 showcases that the central planner is able to greatly benefit from distributional knowledge.

The proof of Theorem 2 is inspired by ideas in Jasin and Kumar [46] and Balseiro *et al.* [47] but catered to our problem structure. It is worth noting our analysis here is different from that in the classical revenue management problem in various ways. The egalitarian welfare objective demands a special treatment that connects the final welfare and the fluid benchmark. In addition, the $(\rho_t : t \in [T])$ process denotes the time-normalized cumulative utilities, which is not a martingale. In the classical revenue management problem, the state of the algorithm is the ratio between the remaining capacity to the remaining time, which behaves like a martingale. In contrast, the $(\tilde{\rho}_t : t \in [T])$ process has a positive drift, which needs to be taken care of by the tilde operator, so that the process has zero mean and zero drift and is martingale-like.

We give an outline of the proof to provide the reader with some intuition. Since the BIR policy does not necessarily re-solve at every period, the $(\tilde{\rho}_t : t \in [T])$ process is not exactly a martingale but demonstrates similar behavior. The first step of the proof is to introduce an auxiliary martingale process $(M_t : t \in [T])$ that evolves closely related to $(\tilde{\rho}_t : t \in [T])$. To estimate the performance of the policy, we investigate the hitting time τ when the $(\tilde{\rho}_t : t \in [T])$ process first deviates too much from zero and exits the neighborhood $D_\delta(0)$ where $J(\tilde{\rho})$ is well-behaved by Assumption 1; the performance of the policy by τ provides a lower bound on the final egalitarian welfare, in comparison with the fluid benchmark $\mathsf{FLU} = TJ(0)$, to yield a bound on the regret. More specifically, we are interested in the expected hitting time $\mathbb{E}[\tau]$ and the performance of the algorithm $\min_i U^i_\tau$ by then.

We start by defining the hitting time

$$\tau := T \wedge \inf\{t \in [T] : \|\tilde{\rho}_t\|_\infty \geq \delta\} \tag{1.8}$$

when the random zero-mean process $(\tilde{\rho}_t : t \in [T])$ exits the neighborhood $D_\delta(0)$. Note τ is a well-defined stopping time because $\tilde{\rho}_t \in \mathcal{F}_t$ for any $t \in [T]$. Since for resource t the central planner adopts the fluid policy $\mathsf{F}_{L(t)}$ computed at the end of period $L(t)$, the two conditions in Assumption 1 hold if $L(t) < \tau$, allowing for good control over the performance. The following proposition provides a sufficient condition for the exiting to happen close to the end in expectation. More specifically, Proposition 2 shows the hitting time is in expectation within a constant before the end of the horizon if the BIR_f policy re-solves at least $\Theta(\log\log T)$ times. The proof can be found in Section A.1.10.

Proposition 2. *Fix $P \in \mathcal{P}$. Under the BIR_f policy with $f(t) = \Omega(\sqrt{t\log^{2+\varepsilon} t})$,*

$$\mathbb{E}[T - \tau] = O(1).$$

The last part of the proof is to estimate the welfare at the hitting time τ, which is at least $\min_i U^i_\tau \geq \overline{U}_\tau - \|\tilde{U}_\tau\|_\infty$. Recall $\overline{U}_\tau = \sum_{i\in[n]}\sum_{t\in[\tau]} u^i_t(x^i_t)/n$ is the average utility agents accumulate by τ, and $\tilde{U}_\tau = U_\tau - \overline{U}_\tau$. For the expected average utility $\mathbb{E}[\overline{U}_\tau]$ by τ, we first invoke part 1 of Assumption 1 and connect it with $J(\tilde{\rho}_t)$ through the optional stopping theorem; we then invoke part 2 of Assumption 1 to give a quadratic-form lower bound, whose first-order term has zero expected values and whose second-order term we estimate using the orthogonality of martingale differences. The maximum variation term $\|\tilde{U}_\tau\|_\infty$ can be estimated by noting $\|\tilde{\rho}_{\tau-1}\|_\infty < \delta$ along with the boundedness of utility increments.

Sufficient condition for Assumption 1

Assumption 2 (Regularity condition, a.k.a. linear independence constraint qualification (LICQ))**.** *At $\boldsymbol{x}_0$, an optimal primal solution to $J(0)$, the gradients of all active constraints in* (1.5) *are linearly independent.*

The regularity condition is a conventional constraint qualification in the local theory of constrained optimization. It states no active constraint at the referenced solution is redundant in the first-order sense. Regular points enjoy numerous properties and can greatly help facilitate analysis (e.g., the tangent cone is identical to the cone of first-order feasible variations). In particular, regularity guarantees the uniqueness of optimal dual solutions (Lemma 7); violating regularity may imply multiple dual variables satisfying the KKT necessary conditions (see discussion after Lemma 8).

The regularity condition can be viewed as a generalization of the non-degeneracy condition in linear programs, i.e., all basic variables take non-zero values; indeed, it reduces to the non-degeneracy condition when the optimization problem is a linear program. Our particular problem (1.5) reduces to an LP if all utilities $u_\theta^i(\cdot)$ are linear, and Assumption 2 reduces to the non-degeneracy condition, namely no more constraints are active than the dimension of optimization variables.

We present the following well-known result between regularity of the optimal primal solution and the uniqueness of the optimal dual solution. It enables us to describe the shape of the dual function $\phi(\cdot)$ and parameterized fluid benchmark $J(\cdot)$ in more detail.

Lemma 7 (Primal regularity and dual uniqueness, Proposition 3.3.1 [84])**.** *If an optimal primal solution $\boldsymbol{x}_\rho$ to $J(\rho)$ is regular, the optimal dual solution ν_ρ is unique.*

Since Assumption 1 describes the local behavior of $J(\cdot)$, it is crucial to gauge how faraway the initial fluid solution $\boldsymbol{x}_0$ is from irregular points, i.e., how much perturbation is allowed before inactive constraints may become active and $J(\cdot)$ is no longer well-behaved. Hence, we introduce the notion of margin γ for the optimal fluid solution $\boldsymbol{x}_0$, which will be useful in determining the size δ of the neighborhood in Assumption 1 where $J(\cdot)$ is well-behaved. By definition, $(x_0)_\theta^i \in \{0\} \cup [\gamma, 1-\gamma] \cup \{1\}$ for every agent i and almost every type $\theta \in \Theta_P$.

Definition 4 (Margin)**.** *Fix $P \in \mathcal{P}$ with finite support $|\Theta_P| < \infty$. The* margin *of an optimal primal solution $\boldsymbol{x}_0$ to $J(0)$ is defined as $\gamma := \min(\min\{(x_0)_\theta^i : (x_0)_\theta^i > 0\}, \min\{1-(x_0)_\theta^i : (x_0)_\theta^i < 1\})$, where the minima are taken over almost every $\theta \in \Theta_P$.*

We now present the following proposition that exhibits regularity of the initial fluid solution $\boldsymbol{x}_0$ (Assumption 2) as a sufficient condition for the local well behavior of the parameterized fluid benchmark $J(\cdot)$ in Assumption 1, which is in turn crucial in the performance guarantee of the

BIR policy (Theorem 2). The proposition converts Assumption 1 on the entire neighborhood to a simple and classical constraint qualification at a single point, namely the initial fluid solution $\boldsymbol{x}_0$.

Proposition 3. *Fix $P \in \mathcal{P}$ with finite support $|\Theta_P| < \infty$. Let Assumption 2 hold. There is a positive constant $\varphi \in (0,2)$ such that Assumption 1 holds with $\delta = \frac{\varphi^2}{4\sqrt{n}} \min(\gamma, (\kappa_u)^{-1} \min_i v_0^i)$ and $\kappa_J = 2\kappa_u/\varphi^2$, i.e., if $\|\tilde{\rho}\|_\infty < \delta$, then*

1. *for any optimal solution $\boldsymbol{x}_\rho$ to the fluid problem $J(\rho)$, the expected utilities are equal across agents, i.e.,*

$$J(\rho) = \rho^i + \mathbb{E}_\theta \left[u_\theta^i((x_\rho)_\theta^i)\right] \text{ for any agent } i;$$

2. *$J(\cdot)$ is differentiable at 0 and admits a lower quadratic envelope around zero, i.e.,*

$$J(\rho) \geq J(0) + \nabla J(0)^\top \rho - \frac{\kappa_u}{\varphi^2} \|\tilde{\rho}\|_2^2;$$

The proof is deferred to Section A.1.11. We give some intuition of the proof of Proposition 3. The proof hinges heavily on the shape of the dual function $\phi(\cdot)$ because of the definition $J(\rho) = \inf_v\{v^\top \rho + \phi(v)\}$. It invokes Lemma 8, presented below, which provides a lower envelope for $\phi(\cdot)$ around the initial dual solution v_0, to observe the stability of the optimal dual solution against initial time-normalized utilities. More precisely, when $\tilde{\rho}$ is close enough to zero, the corresponding optimal dual solution v_ρ does not deviate from the initial v_0 too much, namely $\left\|v_\rho - v_0\right\|_2 \leq \min(\kappa_u \gamma, \min_i v_0^i)$; in particular, if all utilities are linear, it does not deviate whatsoever.

Lemma 8 describes the shape (or curvature) of the locally strongly convex dual function $\phi(\cdot)$ around the initial optimal dual solution v_0. More specifically, it gives a lower envelope of the dual function based on the semimetric ℓ and parameterized by the distance $\|v - v_0\|$ from v_0, helping analyzing the response of the optimal dual solution to perturbations in time-normalized utilities $\tilde{\rho}$. When all agents have linear utilities, ℓ grows linearly, yielding a conic envelope; otherwise, ℓ grows locally quadratically and linearly beyond.

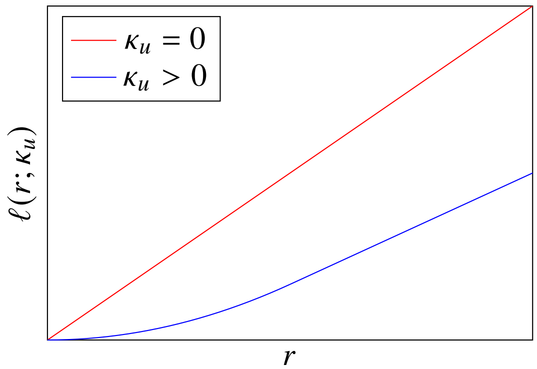


Figure 1.6: Semimetric $\ell(\cdot;\kappa_u)$ paramaterized by the maximum curvature κ_u of utility functions

Lemma 8 (Lower envelope of dual function $\phi(\cdot)$)**.** *Fix $P \in \mathcal{P}$ with finite support $|\Theta_P| < \infty$. Let Assumption 2 hold. There is a positive constant $\varphi \in (0, 2)$ such that for any $\nu \in \Delta_n$,*

$$\phi(\nu) \geq \phi(\nu_0) + \ell(\|\nu - \nu_0\|_2\,; \kappa_u),$$

where we define a the parameterized semimetric $\ell(\cdot;\cdot)$ is given by

$$\ell(r;\kappa_u) = \begin{cases} \varphi\gamma r, & \text{if } \kappa_u = 0, \\ \dfrac{\varphi^2}{4\kappa_u} r^2, & \text{if } \kappa_u > 0 \text{ and } r < r_\nu, \\ \dfrac{\varphi^2}{2\kappa_u} r_\nu \left(r - \dfrac{r_\nu}{2}\right), & \text{if } \kappa_u > 0 \text{ and } r \geq r_\nu, \end{cases}$$

for any fixed $r_\nu \in [0, \kappa_u\gamma]$.

The proof of Lemma 8 is intricate and contingent upon the regularity condition (Assumption 2). The main idea is to leverage the strong smoothness of admissible utility functions (Definition 1) and to consider specific feasible solutions $\hat{\boldsymbol{x}}$ that would give rise to meaningful lower bounds to $\phi(\nu)$ with a coefficient φ. The last part is to show $\varphi > 0$ to render this lower envelope non-trivial by arguing the otherwise linear dependence of active constraints at $\boldsymbol{x}_0$ in the initial fluid problem (1.5), which would violate the regularity condition (LICQ).

Discussion

We now reflect the essential underlying assumptions and elements in the theoretical results in this section. Proposition 3 states that regularity is a sufficient condition for the key local well-behavior assumption for $J(\cdot)$ (Assumption 1), which yields the performance guarantee of the BIR policy (Theorem 2). The proposition is built upon two constants relevant to the initial fluid solution $\boldsymbol{x}_0$: φ indicating its regularity, and γ, the margin of the fluid solution. In the absence of regularity, φ can be zero, in which case not only the neighborhood $D_\delta(0)$ becomes trivial as $\delta = 0$, but the curvature of the lower quadratic envelope for $J(\cdot)$ explodes, nullifying the result. Hence, regularity of the initial fluid solution $\boldsymbol{x}_0$, which implies the uniqueness of the initial dual solution ν_0 by Lemma 7, serves as the indispensable foundation to bounding the gap $\mathsf{FLU} - \mathbb{E}[\mathsf{ALG}(\mathsf{BIR})]$ and hence the regret $\mathcal{R}_T(\mathsf{BIR}) = \mathbb{E}[\mathsf{OPT} - \mathsf{ALG}(\mathsf{BIR})]$ in the proof of Theorem 2.

We would also like to remark that this approach relies on the fact that the fluid benchmark FLU is an upper bound on the expected hindsight optimum $\mathbb{E}[\mathsf{OPT}]$, but we show in Lemma 9 that it is always too loose as long as the initial dual problem $D_J(0)$ (1.6) has non-unique optimal solutions. More specifically, the expected hindsight optimum $\mathbb{E}[\mathsf{OPT}]$ is always at least $\Omega(\sqrt{T})$ less than the fluid benchmark FLU, undermining the role of the fluid benchmark and hence the approach in proving Theorem 2. The proof of Lemma 9 can be found in Section A.1.17, which gives the explicit value of the asymptotic gap.

Lemma 9. *Fix $P \in \mathcal{P}$ with finite support $|\Theta_P| < \infty$. If the initial dual problem $D_J(0)$ (1.6) has more than one optimal solution, then* $\mathsf{FLU} - \mathbb{E}[\mathsf{OPT}] = \Omega(\sqrt{T})$.

In addition, even if the fluid solution is regular and $\varphi > 0$, having too small a margin γ would imply that the solution is too close to an irregular point, and that the neighborhood $D_\delta(0)$ shrinks alongside γ. The constant factors in the performance guarantees in Theorem 2 can be explosively large, indicating the vulnerability of the BIR policy against irregularity. See Section 1.4.4 for a numerical illustration of BIR's deteriorating performance when the initial fluid solution is nearly irregular.

In the next section, we will consider enhancing the BIR policy with an extra thresholding adjustment to defend against irregularity; furthermore, we are to show comparable asymptotic performance of the enhanced re-solving policy regardless of whether or not the initial fluid solution is regular.

Technical novelty

Before ending this section, we would like to comment on some key distinctions between our problem and analyses and those in Jasin and Kumar [46] and Balseiro *et al.* [47], whose works inspire the BIR policy and its performance analysis.

The most salient feature of our problem, in contrast with classical dynamic resource allocation and especially revenue management problems, is that its objective is non-linear and hence non-separable over time. Unlike in [46] and [47], the final welfare in our dynamic fair allocation problem is not a simple sum of welfare gains in each period.

Our problem allows for non-linear utility functions, in contrast to the revenue management problem in Jasin and Kumar [46], where reward is linear and additive. While Balseiro *et al.* [47] allows for non-linear reward functions, they only consider the frequent re-solving (FR) policy, which requires a simpler analysis than BIR, and provide separate analyses for linear and non-linear utility functions. In contrast, we provide a consolidated analysis that sheds more light onto the evolution of the problem state and corresponding dual solutions, based on problem primitives such as utility functions.

The most important novelty in our analysis is in the assumptions. Our analysis only requires regularity of the initial fluid solution. In contrast, Balseiro *et al.* [47] assume the initial fluid solution to be interior (SC 9), which is highly restrictive, to show the dual function is strongly convex (SC 7); in fact, this often does not hold in our dynamic fair allocation problem. In addition, they assume the dual function is differentiable everywhere (SC 4), which is also highly restrictive. For example, this assumption is violated by linear programs; this partly obstructs the path to a consolidated proof for claims under linear and non-linear utilities.

The last highlight of the BIR policy is its maximum flexibility in the re-solving schedule (Lemma 27). The BIR policy can choose virtually arbitrary re-solving epochs, and we develop a customized performance guarantee tailored to every scheduling function f. Moreover, the guarantee take a clean and interpretable form that depends on the convergence or growth rate of the series $\sum_t 1/f(t)$. In contrast, Balseiro *et al.* [47] focuses only on frequent re-solving, and Jasin and Kumar [46] provides a highly intricate result through a Chernoff-like bound.

1.3.3 Re-solving the Fluid Problem with Thresholding

In this section, we continue exploring improvement of fluid policies, which attain a regret of order $\Theta(\sqrt{T})$ under the egalitarian welfare, through updating and re-solving the fluid problem. As an enhancement to the BIR policy proposed in the previous section, we propose the **Backward Infrequent Re-solving with Thresholding (BIRT)** policy, detailed in Algorithm 3, inspired by previous fluid-based re-solving policies proposed in the literature, especially by Bumpensanti and Wang [48]. Similar to the BIR policy, the BIRT policy's re-solving schedule is based on an admissible scheduling function $f : \mathbb{N} \to \mathbb{N}$ (see Figure 1.7). In addition to re-solving the fluid problem, the BIRT policy is marked by a thresholding adjustment to each fluid policy solution, parameterized by positive tuning constants $m, \varepsilon > 0$ (see Algorithm 3), before acting accordingly. Hence, we may denote the BIRT policy along with its parameters as $\text{BIRT}_{f,m,\varepsilon}$.

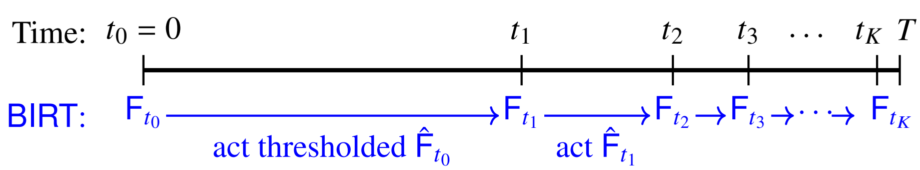


Figure 1.7: Re-solving schedule of the BIRT policy ($t_0 = 0, t_1, \ldots, t_K = T - 1$)

The highlight of the BIRT policy is that it can attain similar and even better regret compared with that of the BIR policy without the regularity assumption—in fact *without any assumption*. To be more precise, BIRT attains an $O(1)$ regret by re-solving only $O(\log \log T)$ times when utilities are almost surely linear, and otherwise an $O(\log^{2+\varepsilon} T)$ regret by re-solving only $O(\log T)$ times.

Algorithm 3 Backward Infrequent Re-solving with Thresholding (BIRT)

Input: horizon length $T \in \mathbb{N}$, admissible arrival distribution $P \in \mathcal{P}$ supported on Θ_P, welfare metric $w : \mathbb{R}_+^n \to \mathbb{R}$, admissible scheduling function $f : \mathbb{N} \to \mathbb{N}$, constants $m > 0$, $\varepsilon > 0$.

Initialize: set $U_0^i \leftarrow 0$ for $i \in [n]$

set $t_0 \leftarrow 0$ and $t_k \leftarrow T - f(T - t_{k-1})$ for $k = 1, \ldots$ until $t_K \geq T - 1$

for $k = 0, 1, \ldots, K$ **do**

solve the updated fluid problem with updated utilities

$$\mathsf{F} \in \arg\max \left\{ w\left(U_{t_k} + (T - t_k)\mathbb{E}_{\theta \sim P}\left[u_\theta(x_\theta)\right]\right) : \boldsymbol{x} \in \Delta_n^{\Theta_P} \right\};$$

set threshold

$$\gamma_{t_k} \leftarrow \frac{2m^2}{3n^2} \frac{\sqrt{(T - t_k)\log^{1+\varepsilon}(T - t_k)}}{(T - t_{k+1})}$$

for $\theta \in \Theta_P$ **do** `// threshold the fluid policy`

pick some agent $j \in \arg\max_i \mathsf{F}_\theta^i$ **for** $i \neq j$ **do**

set $\hat{\mathsf{F}}_\theta^i \leftarrow \mathsf{F}_\theta^i \mathbf{1}\{\mathsf{F}_\theta^i \geq \gamma_{t_k}\}$

set $\hat{\mathsf{F}}_\theta^j \leftarrow 1 - \sum_{i \neq j} \hat{\mathsf{F}}_\theta^i$

for $t = t_k + 1, \ldots, t_{k+1}$ **do** `// implement thresholded fluid policy`

observe incoming resource type $\theta_t \sim P$

allocate $x_t \leftarrow \hat{\mathsf{F}}_{\theta_t}$

update utilities $U_t^i \leftarrow U_{t-1}^i + u_t^i(x_t^i)$ for every agent $i \in [n]$

Output: allocations $(x_t : t \in [T])$ and final welfare $w(U_T^i)$.

Thresholding adjustment We next detail the thresholding adjustment used in BIRT (Algorithm 3) on fluid solutions, parameterized by the threhold γ. Given a fluid type-based policy $\mathsf{F} \in \Delta_n^{\Theta_P}$ and a threshold $\gamma \in (0, 1/n)$, we define its corresponding thresholded type-based policy $\hat{\mathsf{F}}$ by implementing the following procedure for each $\theta \in \Theta_P$. After thresholding we will arrive at a type-based policy $\hat{\mathsf{F}}$ such that $\hat{\mathsf{F}}^i_\theta \in \{0\} \cup [\gamma, 1-\gamma] \cup \{1\}$ for all $i \in [n]$ and $\theta \in \Theta_P$.

1. Pick an agent $j \in \arg\max_i \mathsf{F}^i_\theta$ receiving the largest type-θ allocation under F.

2. For other agents $i \neq j$, implement allocations only if they exceed the threshold γ and withhold the rest, i.e., set $\hat{\mathsf{F}}^i_\theta \leftarrow \mathsf{F}^i_\theta \mathbf{1}\{\mathsf{F}^i_\theta \geq \gamma\}$.

3. Allocate the remaining resource to agent j, i.e. set $\hat{\mathsf{F}}^j_\theta \leftarrow 1 - \sum_{i \neq j} \hat{\mathsf{F}}^i_\theta$.

Thresholding is a defining characteristic of the BIRT policies and, as shown in the experiments in Sections 1.4.2 and 1.4.3, it is crucial to obtain low regret guarantees especially when the initial fluid problem is not regular. We provide some intuition about the need for thresholding in Section 1.3.3.

Performance Analysis

Theorem 3 gives the performance guarantee for the $\mathsf{BIRT}_{f,m,\varepsilon}$ policies, parameterized by the admissible scheduling function f and thresholding constants $m, \varepsilon > 0$.

Theorem 3. *Fix $P \in \mathcal{P}$ with finite support $|\Theta_P| < \infty$. Consider the egalitarian welfare metric. If all utilities are linear ($\kappa_u = 0$) and there is a positive constant $\varepsilon > 0$ such that $\Omega\left(\sqrt[4]{\log^{1+\varepsilon} t/t}\right) \leq f(t)/t < 1$ for any $t \geq 1$, then for some $m > 0$,*

$$\mathcal{R}_T(\mathsf{BIRT}_{f,m,\varepsilon}) = O(1);$$

in general, if there are positive constants $\underline{M}, \overline{M} \in (0,1)$ such that $f(t)/t \in (\underline{M}, \overline{M})$ for any $t \geq 2$, *then for any $\varepsilon > 0$ and $m \in (0, \underline{M})$,*

$$\mathcal{R}_T(\mathsf{BIRT}_{f,m,\varepsilon}) = O(1) + \kappa_u O(\log^{2+\varepsilon} T).$$

The constant factors omitted in the big O notation are determined solely by P and f.

A full proof is provided in Section A.1.18. Theorem 3 gives performance guarantees for the BIRT policies based on re-solving schedules, thresholding adjustments and the curvature of the utility functions. When all utility functions are linear, the BIRT policy can attain $O(1)$ regret with re-solving as few as $O(\log \log T)$ times; when utility functions are nonlinear and concave ($\kappa_u > 0$), the BIRT policy can attain $O(\log^{2+\varepsilon} T)$ with re-solving $O(\log T)$ times.

Without assuming the initial fluid solution is regular, it is no longer viable to obtain regret bounds by comparing performances to the fluid benchmark, as the gap between it and the expected hindsight optimum can be as loose as $\Theta(\sqrt{T})$ (see Lemma 9). Instead, we need to resort to a more granular regret analysis directly involving the hindsight optimum. We next give an outline of the proof, inspired by the analysis in Bumpensanti and Wang [48], hinging on the structure of BIRT; in particular, the infrequent re-solving schedule warrants a serial regret analysis. More specifically, the regret of BIRT can be decomposed into a series of expected egalitarian welfare losses against the hindsight optimum, each attributed to one re-solving epoch in the schedule. To this end, we introduce a series of auxiliary policies OPT^k for $k = 0, 1, \ldots, K$, where OPT^k acts exactly the same as BIRT does until t_k and allocates in the hindsight optimal way afterwards, i.e., it picks the optimal allocation for the remaining resources given the decisions on the first t_k resources. With an abuse of notation, we also denote by OPT^k the egalitarian welfare generated by the OPT^k policy, i.e., the hindsight optimal egalitarian welfare starting with utility vector U_{t_k} at time t_k. More precisely,

$$\mathsf{OPT}^k := \max\left\{ w\left(U_{t_k} + \sum_{t=t_k+1}^{T} u_{\theta_t}(x_t) \right) : x_t \in \Delta_n \text{ for } t = t_k + 1, \ldots, T \right\}, \tag{1.9}$$

where U_t^i is agent i's utility by t under the BIRT policy. In particular, $\mathsf{OPT} = \mathsf{OPT}^0$. Notice, however, the auxiliary OPT^k policy is not an online policy, as it is anticipative starting from $t_k + 1$.

The auxiliary policies are introduced to quantify the egalitarian welfare loss due to each re-solving epoch. For example, comparing OPT and OPT^1 would give us the loss in the first epoch, as OPT^1 acts optimally with clairvoyance after t_1.

Hence, we can decompose the regret of BIRT into

$$\mathcal{R}_T(\mathsf{BIRT}) = \mathbb{E}\left[\mathsf{OPT} - \mathsf{ALG}(\mathsf{BIRT})\right] = \sum_{k=1}^{K} \mathbb{E}\left[\mathsf{OPT}^{k-1} - \mathsf{OPT}^k\right] + \mathbb{E}\left[\mathsf{OPT}^K - \mathsf{ALG}(\mathsf{BIRT})\right], \tag{1.10}$$

The telescopic decomposition is useful for two reasons. First, since the problems OPT^{k-1} and OPT^k share the same history up to time t_{k-1}, their difference in egalitarian welfare $\mathsf{OPT}^{k-1} - \mathsf{OPT}^k$ is isomorphic to $\mathsf{OPT} - \mathsf{OPT}^1$, with appropriately chosen initial welfare and horizon lengths. Second, we can analyze each term $\mathbb{E}[\mathsf{OPT}^{k-1} - \mathsf{OPT}^k]$ without considering inter-epoch interactions, i.e., consequences of previous allocations on subsequent periods. In other words, analyses of disjoint epochs are decoupled from one another. This follows because we can provide guarantees for each term that are *uniform over the agents' initial utilities*.

The decomposition warrants an inductive argument to show the main theorem using the following result on one-epoch expected egalitarian welfare loss. For the purpose of providing a uniform framework for analyzing each re-solving epoch, we assume the agents start with initial utilities $U_0 \in \mathbb{R}_+^n$ that may be positive, so that the analysis can be directly deployed to subsequent epochs.

Proposition 4. *Fix $P \in \mathcal{P}$ with finite support $|\Theta_P| < \infty$. Under the egalitarian welfare, for any initial utilities $U_0 \in \mathbb{R}_+^n$ and horizon length $T > 1$, if the first re-solving occurs at some $t_1 \in [T]$, and the initial threshold is set as $\gamma \in [0, 2(T - t_1)/3n^2 T]$,*

$$\mathbb{E}\left[\mathsf{OPT} - \mathsf{OPT}^1 | U_0\right] \leq 4|\Theta| t_1 \exp\left(-\frac{2\underline{p}^2 (T - t_1)^2 \gamma^2}{9L^2 T}\right) + \frac{3\kappa_u n^2}{4} \frac{\gamma^2 T^2}{T - t_1}, \tag{1.11}$$

where $\underline{p} = \min_\theta p_\theta$.

The complete proof is postponed to Section A.1.19. Its main idea is to focus on the coupling event E (A.1.9) where an offline policy exists after t_1 that can eventually replicate aggregate allocations of each type to each agent made by the hindsight optimal policy OPT. On the complementary decoupling event $E^{\complement}$, we can control the difference using the almost sure uniform bound $\mathsf{OPT} - \mathsf{OPT}^1 \leq t_1$ (Lemma 17), which follows because any loss the clairvoyant OPT^1 can incur against OPT must only stem from the sub-optimal allocations by the type-based policy during the first t_1 periods. The remaining two steps of the proof involve showing that the welfare loss is small on the coupling event, and that the coupling event is a high probability event, i.e., after t_1 it is still most likely to replicate the exact per-type per-agent aggregate allocations by the hindsight optimal policy. For the first, we leverage the strong smoothness property of admissible utility functions (Definition 1) to bound the welfare loss on the event. To show the second, we invoke the Lipschitz continuity property of the optimal allocations (Lemma 10 in the next section) to provide sufficient conditions in terms of the arrival process; we then prove that these sufficient conditions hold with high probability using concentration inequalities. Thresholding the fluid policy is critical to proving that the concentration inequalities hold uniformly over the initial agents' utilities as, otherwise, BIRT would likely fall into the trap of not being able to replicate hindsight optimal allocations due to irrevocability of allocations.

Lipschitz Continuity Property

The Lipschitz continuity property of the optimal allocation multi-function serves as a theoretical foundation for the one-epoch result (Proposition 4) and hence the main performance guarantee for BIRT (Theorem 3). To be more precise, the property is crucial in proving the existence of a hindsight optimal policy that is close to the fluid policy $\mathbf{F}$ and hence to the thresholded fluid policy $\hat{\mathbf{F}}$ used in practice as part of BIRT; the closeness of the said policies can then be leveraged to argue the possibility of recovering per-type per-agent aggregate allocations $N_\theta x^i_\theta$ by the hindsight optimal policy.

We first invoke the classical literature to generalize the concept of Lipschitz continuity to multi-functions (i.e., set-valued functions). In the generalization, the usual Euclidean distance in $\mathbb{R}^m$ based on the norm $\|\cdot\|$ is replaced by the Pompeiu-Hausdorff metric d_{PH} on the space of all nonempty and compact subsets of the Euclidean space $\mathbb{R}^m$, i.e., $d_{\mathsf{PH}}(G_1, G_2) = \min\{\varepsilon \geq 0 : G_1 \subseteq G_2 + D_\varepsilon(0), G_2 \subseteq G_1 + D_\varepsilon(0)\}$.

Definition 5 (Lipschitz continuity of multi-functions [85])**.** *A multi-function $\Gamma : \mathbb{R}^d \rightrightarrows \mathbb{R}^m$ is* Lipschitz continuous *on $V \subseteq \mathbb{R}^d$ if $\Gamma(v)$ is nonempty and compact for every $v \in V$, and there is a constant $L \geq 0$ (the* modulus *of Lipschitz continuity) such that for any $v_1, v_2 \in V$,*

$$d_{\mathsf{PH}}(\Gamma(v_1), \Gamma(v_2)) \leq L \|v_1 - v_2\|,$$

i.e., for every $g_1 \in \Gamma(v_1)$, there is some $g_2 \in \Gamma(v_2)$ such that $\|g_1 - g_2\| \leq L \|v_1 - v_2\|$.

We now proceed to show the proximity of hindsight optimal policies to fluid policies by arguing the optimal solution multi-function for the following parameterized optimization problem is Lipschitz continuous. The offline problem under the egalitarian welfare, parameterized by the initial utilities U_0 and arrivals $N \in \mathbb{Z}_+^{\Theta_P} \subseteq \mathbb{R}_+^{\Theta_P}$ can be formulated as follows.

$$\begin{aligned}
\mathsf{OPT}(N; U_0) = \underset{\Upsilon \in \mathbb{R}, \boldsymbol{x} \in \mathbb{R}^{|\Theta_P| \times n}}{\text{maximize}} \quad & \Upsilon \\
\text{subject to} \quad & U_0^i + \sum_\theta N_\theta u_\theta^i(x_\theta^i) - \Upsilon \geq 0, && i \in [n] \\
& \sum_{i \in [n]} x_\theta^i = 1, && \theta \in \Theta \\
& x_\theta^i \geq 0, && i \in [n], \theta \in \Theta.
\end{aligned} \tag{1.12}$$

Lemma 10. *Fix a finite set of types $\Theta_P \subseteq \Theta$, i.e., $|\Theta_P| < \infty$, and initial utilities $U_0 \in \mathbb{R}_+^n$. Suppose $u_\theta^i(\cdot)$ is an admissible utility function for almost every type $\theta \in \Theta_P$ and every agent i. Then the multi-function $\Gamma : \mathbb{R}^{\Theta_P} \rightrightarrows \mathbb{R}^{|\Theta_P| \times n}$ given by*

$$\Gamma(N) := \left\{(N_\theta x_\theta^i : \theta \in \Theta_P, i \in [n]) : (\Upsilon, \boldsymbol{x}) \text{ is an optimal solution to } \mathsf{OPT}(N; U_0) \text{ for some } \Upsilon\right\}$$

is Lipschitz continuous on $\mathbb{R}_+^{\Theta_P}$ *with modulus* $L = 2(|\Theta_P| + 1)/\underline{\beta}$*, where* $\underline{\beta} := \min\{(u_\theta^i)'(1) : \theta \in \Theta_P, i \in [n]\} \in (0, 1]$.

In other words, given any $N, \hat{N} \in \mathbb{R}_+^{\Theta_P}$*, if* $\boldsymbol{x}$ *is an optimal solution to the offline problem* $\mathsf{OPT}(N; U_0)$ (1.12)*, then there is an optimal solution* $\hat{\boldsymbol{x}}$ *to the offline problem* $\mathsf{OPT}(\hat{N}; U_0)$ *such that*

$$\max_{\theta\in\Theta_P} \left\|\hat{N}_\theta \hat{x}_\theta - N_\theta x_\theta\right\|_\infty \leq L \left\|\hat{N} - N\right\|_\infty .$$

The proof can be found in Section A.1.15. It is inspired by Theorem 10.5 in Schrijver [86], which states in linear programs the optimal solution multi-function is Lipschitz against RHS of the linear constraints. The main idea of their proof is to argue by Farkas' lemma the feasibility of a linear system constructed by augmenting the original LP with inequalities with Lipschitz perturbations. In contrast, although we can invoke a generalized version of Farkas' lemma (Lemma 31), the non-linearity of utility functions u_θ^i still poses a challenge to analyzing our offline problem (1.12) in a similar manner. Fortunately, the non-linearity can be absorbed into the feasible set of $\mathcal{Y} = \{\boldsymbol{y} : y_\theta^i = u_\theta^i(x_\theta^i), \boldsymbol{x} \in \Delta_n^{\Theta_P}\}$, so that the constraints are now linear. An important observation is that the feasible set $\mathcal{Y}$ is full dimension, the generalized Farkas' lemma acts as if on a convex cone, essentially the entire positive hyper-quadrant. The proof can then proceed as if the optimization were an LP.

Lemma 10 extends Theorem 10.5 in Schrijver [86], a sensitivity result in linear programming attributed to [87, 88, 89], to our problem where utility functions can be non-linear. In addition, we provide an explicit value of the Lipschitz modulus, providing a considerably cleaner and interpretable result than the said theorem, whose Lipschitz modulus involves the largest entry of the inverse matrices of all possible basis matrices of the linear program.

More broadly speaking, our result contributes to a stream of literature that studies the Lipschitz continuity of the solution multi-function [90] and similar properties. Indeed, research has been done on pseudo-Lipschitz continuity (a.k.a., the Aubin property, a local version of the Lipschitz continuity) [91], upper Lipschitz continuity (a "one-point" version of the vanilla "two-point" Lipschitz continuity) [92], and calmness (a local one-point version) [93]. Researchers have also studied generalizations such as the Hölder continuity of solutions [94, 95]. We refer the reader to [96, 85] for comprehensive reviews of relevant concepts and results.

Regularity and the Need for Thresholding

Recall any optimal solution $\boldsymbol{x}_0$ to the initial fluid problem $J(0)$ (1.5) is regular if the gradients of all active constraints at $\boldsymbol{x}_0$ are linearly independent. In the case where all utilities are linear and the optimization problem is an LP, this reduces to the non-degeneracy condition, namely every basic variable is nonzero. Drawing an analogy to LP, we refer to the set of linearly independent active constraints as a "basis". We say the initial fluid problem $J(0)$ (1.5) as *regular* if every optimal solution is regular.

When the initial fluid problem is regular, the same set of linearly independent constraints are active, or, the same "basis" is optimal, both in the offline problem (1.12) when the number of arrivals of each type N_θ is close to its expected value Tp_θ and in the updated fluid problem (1.7) when the cumulative agents' utilities U_t^i follow likely trajectories. Therefore, with a high probability, the offline problem and the updated fluid problems have the same set of active constraints (the same optimal "bases") at optimal solutions, which leads to similar allocations. In this case, the re-solving BIR policy suffices to achieve a low regret without thresholding.

Thresholding plays a key role when any initial fluid problem is not regular (or nearly irregular) and active constraints are first-order redundant. Slight perturbations in the number of arrivals of each type or the agents' utilities can lead to changes in the optimal "basis" of the offline problem (1.12). A vanilla re-solving policy without thresholding could very likely introduce welfare losses by choosing active constraints different from those of the hindsight optimal policy. Intuitively, by adjusting small allocation components to zero, thresholding essentially forces adopting an irregular feasible solution to the fluid problem that simultaneously corresponds to multiple "bases". Thresholding postpones selecting any "basis" by placing BIRT on multiple bases that are sufficiently close by so that past sub-optimal allocation decisions could be made up for later once the final choice of basis is revealed. In a word, the gist of thresholding is to *postpone choosing any* *of neighboring "bases" until subsequent arrivals reveal more information.*

An alternative way to understand thresholding is through our theoretical regret analysis in Section 1.3.3. The intuition behind thresholding is that BIRT aims to recover the hindsight optimum by the end of the horizon; in particular and more ambitiously, it aspires to replicate the exact amount of resources of each type θ allocated to each agent i by the hindsight optimal policy. Though re-solving helps approximate these hindsight optimal allocations, exact replication is hindered by stochasticity, especially for (θ, i) where the fluid policy yields allocations F^i_θ is close to zero. As a result of stochasticity, the hindsight optimal policy may possibly end up allocating none of type-θ resources to agent i, in which case the replication goal fails completely as allocations are irrevocable. Hence, it is best to altogether adjust these allocations to zero, leaving more leeway for better informed allocations in the future.

Technical novelty

Before ending this section, we would like to comment on some key distinctions between our problem and analyses from those in Bumpensanti and Wang [48], whose work in network revenue management inspired the BIRT policy.

As discussed in Section 1.3.2, the most salient feature of our problem is that its objective is non-linear and hence non-separable over time. This renders inapplicable a direct decomposition of the horizon into several, disjoint re-solving epochs and plugging in single-epoch regret guarantees (assuming memorylessness of the allocations), as is done by [48]. Rather, in our problem, each re-solving epoch starts with the random utilities garnered from stochastic arrivals in previous epochs, and these initial conditions have to be taken into account in epoch-wise analyses.

In addition to linear utilities, as in Bumpensanti and Wang [48], our problem setup allows for non-linear utility functions, and we provide provide low-regret performance guarantees in that case.

Another characteristic of the dynamic fair allocation problem is that the central planner needs to assign resources across multiple agents, i.e., allocation decisions lie in the simplex Δ_n across the n agents. In the network revenue management problem, the decision maker needs to decide whether to fulfill demand requests, i.e., allocation decisions are binary accept-reject. This discrepancy complicates the thresholding procedure as thresholding one agent's assignment affects the decisions of all other agents.

Our BIRT policy is highly flexible in its re-solving schedule as well as thresholding adjustments, compared with the IRT policy by [48]. By defining the re-solving schedule by the scheduling function f and the time remaining until the end of the horizon, we are able to better attribute parts of the total welfare loss to epochs across the horizon. We also provide more granular guarantees and understanding of the regimes of re-solving frequencies.

In analyzing the performance, Bumpensanti and Wang [48] invoke existing results by Reiman and Wang [49] for network revenue management problems, whereas we provide a new and simpler stand-alone analysis that exploits the structure of our problem. Furthermore, in the Lipschitz continuity result, we provide an explicit Lipschitz modulus that is clearly interpretable; in contrast, their result involves enumerating the inverse matrices of all possible basis matrices of the fluid problem, which can be computationally intense. Our analysis also facilitates a more fundamental understanding of the impact of stochastic arrivals on the optimal solution through a new sensitivity analysis and of the coupling between the hindsight optimal solution and auxiliary policies.

Lastly, we would also like to briefly compare our work with Freund and Banerjee [54], who study the online decision-making problem whose objective depends only on aggregate counts of different actions taken over the horizon in response to each arrival type. With a *discrete action space* that translates to indivisible resources, they propose algorithms that dynamically re-solves the empirical fluid problem and triggers re-solving by hitting lower confidence bounds on the expected allocations, and show they attain an $O(1)$ regret. We note that the result in [54] relies heavily on the discreteness and finiteness of the action space. In the case of linear utilities, their algorithm can be directly applied to our problem with algorithm as it suffices to count the number of resources the central planner allocates to each agent. When utilities are non-linear and resources are divisible, however, the decision maker needs to divide resources across agents optimally, and focusing on integral decisions (only allocating one resource to each agent) would lead to poor performance; at times this can even lead to linear regret (see Experiment 1.4.3).

1.4 Experiments

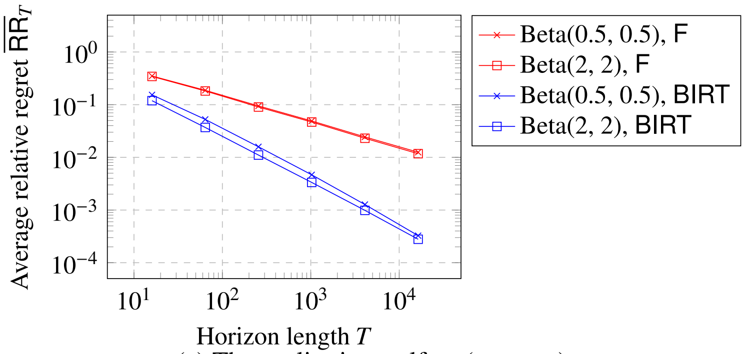
In this section, we present numerical results from our simulation experiments to illustrate the performances of the fluid F, BIR and BIRT policies. The first experiment evaluates these policies on several randomized arrival distributions under the egalitarian welfare, the harmonic-mean welfare and the Nash social welfare, assuming all agents have linear utilities. The second experiment focuses on the special case of the egalitarian welfare and exposes the impact of degeneracy on re-solving heuristics. It is built upon two artificially designed arrival distributions, whose fluid problems are degenerate and non-degenerate respectively.

1.4.1 Randomized Experiment Under Various Welfare Metrics

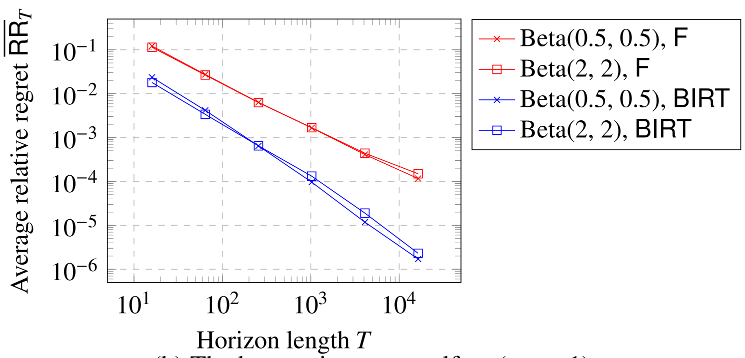
In the first experiment, we illustrate the performances of the F and BIRT policies under the egalitarian welfare, the harmonic-mean welfare and the Nash social welfare, using a number of randomized distribution instances. More specifically, we consider a problem setup with $n = 4$ agents and $|\Theta_P| = 5$ types of resources. We simulate probabilities of the resource types uniformly at random from the simplex, i.e., $(p_\theta : \theta \in \Theta) \sim \mathsf{Unif}(\Delta_5)$, and the linear utility functions with marginal utilities independently and identically from a parameterized Beta distribution, i.e., $\beta^i_\theta \overset{\text{iid}}{\sim} \mathsf{Beta}(\cdot, \cdot)$ across θ, i. In particular, we make the two choices for the utility functions:

- $\beta^i_\theta \sim \mathsf{Beta}(0.5, 0.5)$, under which values are concentrated around the two extremes of 0 and 1, or in other words, the agents have dichotomous valuations for the resources;
- $\beta^i_\theta \sim \mathsf{Beta}(2.0, 2.0)$, under which values are concentrated near the middle at 0.5, or in other words, the agents have similar valuations for the resources.

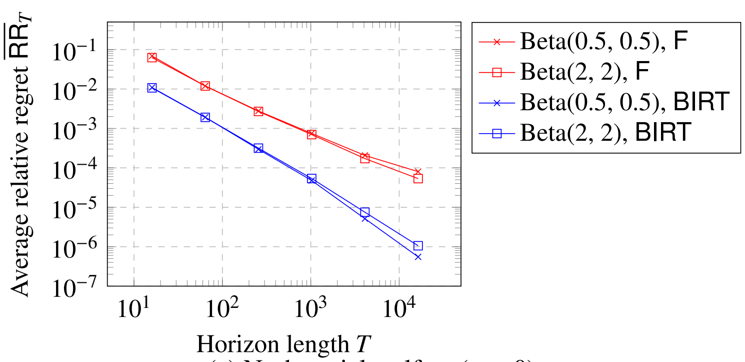
In each case, we simulate 30 different arrival distributions P and compute the average relative regrets over the randomized arrival distributions. Note that relative regrets are used to normalize the problems for a reasonable comparison regardless of welfare dimensions. Results are shown in Figure 1.8, where the meta-distributions for generating arrival distributions are distinguished by marks, and the policies are distinguished by colors.



(a) The egalitarian welfare ($q = -\infty$)



(b) The harmonic-mean welfare ($q = -1$)



(c) Nash social welfare ($q = 0$)

Figure 1.8: Average relative regret $\overline{\rho}(\pi)$ over randomized arrival distributions, under the egalitarian welfare ($q = -\infty$), the harmonic-mean welfare ($q = -1$) and the Nash social welfare ($q = 0$), plotted against the horizon length T on a double logarithmic scale

We summarize our findings under the three welfare metrics.

- Figure 1.8a shows that under the egalitarian welfare ($q = -\infty$), the F policy attains a relative regret of $\mathsf{RR}_T(\mathsf{F}) = \Theta(T^{-1/2})$, translating to a regret of $\mathcal{R}_T(\mathsf{F}) = \Theta(\sqrt{T})$; the BIRT policy attains a relative regret of $\mathsf{RR}_T(\mathsf{BIRT}) = \Theta(T^{-1})$, translating to a uniformly bounded regret of $\mathcal{R}_T(\mathsf{BIRT}) = \Theta(1)$.

- Figure 1.8b shows that under the harmonic-mean welfare ($q = -1$), the F policy attains a relative regret of $\mathsf{RR}_T(\mathsf{F}) = \Theta(T^{-1})$, translating to a regret of $\mathcal{R}_T(\mathsf{F}) = \Theta(1)$; the BIRT policy attains a relative regret of $\mathsf{RR}_T(\mathsf{BIRT}) = o(T^{-1})$, translating to a vanishing regret of $\mathcal{R}_T(\mathsf{BIRT}) = o(1)$.

- Figure 1.8c shows that under the Nash social welfare ($q = 0$), the F policy attains a relative regret of $\mathsf{RR}_T(\mathsf{F}) = \Theta(T^{-1})$, translating to a regret of $\mathcal{R}_T(\mathsf{F}) = \Theta(1)$; the BIRT policy attains a relative regret of $\mathsf{RR}_T(\mathsf{BIRT}) = o(T^{-1})$, translating to a vanishing regret of $\mathcal{R}_T(\mathsf{BIRT}) = o(1)$.

These empirical findings from the randomized experiments exactly match out theoretical guarantees. Interestingly, we notice the BIRT policy attains a resounding $o(1)$ vanishing regret under the harmonic-mean and Nash social welfare metrics, an even more impressive improvement upon the $\Theta(1)$ regret attained by the static F policy. The precise mechanism to attaining the vanishing regret is unknown to us, but the empirical findings do highlight the excellent performance that BIRT has under all welfare metrics.

1.4.2 Egalitarian Welfare: Linear Utilities

In the second experiment, we zoom in and focus on the egalitarian welfare and the case where all agents have linear utilities. We aim to corroborate our claims on the performances of the fluid and BIRT policies under the egalitarian welfare and, moreover, to understand what driving forces are behind the outstanding performance of BIRT. To be more precise, we are including the following heuristic policies for comparison.

1. F: the fluid type-based policy based on the initial fluid problem (Algorithm 1).

2. BIR: the Backward Infrequent Re-solving policies (Algorithm 2).

3. BIRT: the Backward Infrequent Re-solving with Thresholding policies (Algorithm 3).

4. FR: the Frequent Re-solving policy, a specific example of BIR that re-solves the fluid problems in every period before making allocation decisions.

Comparing the BIR and BIRT policies can illustrate the crucial role that thresholding is playing in improving performance when the initial fluid solution is not regular, and the FR policy is included to evaluate whether performance could be improved by more frequent re-solving.

In addition to inter-policy comparison, we would like to investigate numerically how the absence of regular (a.k.a., non-degenerate in LP) initial fluid solutions can have on the performance of re-solving heuristics, as widely discussed in previous literature.

To illustrate the impact of degeneracy, we consider a simple problem setup with $n = 2$ agents and $|\Theta_P| = 2$ types of resources where the utility functions are linear with marginal utilities given by $\beta_1 = (1, 1/2)$ and $\beta_2 = (1/2, 1)$. We then consider two arrival distributions $(p_1, p_2) = (1/2, 1/2)$ and $(p_1, p_2) = (2/5, 3/5)$. The first case (detailed in Section A.1.4) gives rise to a degenerate fluid problem, as a total of $6 > 5$ constraints are active at the fluid policy $\mathbf{F}$ with $(\mathsf{F})_1 = (1, 0)$ and $(\mathsf{F})_2 = (0, 1)$, corresponding to the two bases $(\Upsilon, x_1^1, x_2^1, x_2^2)$ and $(\Upsilon, x_1^1, x_1^2, x_2^2)$. Therefore, slight perturbations on the number of arrivals can lead to optimal hindsight allocations of the form $(x)_1 = (1, 0)$ and $(x)_2 = (\epsilon, 1 - \epsilon)$ or $(x)_1 = (1 - \epsilon, \epsilon)$ and $(x)_2 = (0, 1)$ for $\epsilon \geq 0$, and, without thresholding, the re-solving policy might allocate resources to the wrong agent (e.g., allocating type-1 resource to agent 2 when the hindsight optimal basis is the first one). In the second case, the original fluid problem is non-degenerate as a total of 5 constraints are active at the fluid policy $\mathbf{F}$ with $(\mathsf{F})_1 = (1, 0)$ and $(\mathsf{F})_2 = (2/9, 7/9)$, corresponding to the unique basis $(\Upsilon, x_1^1, x_2^1, x_2^2)$. Therefore, with high probability, an optimal hindsight allocation is of the form $(x)_1 = (1, 0)$ and $(x)_2 = (2/9 \pm \epsilon, 7/9 \mp \epsilon)$ and it is less likely to allocate resources to the wrong agent. In both cases, the BIRT and BIR policies have the same exponential-geometric re-solving schedule determined by the scheduling function $f(t) = t^{1/1.05}$. Regrets of respective policies are shown in Figure 1.9.

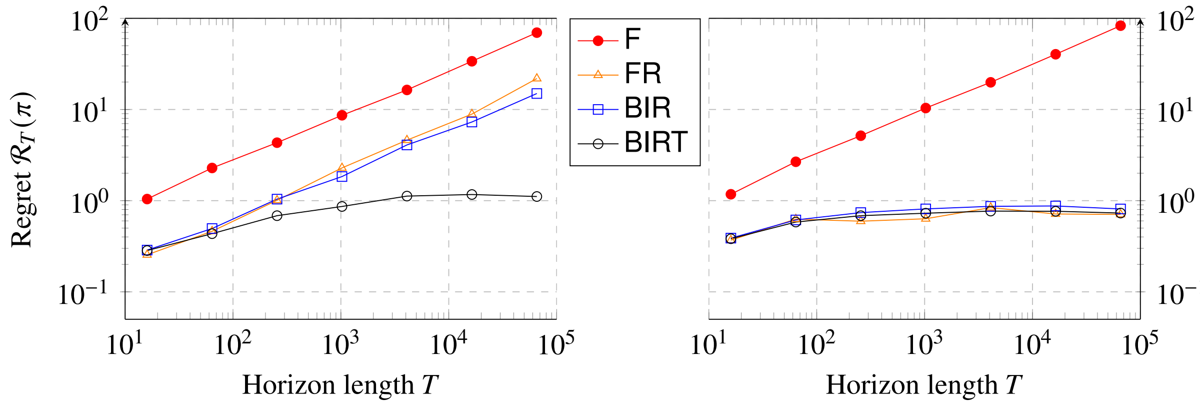


Figure 1.9: Comparison of regrets of F, BIR, FR, and BIRT policies under linear utilities on a double logarithmic scale. Results for the degenerate instance are shown on the left, and those for the non-degenerate instance are shown on the right.

First, we observe from Figure 1.9 that in both cases BIRT's regret converges, corroborating our claim (Theorem 3) that it attains $O(1)$ regret. Equipped with distributional knowledge, the central planner is shown to be able to improve the performance upon the fluid type-based policy by re-solving infrequently (BIR policy), either by improving the factor of $\sqrt{T}$ in the regret in the degenerate case, or by attaining a constant regret in the non-degenerate case. Interestingly, in both cases, further increasing the re-solving frequency does not lead to performance improvement; in fact, the BIR and FR policies attain similar regrets. In other words, carefully designing the re-solving schedule suffices to guarantee good performance while drastically reducing computational complexity. The most significant difference between the two cases is that the BIR and FR policies attain $\Theta(\sqrt{T})$ regret in the degenerate case, whereas they can attain $O(1)$ regret in the non-degenerate case, indicating their vulnerability against degeneracy. By contrast, the more robust BIRT policy steadily attains a $O(1)$ regret in both cases, demonstrating that the thresholding rule is crucial for controlling the excessive variation in allocations and obtain remarkable performance even when the initial fluid solution is non-degenerate.

1.4.3 Egalitarian Welfare: Non-linear Utilites

We next investigate the performances of the policies under the egalitarian welfare when utility functions are non-linear. In particular, we consider the setup with $n = 2$ agents and $|\Theta_P| = 2$ types of resources where the utility functions are of the form $u^i(x) = \sqrt{1 + \beta^i x} - 1$ with two types given by $\beta_1 = (1, 1/2)$ and $\beta_2 = (1/2, 1)$. As before, we consider two arrival distributions $(p_1, p_2) = (1/2, 1/2)$ and $(p_1, p_2) = (2/5, 3/5)$. The first case gives rise to an irregular fluid problem, as the active constraints are linearly dependent at the fluid policy F with $(\mathsf{F})_1 = (1, 0)$ and $(\mathsf{F})_2 = (0, 1)$. In the second case, the initial fluid problem is regular as the active constraints at the fluid policy F with $(\mathsf{F})_1 = (1, 0)$ and $(\mathsf{F})_2 \approx (0.23, 0.77)$ are linearly independent.

In addition to the fluid F, BIR and BIRT policies, we include the BF algorithm proposed in [54], with a slight improvement of substituting the known underlying distribution in its optimization step, and a coefficient of $m = 0.5$ (see Algorithm 4). To be more precise, we are adopting BIR and BIRT policies with midpoint re-solving schedules determined by the scheduling function $f(t) = \max(1, \lfloor t/2 \rfloor)$, and BIRT has a threshold coefficient of 1. Results are summarized in Figure 1.10.

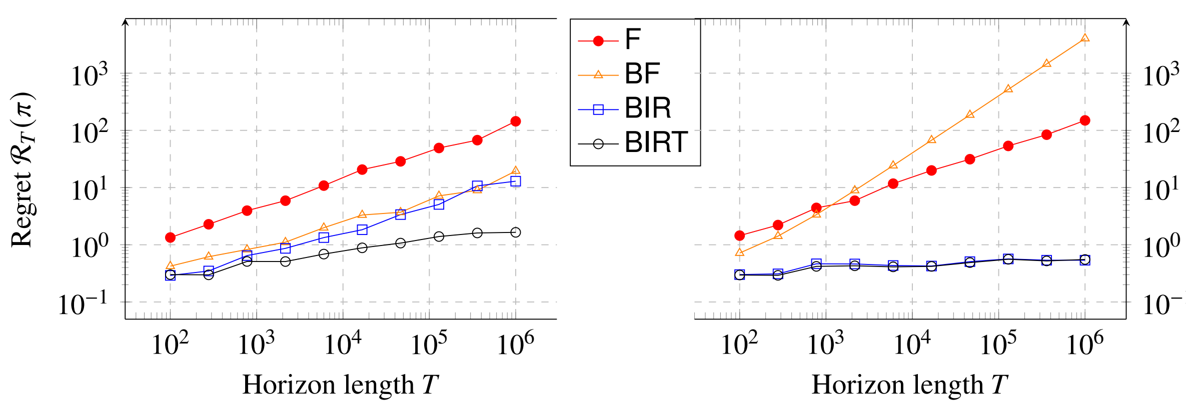


Figure 1.10: Comparison of regrets of F, BF, BIR, and BIRT policies under non-linear utilities on a double logarithmic scale. Results for the irregular instance are shown on the left, and those for the regular instance are shown on the right.

First, the BIRT policy is shown to have a rather low regret in both cases, signifying its outstanding performance regardless of whether the initial fluid problem is regular. The BIR policy almost performs identically to BIRT, believably because thresholding adjustments do not occur in

Algorithm 4 Freund and Banerjee [54] (BF)

Input: horizon length $T \in \mathbb{N}$, welfare metric $w : \mathbb{R}_+^n \to \mathbb{R}$, positive coefficient $m > 0$, admissible arrival distribution $P \in \mathcal{P}$ supported on Θ_P.

Initialize: set $U_0^i \leftarrow 0$ for $i \in [n]$

set cumulative allocations $\xi_\theta^i \leftarrow 0$ for $i \in [n]$ and $\theta \in \Theta_P$

set allocation allowances $A_\theta^i \leftarrow 0$ for $i \in [n]$ and $\theta \in \Theta_P$

for $t = 1, \ldots, T$ **do**

  if $A_{\theta_t}^i < 1$ *for every agent i* **then**

    solve the updated fluid problem with updated utilities

$$\mathsf{F} \in \arg\max \left\{ w\left(U_{t-1} + (T - t + 1)\mathbb{E}_{\theta \sim P}\left[u_\theta(x_\theta)\right]\right) : \boldsymbol{x} \in \Delta_n^{\Theta_P} \right\};$$

    set time remaining $\bar{t} \leftarrow T - t + 1$

$$A_\theta^i \leftarrow \left\lfloor \mathsf{F}_\theta^i \bar{t} - \xi_\theta^i - m\bar{t}^{3/4}\sqrt{\log \bar{t}} \right\rfloor \qquad \text{for any } \theta, i.$$

  observe incoming resource type $\theta_t \sim P$

  allocate $x_t \leftarrow \mathrm{e}^j$ for some $j \in \arg\max A_{\theta_t}^j$

  reduce allowance $A_{\theta_t}^j \leftarrow A_{\theta_t}^j - 1$

  increase cumulative allocation $\xi_{\theta_t}^j \leftarrow \xi_{\theta_t}^j + 1$

  update utilities $U_t^i \leftarrow U_{t-1}^i + u_t^i(x_t^i)$ for every agent $i \in [n]$

Output: allocations $(x_t : t \in [T])$ and final welfare $w(U_T^i)$.

the regular case; nevertheless, the BIR policy incurs a quite large regret of order $O(\sqrt{T})$ in the irregular case, comparably to the naïve fluid policy, which attains the vanilla $\sqrt{T}$ regret in both cases. The observations match what we expect from a theoretical perspective—the regularity condition is essential to the good performance of BIR, whereas BIRT is immune.

We also notice Algorithm 4 from [54] incurs a regret of order $O(\sqrt{T})$ when the initial fluid solution is not regular, but one of order as high as $O(T)$ otherwise. This is largely due to the design of the algorithm that relies heavily on the finiteness of the action space, which in our case means always allocating entire resources to single agents as if they are indivisible. In the first case, this is precisely what the fluid policy does, which explains the resemblance between the performances of the two policies. In the second, however, even the fluid policy makes divided allocations across agents, and this leads to a non-trivial constant utility loss in every single period by Algorithm [54]. Hence, the algorithm itself is not directly applicable to our problem when utilities are non-linear.

1.4.4 Robustness of BIRT

We complement the numerical findings with a last experiment to help illustrate the robustness of BIRT against not only whether the fluid solution is regular, but how close it is to any irregular point. To this end, we further exploit the same experimental setup in the preceding section, and compare the performances of BIR and BIRT when the fluid problem is nearly irregular.

Consider the same utility functions of the form $u^i(x) = \sqrt{1 + \beta^i x} - 1$ with two types given by $\beta_1 = (1, 1/2)$ and $\beta_2 = (1/2, 1)$. We zoom in and investigate performances when p_1 varies around 0.5, the value that gives rise to an irregular fluid solution. Figure 1.11 shows a huge spike in the regret of BIR in the irregular case as well as elevated regret nearby. In contrast, the response of BIRT remains muted in and around the irregular case, marking its robustness.

1.5 Conclusion

In this work we study dynamic policies for allocating T divisible items arriving sequentially to a fixed set of n agents with possibly non-linear and concave utility functions. We consider a

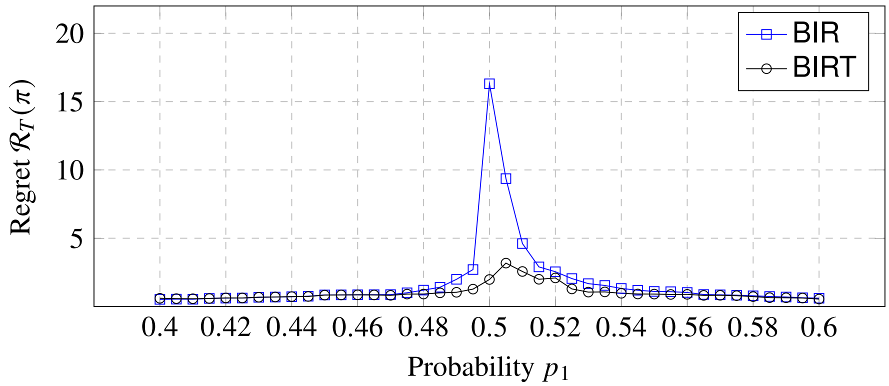


Figure 1.11: Regrets of the BIR and BIRT policies when the initial fluid problem is nearly irregular, for $T = 1,000,000$. Both policies have the midpoint schedule, i.e., $f(t) = \max(1, \lfloor t/2 \rfloor)$.

collection of Hölder-mean welfare metrics, including the Nash social welfare and the egalitarian welfare, and evaluate policies based on the expected welfare generated, which entails addressing the trade-off between efficiency and fairness. We show the classical fluid policy based on certainty equivalents incurs a $\Theta(\sqrt{T})$ regret against the expected hindsight optimal egalitarian welfare and an $O(1)$ regret under all other welfare metrics. For the egalitarian welfare, we propose the Backward Infrequent Re-solving (BIR) and Backward Infrequent Re-solving with Thresholding (BIRT) policies. We show under a mild regularity condition, the BIR policy attains an asymptotic regret of order $O(1)$ when all utilities are linear and of order $O(\log T)$ otherwise; the BIRT policy, in contrast, does not require the regularity condition, and attains a comparable asymptotic regret of order $O(1)$ when all utilities are linear and of order $O(\log^{2+\varepsilon} T)$ otherwise.

Chapter 2: Evaluating Model Performance Under Worst-case Sub-populations

2.1 Introduction

[1]The training population typically does not accurately represent what the model will encounter under operation. Model performance has been observed to substantially degrade under distribution shift [97, 98, 99, 100, 101] in speech recognition [102], automated essay scoring [103], and wildlife conservation [104]. Similar trends persist for state-of-the-art NLP and computer vision models [105, 106], even on new data constructed under a near-identical process [107, 108]. Heavily engineered commercial models are no exception [109], performing poorly on rare entities in named entity linking and examples that require abstraction and distillation in summarization tasks [110].

A particularly problematic form of distribution shift comes from embedded power structures in data collection. Data forms the infrastructure on which we build prediction models [111], and they inherit socioeconomic and political inequities against marginalized communities. For example, out of 10,000+ cancer clinical trials the National Cancer Institute funds, less than 5% of participants were non-white [112]. Typical models replicate and perpetuate such bias, and their performance drops significantly on underrepresented groups. Speech recognition systems work poorly for Blacks [102] and those with minority accents [113]. More generally, model performance degrades across demographic attributes such as race, gender, or age, in facial recognition, video captioning, language identification, and academic recommender systems [114, 115, 116, 117, 118, 109].

[1]This work is collaborative with Mike Li and Dr. Hongseok Namkoong at the Graduate School of Business at Columbia University.

Model training typically relies on varied engineering practices. It is crucial to *rigorously certify* model robustness prior to deployment for these heuristic approaches to bear fruit and transform consequential applications. Ensuring that models perform uniformly well across subpopulations is simultaneously critical for reliability, fairness, satisfactory user experience, and long-term business goals. While a natural approach is to evaluate performance across a small set of groups, disadvantaged subpopulations are hard to define a priori because of *intersectionality*. The most adversely affected are often determined by a complex combination of variables such as race, income, and gender [109]. For example, performance on summarization tasks varies across demographic characteristics and document specific traits such as abstractiveness, distillation, and location and dispersion of information [110].

Motivated by these challenges, we study the worst-case subpopulation performance across *all* subpopulations of a given size. This conservative notion of performance evaluates robustness to unanticipated distribution shifts in Z, and automatically accounts for complex intersectionality by virtue of being agnostic to demographic groupings. Formally, let Z be a set of core attributes that we wish to guarantee uniform performance over. It may include protected demographic variables such as race, gender, income, age, or task-specific information such as length of the prompt or metadata on the input; notably, it can contain any continuous or discrete variables. We let $X \in \mathcal{X}$ be the input / covariate, and $Y \in \mathcal{Y}$ be the label. In NLP and vision applications, X is high-dimensional and typically $\dim(Z) \ll \dim(X)$.

We use $\theta(X)$ to denote a fixed prediction model and consider flexible and abstract losses $\ell(\theta(x); y)$. Our goal is to ensure that the model θ performs well over all subpopulations defined over Z. We evaluate model losses on a mixture component, which we call a subpopulation. Postulating a lower bound $\alpha \in (0, 1]$ on the demographic proportion (mixture weight), we consider the set of subpopulations of the data-generating distribution P_Z

$$\mathcal{Q}_\alpha := \left\{Q_Z \mid P_Z = aQ_Z + (1-a)Q'_Z \text{ for some } a \geq \alpha, \text{ and subpopulation } Q'_Z\right\}. \tag{2.1.1}$$

The demographic proportion (mixture weight) a represents how underrepresented the subpopulation is under the data-generating distribution P_Z. Before deploying the model θ, we wish to evaluate the worst-case subpopulation performance

$$\mathsf{W}_\alpha(\theta) := \sup_{Q_Z \in \mathcal{Q}_\alpha} \mathbb{E}_{Z \sim Q_Z} \left[\mathbb{E}[\ell(\theta(X), Y) \mid Z] \right]. \tag{2.1.2}$$

The worst-case subpopulation performance (2.1.2) guarantees uniform performance over subpopulations (2.1.1) and has a clear interpretation that can be communicated to diverse stakeholders. The minority proportion α can often be chosen from first principles, e.g., we wish to guarantee uniformly good performance over subpopulations comprising at least $\alpha = 20\%$ of the collected data. Alternatively, it is often informative to study the threshold level of $\alpha^\star$ when $\alpha \mapsto \mathsf{W}_\alpha(\theta)$ crosses the *maximum level of acceptable loss*. The threshold $\alpha^\star$ provides a *certificate of robustness* on the model $\theta(\cdot)$, guaranteeing that all subpopulations large than $\alpha^\star$ enjoy good performance.

We develop a principled and scalable procedure for estimating the worst-case subpopulation performance (2.1.2) and the certificate of robustness $\alpha^\star$. A key technical challenge is that for each data point, we observe the loss $\ell(\theta(X); Y)$ but never observe the conditional risk evaluated at the attribute Z

$$\mu^\star(Z) := \mathbb{E}[\ell(\theta(X); Y) \mid Z]. \tag{2.1.3}$$

In Section 2.2, we propose a two-stage estimation approach where we compute an estimate $\widehat{\mu}_1(\cdot)$ of the conditional risk $\mu^\star(\cdot)$. Then, we compute a plug-in estimate of the worst-case subpopulation performance under $\widehat{\mu}_1(\cdot)$ using a dual reformulation of the worst-case problem (2.1.2). We show several theoretical guarantees for our estimator of the worst-case subpopulation performance (2.1.2). Our first finite-sample result (Section 2.3.1) shows convergence at the rate $O_p\left(\sqrt{\mathfrak{Comp}_n(\mathcal{H})/n}\right)$, where $\mathfrak{Comp}_n$ denotes a notion of complexity for the model class estimating the conditional risk (2.1.3).

In some applications, it may be natural to define Z using images or natural languages describing the input and use deep networks to predict the conditional risk (2.1.3). As the complexity term

$\mathfrak{Comp}_n(\mathcal{H})$ becomes prohibitively large in this case [119, 120], our second result (Section 2.3.2) shows data-dependent *dimension-free* concentration of our two-stage estimator: our bound only depends on the complexity of the model class $\mathcal{H}$ through the out-of-sample error for estimating the conditional risk (2.1.3). This error can be made small using overparameterized deep networks, allowing us to estimate the conditional risk (2.1.3) using even the largest deep networks and still obtain a theoretically principled upper confidence bound on the worst-case subpopulation performance. Leveraging these guarantees, we develop principled procedures for estimating the certificates of robustness $\alpha^\star$ in Section 2.3.3.

In Section 2.4, we demonstrate the effectiveness of our procedure on real data. By evaluating model robustness under subpopulation shifts, our methods allow the selection of robust models before deployment as we illustrate using the recently proposed CLIP model [121].

Related work. The long line of works on distributionally robust optimization (DRO) aims to *train models* to perform well under distribution shifts. Previous approaches considered finite-dimensional worst-case regions such as constraint sets [122, 123, 124] and those based on notions of distances for probability measures such as f-divergences [125, 126, 127, 128, 129, 130, 131], Levy-Prokhorov [132], Wasserstein distances [133, 134, 135, 136, 137, 138], and integral probability metrics based on reproducing kernels [139, 140]. The distribution shifts considered in these approaches are often contrived and difficult to interpret and often result in overly conservative models. Furthermore, these approaches do not currently scale to modern large-scale NLP or vision applications.

Our work is most closely related to Duchi *et al.* [141], who proposed algorithms for *training* models with respect to the worst-case subpopulation performance (2.1.2), which is a more ambitious goal than our narrower viewpoint of *evaluating* model performance pre-deployment. Their (full-batch) training procedure requires solving a convex program with n^2 variables per gradient step, which is often prohibitively expensive. Furthermore, training with respect to the worst-case conditional risk $\mathbb{E}[\ell(\theta(X); Y) \mid Z]$ do not scale to deep networks that can overfit to the training data [142]. By contrast, our evaluation perspective aims to take advantage of the rapid progress in deep learning. We build scalable evaluation methods that apply to arbitrary models, which allows leveraging state-of-the-art engineered approaches for training $\theta(\cdot)$. Our narrower focus on evaluation allows us to provide convergence rates that scale advantageously with the dimension of Z, compared to the nonparametric $O_p(n^{-1/d})$ rates for training [141]. Recently, Jeong and Namkoong [143] studied a similar notion of worst-case subpopulation performance in causal inference.

Our notion of worst-case subpopulation performance is also related to the by now vast literature on fairness in ML. We give a necessarily abridged discussion and refer readers to Barocas *et al.* [144] and Corbett-Davies and Goel [145] for a comprehensive treatment. A large body of work studies *equalizing* a notion of performance over fixed, pre-defined demographic groups for *classification tasks* [146, 147, 148, 149, 150, 151]. Kearns *et al.* [152, 153] and Hébert-Johnson *et al.* [154] consider finite subgroups defined by a structured class of functions over Z, and study methods of equalizing performance across them. By contrast, our approach instantiates Rawls' theory of distributive justice [16, 155], where we consider the allocation of the loss $\ell(\cdot;\cdot)$ as a resource. Rawls' difference principle maximizes the welfare of the worst-off group and provides incentives for groups to maintain the status quo [16]. Similarly, Hashimoto *et al.* [156] studied negative feedback loops generated by user retention—they use a more conservative notion of worst-case loss than ours—as poor performance on a currently underrepresented user group can have long-term consequences.

Our diagnostics complement the recent approaches to benchmarking under distribution shifts [100, 108, 106, 101, 157, 105, 107] as our procedure does not require out-of-distribution data. Since good performance on a particular distribution shift does not necessitate robustness, we evaluate models using the worst-case subpopulation performance (2.1.2).

2.2 Approach

We begin by contrasting our approach to standard alternatives that consider pre-defined, fixed demographic groups [158]. Identifying disadvantaged subgroups a priori is often challenging as they are determined by *intersections* of multiple demographic variables. To illustrate such complex

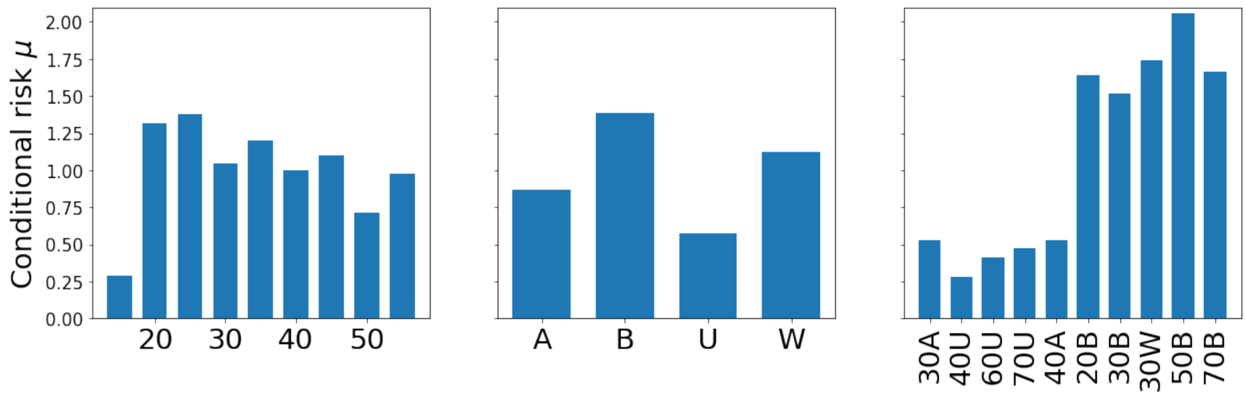


Figure 2.1. Conditional risk $\mu^\star(Z) = \mathbb{E}[(Y - \theta(X))^2 \mid Z]$. Here Z = *age* on the left panel, Z = *race* in the center, and Z = (*age*, *race*) on the right. A = Asian, B = Black, U = Unknown, W = White.

intersectionality, consider a drug dosage prediction problem for Warfarin [159], a common anti-coagulant (blood thinner). Taking the best prediction model for the optimal dosage on this dataset based on genetic, demographic and clinical factors [159], we present the squared error on the root dosage. In Figure 2.1, when age and race are considered *simultaneously* instead of *separately*, subpopulation performance vary significantly across intersectional groups.

The worst-case subpopulation performance (2.1.2) automatically accounts for latent intersectionality. It is agnostic to demographic groupings and allows considering infinitely many subpopulations that represent at least α-fraction of the training population P. By allowing the modeler to select arbitrary protected attributes Z, we are able to consider potentially complex subpopulations. For example, Z can even be defined with respect to a natural language description of the input X. The choice of Z—and subsequent worst-case subpopulation performance (2.1.2) of the conditional risk $\mu^\star(Z) = \mathbb{E}[\ell(\theta(X); Y) \mid Z]$—interpolates between the most conservative notion of subpopulations (when $Z = (X, Y)$) and simple counterparts defined over a single variable.

The choice of the subpopulation size α should be informed by domain knowledge—desired robustness of the system—and the dataset size relative to the complexity of Z. Often, proxy groups can be used for selecting α. If we wish to ensure good performance over patients of all races aged 50 years or older, we can choose α to be the proportion of the least represented $(race, age \geq 50)$ group—this leads to $\alpha = 5\%$ in the Warfarin data. The corresponding worst-case subpopulation performance (2.1.2) guarantees good performance over all groups of similar size.

When it is challenging to commit to a specific subpopulation size, it may be natural to postulate a *maximum level of acceptable loss* $\bar{\ell}$. To measure the robustness of a model, we define the smallest subpopulation size $\alpha^\star(\theta)$ for which the worst-case subpopulation performance is acceptable

$$\alpha^\star(\theta) := \inf\{\alpha : \mathsf{W}_\alpha(\theta) \le \bar{\ell}\}. \tag{2.2.1}$$

This provides a *certificate of robustness*: if $\alpha^\star(\theta)$ is large, then θ is brittle against even majority subpopulations; if it is sufficiently small, then θ performs well on underrepresented subpopulations.

We now derive estimators for the worst-case subpopulation performance (2.1.2) and the certificate of robustness (2.2.1), based on i.i.d. observations $(X_i, Y_i, Z_i)_{i=1}^n \sim P$. We assume our observations are independent from the data used to train the model $\theta(\cdot)$.

Dual reformulation The worst-case subpopulation performance (2.1.2) is unwieldy as it involves an infinite dimensional optimization problem over probabilities. Instead, we use its dual reformulation for tractable estimation. We denote $(\cdot)_+ = \max(\cdot, 0)$, and abuse notation by letting $\mathsf{W}_\alpha(h)$ be the worst-case subpopulation performance for $h(Z)$ (so that $\mathsf{W}_\alpha(\theta) = \mathsf{W}_\alpha(\mu^\star)$).

Lemma 11 (Shapiro *et al.* [160, Example 6.19] and Rockafellar and Uryasev [161])**.** *If* $\mathbb{E}[\mu(Z)_+] < \infty$*,*

$$\mathsf{W}_\alpha(h) := \sup_{Q_Z \in \mathcal{Q}_\alpha} \mathbb{E}_{Z \sim Q_Z}[\mu(Z)] = \inf_{\eta \in \mathbb{R}} \left\{ \frac{1}{\alpha} \mathbb{E}_P (\mu(Z) - \eta)_+ + \eta \right\}. \tag{2.2.2}$$

The dual optimum is attained at the $(1-\alpha)$-quantile of the $\mu(Z)$ [162, Theorem 10]. The dual (2.2.2) hence shows $\mathsf{W}_\alpha(\theta)$ is a tail-average of $\mu^\star(Z)$, a popular risk measure known as the conditional value-at-risk (CVaR) in portfolio optimization [161].

Algorithm 5 Two-stage procedure for estimating worst-case subpopulation performance (2.1.2)

1: INPUT: Subpopulation size α, model class $\mathcal{H}$, samples S_1 and S_2
2: On S_1, solve $\widehat{\mu}_1 \in \operatorname{argmin}_{\mu \in \mathcal{H}} \frac{1}{|S_1|} \sum_{i \in S_1} (\ell(\theta(X_i); Y_i) - \mu(Z_i))^2$.
3: On S_2, compute the plug-in estimator $\widehat{\mathsf{W}}_{\alpha,k}(\widehat{\mu}_1) = \inf_\eta \left\{ \frac{1}{\alpha |S_2|} \sum_{i \in S_2} (\widehat{\mu}_1(Z_i) - \eta)_+ + \eta \right\}$.

Two-stage procedure A key remaining challenge in estimating $\mathsf{W}_\alpha(\theta)$ is that we can only observe losses $\ell(\theta(X_i); Y_i)$ and never observe the conditional risk $\mu^\star(\cdot)$ (2.1.3). We propose a two-stage procedure (Algorithm 5), where we split the sample into two sets S_1 and S_2. On the first sample S_1, we fit an estimator $\widehat{\mu}_1(Z)$ of the conditional risk $\mu^\star(Z)$, using any model class $\mathcal{H}$ (class of mappings $\mathcal{Z} \to \mathbb{R}$), by solving an empirical approximation to the loss minimization problem

$$\underset{\mu \in \mathcal{H}}{\text{minimize}} \quad \mathbb{E}\left[(\ell(\theta(X); Y) - \mu(Z))^2\right]. \tag{2.2.3}$$

We denote by $\bar{\mu}$ a minimizer of (2.2.3); for a sufficiently rich model class $\mathcal{H}$, the minimizer is given by $\mu^\star(Z) = \mathbb{E}[\ell(\theta(X); Y) \mid Z]$. The loss minimization formulation (2.2.3) allows the use of any ML estimator, as well as standard tools for model selection (e.g. cross validation). In the second stage, on S_2 we construct a plug-in estimator for the dual form (2.2.2), under the estimated conditional risk $\widehat{\mu}_1(\cdot)$. In practice, we switch the roles of S_1 and S_2 and average the resulting estimates to leverage the entire sample.

To estimate the threshold subpopulation size $\alpha^\star(\theta)$, we simply take the plug-in estimator

$$\widehat{\alpha} := \inf\{\alpha : \widehat{\mathsf{W}}_{\alpha,k}(\widehat{\mu}_1) \le \bar{\ell}\}. \tag{2.2.4}$$

Since $\alpha \mapsto \widehat{\mathsf{W}}_{\alpha,k}(\widehat{\mu}_1)$ is decreasing, the threshold can be efficiently found by a simple bisection search.

2.3 Convergence guarantees

To *rigorously* verify the robustness of a model prior to deployment, we present convergence guarantees for our estimator (Algorithm 5). In Section 2.3.1, we first give finite-sample convergence at the rate $O_p(\sqrt{\mathfrak{Comp}_n(\mathcal{H})/n})$, where $\mathfrak{Comp}_n(\mathcal{H})$ is the localized Rademacher complexity [163] of the model class $\mathcal{H}$ for estimating the conditional risk $\mu^\star(Z)$. In some situations, it may be appropriate to define subpopulations (Z) over features of an image, or natural language descriptions. For such high-dimensional variables Z and complex model classes $\mathcal{H}$ such as deep networks, the complexity measure $\mathfrak{Comp}_n$ is often prohibitively conservative and renders the resulting concentration guarantee meaningless. In Section 2.3.2, we provide a finite-sample, data-dependent convergence result that depends only on the out-of-sample error for estimating $\mu^\star(\cdot)$. In particular, the out-of-sample error can grow smaller as $\mathcal{H}$ gets richer, and as a result of hyperparameter tuning and model selection, it is often very small for overparameterized models such as deep networks. This allows us to construct valid finite-sample upper confidence bounds for the worst-case subpopulation performance (2.1.2) even when Z is defined over high-dimensional features and $\mathcal{H}$ represent deep networks. Finally, in Section 2.3.3, we provide convergence guarantees for our estimator (2.2.4) for the certificate of robustness (2.2.1). By building on previous guarantees, we are again able to obtain both types of results.

We restrict attention to nonnegative and bounded losses.

Assumption C. *There is a B such that $\ell(\theta(X); Y) \in [0, B]$, and $\mu(Z) \in [0, B]$ a.s. for all $\mu \in \mathcal{H}$.*

Throughout, we do not stipulate well-specification, meaning that we allow the conditional risk $\mu^\star(\cdot) = \mathbb{E}[\ell(\theta(X); Y) \mid \cdot]$ not to be in the model class $\mathcal{H}$.

2.3.1 Concentration using the localized Rademacher complexity

To characterize the finite-sample convergence behavior of our estimator $\widehat{\mathsf{W}}_{\alpha,k}(\theta)$, we begin by decomposing the error into two terms relating to the two stages in Algorithm 5. Recalling the notation in Eq. (2.2.2) (so that $\mathsf{W}_\alpha(\mu^\star) = \mathsf{W}_\alpha(\theta)$), we have

$$\mathsf{W}_\alpha(\mu^\star) - \widehat{\mathsf{W}}_{\alpha,k}(\widehat{\mu}_1) = \underbrace{\mathsf{W}_\alpha(\mu^\star) - \mathsf{W}_\alpha(\widehat{\mu}_1)}_{(a):\ \text{first stage}} + \underbrace{\mathsf{W}_\alpha(\widehat{\mu}_1) - \widehat{\mathsf{W}}_{\alpha,k}(\widehat{\mu}_1)}_{(b):\ \text{second stage}}. \tag{2.3.1}$$

To bound term (b), we prove concentration guarantees for estimators of the dual (2.2.2) (see Proposition 5 in Appendix B.1.1). To bound term (a), we use a localized notion of the Rademacher complexity.

Formally, for $\xi_1, \ldots, \xi_n \in \Xi$ and i.i.d. random signs $\varepsilon_i \in \{-1, 1\}$ (independent of ξ_i), recall the standard notion of (empirical) Rademacher complexity of $\mathcal{G} \subseteq \{g : \Xi \to \mathbb{R}\}$

$$\mathfrak{R}_n(\mathcal{G}) := \mathbb{E}_\varepsilon \left[\sup_{g \in \mathcal{G}} \frac{1}{n} \sum_{i=1}^n \varepsilon_i g(\xi_i) \right].$$

We say that a function $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ is *sub-root* [163] if it is nonnegative, nondecreasing, and $r \mapsto \psi(r)/\sqrt{r}$ is nonincreasing for $r > 0$. Any (non-constant) sub-root function is continuous, and has a unique positive fixed point. Let $\psi_n : \mathbb{R}_+ \to \mathbb{R}_+$ be a sub-root upper bound on the localized Rademacher complexity $\psi_n(r) \geq \mathbb{E}\left[\mathfrak{R}_n\left\{g \in \mathcal{G} : \mathbb{E}[g^2] \leq r\right\}\right]$. (The localized Rademacher complexity itself is sub-root.) The fixed point of ψ_n characterizes generalization guarantees [163, 164].

Let $\bar{\mu}$ be the best model in the model class $\mathcal{H}$

$$\bar{\mu} := \operatorname*{argmin}_{\mu \in \mathcal{H}} \mathbb{E}[(\ell(\theta; X, Y) - \mu(Z))^2].$$

Let $\psi_{|S_1|}(r)$ be a subroot upper bound on the localized Rademacher complexity around $\bar{\mu}$

$$\psi_{|S_1|}(r) \geq 2\mathbb{E}\left[\mathfrak{R}_{|S_1|}\left\{\mu \in \mathcal{H} : \mathbb{E}[(\mu(Z) - \bar{\mu}(Z))^2] \leq rB^2/4\right\}\right]. \tag{2.3.2}$$

We define $r^\star_{|S_1|}$ as the fixed point of $\psi_{|S_1|}(r)$.

As we show shortly, we bound the estimation error of our procedure using the *square root* of the excess risk in the first-stage problem (2.2.3)

$$\mathbb{E}\left[(\ell(\theta; X, Y) - \widehat{\mu}_1(Z))^2 \mid S_1\right] - \mathbb{E}\left[(\ell(\theta; X, Y) - \bar{\mu}(Z))^2\right]$$

By using a refined analysis offered by localized Rademacher complexities, we are able to use a fast rate of convergence of $O_p(\mathfrak{Comp}_n(\mathcal{H})/n)$ on the preceding excess risk. In turn, this provides the following $O_p(\sqrt{\mathfrak{Comp}_n(\mathcal{H})/n})$ bound on the estimation error as we prove in Appendix B.1.2. In the bound, we have made explicit the approximation error term $\|\bar{\mu} - \mu\|_{L^2}$. As the model class

$\mathcal{H}$ grows richer, there is tension as the approximation error term will shrink, yet the localized Rademacher complexity of $\mathcal{H}$ will grow.

Theorem 4. *Let Assumption C hold. For some constant* $C > 0$*, with probability at least* $1 - 2\delta$*,*

$$\left|\mathsf{W}_\alpha(\theta) - \widehat{\mathsf{W}}_{\alpha,k}(\widehat{\mu}_1)\right| \le \frac{CB}{\alpha}\left(\sqrt{r^\star_{|S_1|}} + \sqrt{\frac{\log(1/\delta)}{|S_1|}} + \sqrt{\frac{\log(2/\delta)}{|S_2|}}\right) + \frac{1}{\alpha}\left\|\bar{\mu} - \mu^\star\right\|_{L_2}.$$

If we let S_1 be $(1-1/k)$-fraction of the data and S_2 be the remaining $1/k$-fraction for some integer k (e.g. $k = 5$), we have $|S_1| \asymp |S_2| \asymp n$. Thus, by controlling the fixed point $r^\star_{|S_1|}$ of the localized Rademacher complexity, we are able to provide convergence of our estimator (3). For example, when $\mathcal{H}$ is a bounded VC-class [165], it is known that its fixed point satisfy [163, Corollary 3.7]

$$r^\star_{|S_1|} \asymp \log(|S_1|/\mathsf{VC}(\mathcal{H})) \cdot \mathsf{VC}(\mathcal{H})/|S_1|,$$

where $\mathsf{VC}(\cdot)$ is the VC-dimension.

2.3.2 Data-dependent dimension-free concentration

In some applications, it may be natural to model Z as a high-dimensional variable. This may include large subsets of (X, Y), or defining Z using unstructured information such as images or natural languages. In these instances, we may wish to use deep networks as the model class $\mathcal{H}$ for estimating the conditional risk (2.1.3). We now provide an alternative concentration result that depends on the size of model class $\mathcal{H}$ only through the out-of-sample error in the first-stage problem (2.2.3). We denote for simplicity

$$\Delta_S(\mu) := \frac{1}{|S|}\sum_{i\in S}(\ell(\theta(X_i); Y_i) - \mu(Z_i))^2. \tag{2.3.3}$$

for any function $\mu : \mathcal{Z} \to \mathbb{R}$ on any data set S. We prove the following result in Appendix B.1.3.

Theorem 5. *Let Assumption C hold. For some constant $C > 0$, with probability at least $1 - 3\delta$,*

$$|\mathsf{W}_{\alpha}(\theta) - \widehat{\mathsf{W}}_{\alpha,k}(\widehat{\mu}_1)| \leq \frac{1}{\alpha}\left(\sqrt{\left(\Delta_{S_2}(\widehat{\mu}_1) - \Delta_{S_2}(\bar{\mu})\right)_+} + CB\left(\frac{\log(1/\delta)}{|S_2|}\right)^{1/4} + \left\|\bar{\mu} - \mu^\star\right\|_{L^2}\right).$$

Moreover, if the model class $\mathcal{H}$ is convex, then $\left\|\bar{\mu} - \mu^\star\right\|_{L^2}$ can be replaced with $\left\|\bar{\mu} - \mu^\star\right\|_{L^1}$.

Following convention in learning theory, we refer to our data-dependent concentration guarantee *dimension-free*. For overparameterized model classes $\mathcal{H}$ such as deep networks, the localized Rademacher complexity in Theorem 4 becomes prohibitively large [119, 120]. In contrast, the current result can still provide meaningful finite-sample bounds: model selection and hyperparameter tuning provides low out-of-sample performance in practice, and the difference $\Delta_{S_2}(\widehat{\mu}_1) - \Delta_{S_2}(\bar{\mu})$ can be often made very small. Concretely, it is possible to calculate an upper bound on this term as $\Delta_{S_2}(\bar{\mu})$ is lower bounded by $\min_{\mu \in \mathcal{H}} \Delta_{S_2}(\mu)$.

2.3.3 Certificate of robustness

Instead of estimating the worst-case subpopulation performance for a fixed subpopulation size α, it may be natural to posit a level of acceptable performance (upper bound $\bar{\ell}$ on the loss) and study $\alpha^\star(\theta)$, the smallest subpopulation size (2.2.1) over which the model $\theta(\cdot)$ can guarantee acceptable performance. Our plug-in estimator $\widehat{\alpha}$ given in Eq. (2.2.4) enjoys similar concentration guarantees as those given in Sections 2.3.1, 2.3.2. The following theorem—whose proof we give in Appendix B.1.4—states that the true $\alpha^\star(\theta)$ is either close to our estimator $\widehat{\alpha}$ or it is sufficiently small, certifying the robustness of the model against subpopulation shifts.

Theorem 6. *Let Assumption C hold, let $U(\delta) > 0$ be such that for any fixed $\alpha \in (0, 1]$, $|\widehat{\mathsf{W}}_{\alpha,k}(\widehat{h}) - \mathsf{W}_{\alpha}(\theta)| \leq U(\delta)/\alpha$ with probability at least $1 - \delta$. Then given any $\underline{\alpha} \in (0, 1]$, either $\alpha^\star(\theta) < \underline{\alpha}$, or*

$$\left|\frac{\alpha^\star(\theta)}{\widehat{\alpha}} - 1\right| \leq \frac{U(\delta)}{\widehat{\mathbb{E}}\left[\widehat{h}(Z) - \widehat{P}^{-1}_{1-\underline{\alpha}\wedge\widehat{\alpha}}\left(\widehat{h}(Z)\right)\right]_+}$$

with probability at least $1 - \delta$, where $\widehat{\mathbb{E}}$ and $\widehat{P}^{-1}_{1-\alpha}$ denote the expectation and the $(1 - \alpha)$-quantile under the empirical probability measure induced by S_2.

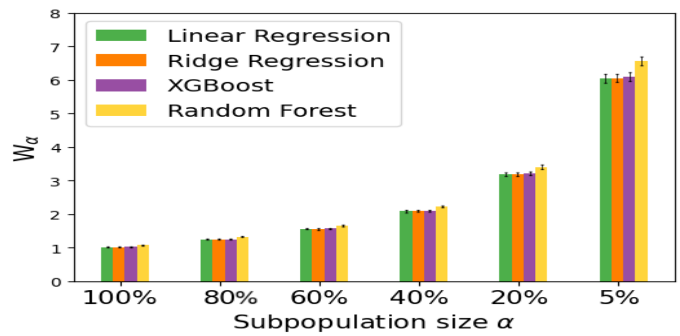


Figure 2.2: Worst-case subpopulation performance $\mathsf{W}_\alpha(\theta)$, where $\mathsf{W}_{1.0}(\theta) = \mathbb{E}[\ell(\theta(X); Y)]$.

Our approach simultaneously provides localized Rademacher complexity bounds and dimension-free guarantees. Our bound becomes large as $\underline{\alpha} \to 0$ and we conjecture this to be a fundamental difficulty as the worst-case subpopulation performance (2.1.2) focuses on α-faction of the data.

2.4 Experiments

On two real datasets, we demonstrate that our diagnostic allows certifying model performance across subpopulations. We first study a drug dosage prediction problem, where our procedure ascertains the robustness of a linear regression model over substantially more expressive model classes. Then, we turn to a large-scale computer vision application based on satellite images [166] where natural distribution shifts were recently studied [101]. In both settings, we illustrate how our worst-case subpopulation approach raises awareness on brittle models without knowledge of out-of-distribution samples. Finally, to verify asymptotic convergence of our proposed two-stage estimator, we present a simultion experiment on a classification task in Appendix B.3. For all experiments, we use gradient boosted decision trees (package XGBoost [167]) to estimate the conditional risk $\mu^\star(Z) = \mathbb{E}[\ell(\theta(X); Y) \mid Z]$.

2.4.1 Warfarin

Warfarin is one of the most widely used anticoagulant, often prescribed to prevent strokes [159]. Its optimal dosage varies substantially across genetics, demographics, and existing conditions (up to ten times). We study a Pharmacogenetics and Pharmacogenomics Knowledge Base dataset constructed from optimal dosages found through trial and error by clinicians. The dataset comprises of 4,788 patients (after excluding missing data) alongside features representing demographics, genetic markers, medication history, pre-existing conditions, and reason for treatment. Consortium [159] found that a linear model outperforms a number of more complicated modeling approaches (e.g. kernel methods, neural networks, splines, boosting) for predicting the optimal dosage.

Such average-case performance needs to translate uniformly to different subpopulations; we need to ensure automated medical models perform well on underrepresented groups [168, 169, 170, 171]. We wish to evaluate and compare the worst-case subpopulation performance of different models over $Z = X$, the entire feature vector. Following Consortium [159], we take the root-dosage as our outcome Y, and consider the squared loss $\ell(\theta(X); Y) = (Y - \theta(X))^2$. In Figure 2.2, we observe that the linear model closely matches the performance of more expressive models even over small subpopulations. Moreover, the trend holds over a range of different subpopulation sizes (up to $\alpha = 5\%$). Our finding instills confidence in the linear regression model: in addition to being simple and interpretable, our diagnostic certifies its advantageous performance even on tail subpopulations. However, our diagnostic raises some concerns about poor subpopulation performance: on $\alpha = 5\%$ of the training population, all models suffer prediction error six times worse than the average-case performance.

2.4.2 Functional Map of the World (fMoW)

Satellite images can impact economic and environmental policies globally by allowing large-scale measurements on poverty [172], population changes, deforestation, and economic growth [173]. An automated approach allows analyzing data from remote regions at a relatively low cost and provides continuous monitoring of land usage. Towards this goal, it is critical that the models perform reliably across time and space. We study this problem on the Functional Map of the World (fMoW) dataset [166], where the goal is to predict building / land use categories (62 classes) based on satellite images. Across different models, we observe that performance remains similar either temporally or spatially when each dimension is considered *separately*, but there is substantial variability across intersections of region and year. For a standard *DenseNet ERM* model [174, 101] that achieves near-state-of-the-art performance, we present these trends in Figure 2.3(a). In Figure 2.3(b), we observe substantial variability in classwise error rates; there is a varying level of difficulty across different classes. (We observed similar patterns for other models.)

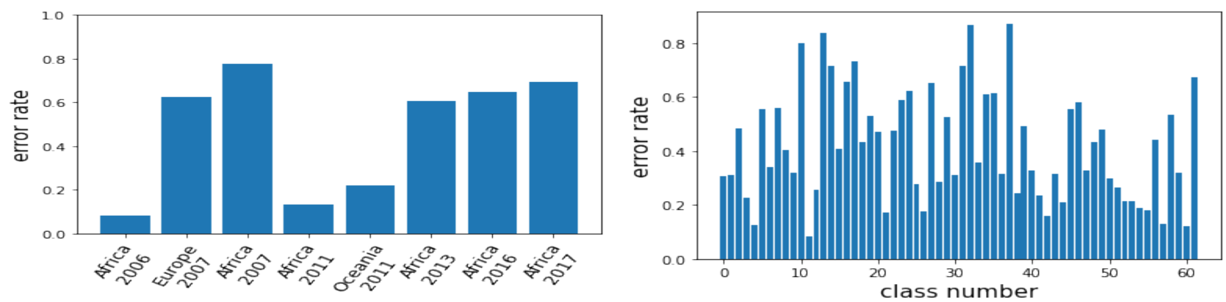


Figure 2.3. For *DenseNet ERM*, spatiotemproal intersectionality (left) and performance by class (right)

We take the perspective of an analyst evaluating prediction models for land usage, based on data collected during 2002-2013. The fMoW dataset provides fertile grounds for demonstrating our method as it includes natural distribution shifts [101], both spatial and temporal. In particular, we demonstrate model robustness on out-of-distribution samples collected in 2016-2018. On validation data collected during 2002-2013, we first evaluate model performance on subpopulations defined across metadata on a satellite image, which consists of (subsets of) {*longitude, latitude, cloud cover, region, year*} and the label Y. Then, we observe how our procedure selects models that perform well "in the future" without requiring out-of-distribution data.

We examine a range of different models trained on the fMoW-WILDS training set (collected in 2002-2013, n =76,863) which fall into two broad categories. First, we consider models pre-trained on ImageNet and finetuned on the fMoW training set. These include *DenseNet* models trained using ERM and the recently proposed invariant risk minimization (IRM) framework [175]. We also study the Dual Path Network-68 (*DPN-68*) model with connection paths that enable feature reusage and feature exploration proposed by Chen *et al.* [176]. We use *DPN-68* trained on fMoW using ERM as reported in [177]. These models all achieve in distribution (ID) accuracy of $\sim$ 60% on a heldout validation set ("ID val", collected in 2002-2013, n =11,483).

Second, we consider models derived from the recently proposed CLIP model [121], which was trained on large and heterogeneous data sources comprising of 40M image-text pairs using natural language supervision and contrastive losses. The pre-training data for CLIP is 400 times bigger than ImageNet, and Radford *et al.* [121] have observed that zero-shot applications of CLIP exhibits substantial *relative robustness gains* over other state-of-the-art methods on natural distribution shifts of ImageNet.

However, on the fMoW in-distribution (2002–2013) validation set, zero-shot CLIP only achieves 19.3% accuracy compared to the 60% accuracy of ImageNet pre-trained models. We thus finetune it using satellite images in the fMoW training data. While finetuning substantially improves ID accuracy on fMoW to 70.2%, the relative robustness gains of the zero-shot CLIP model severely degrade. To address this problem, Wortsman *et al.* [178] proposed a weight-space ensembling method (*CLIP WiSE-FT*) where they average the network weights of the zero-shot CLIP model and its finetuned counterpart. These ensembled networks have been observed to exhibit large Pareto improvements in both in-distribution and out-of-distribution accuracy, including on the fMoW dataset.

Motivated by the observed robustness gains, we average the network weights θ_0 of the *CLIP Zeroshot* model and that of *CLIP fine-tuned* θ_1 to generate a new network $(1-\lambda)\theta_0+\lambda\theta_1$, where $\lambda \in [0, 1]$ controls how much weight is given to the task-specific, fine-tuned model (domain expertise). We select $\lambda = 0.4$ so that the ensembled model (*CLIP WiSE-FT*) achieves similar performance as ImageNet pre-trained counterparts on the in-distribution validation data. To further make models comparable with respect to the cross entropy loss, we calibrate the *CLIP WiSE-FT* model by tuning the temperature parameter so that its average loss on the in-distribution validation set matches the worst average loss of ImageNet pre-trained models (*DenseNet ERM*). See Appendix B.2 for detailed experimental settings and training specifications.

We compute estimators of W_α (Algorithm 5) on the in-distribution validation data (ID val) using the standard cross entropy loss. In Figure 2.4, we summarize the estimated worst-case subpopulation performances defined over the entire *metadata*, across different subpopulation sizes α. First, we note that all models have comparable in-distribution accuracy of $\sim 60\%$ and *DenseNet*

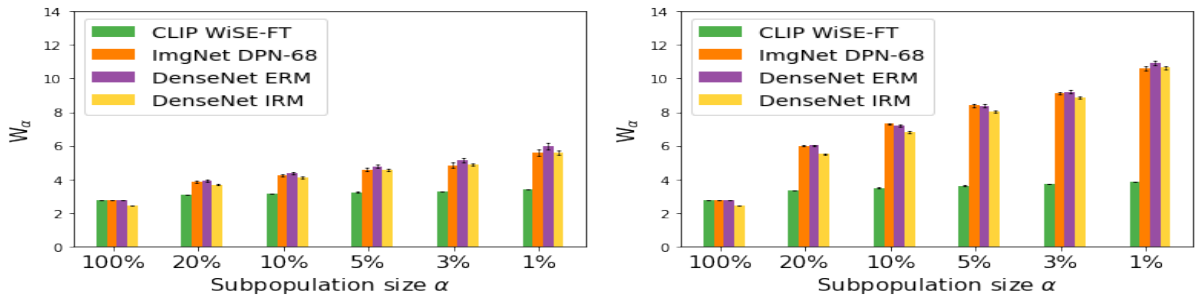


Figure 2.4. Left: Z = (all metadata); Right: Z = (all metadata, Y). Results are averaged over 50 random seeds with error bars corresponding to a 95% confidence interval over the random runs.

	ID, 2002–2013		OOD, 2016–2018			
Models	Accuracy	Loss	Accuracy	Loss	Africa Accuracy	Africa Loss
CLIP WiSE-FT	0.61	2.78	0.56	2.84	0.38	3.08
DenseNet ERM	0.61	2.78	0.53	3.50	0.33	5.41
DenseNet IRM	0.59	2.44	0.51	2.94	0.31	4.46
DPN-68	0.61	2.75	0.53	3.55	0.31	5.61

Table 2.1. Model performance on ID val and OOD test sets. All models suffer a performance drop on the OOD test set in both accuracy and average loss. The performance degradation is particularly significant on images from Africa. On the OOD data, *CLIP WiSE-FT* outperforms other models both in average accuracy/loss and worst-region accuracy/loss.

IRM has the best average-case cross entropy loss. However, the worst-case subpopulation performance of the ImageNet pre-trained models is substantially worse compared to that of *CLIP WiSE-FT*. This gap grows larger as the subpopulation size α becomes increasingly small. Evaluations on worst-case subpopulations suggest that *CLIP WiSE-FT* exhibits robustness against subpopulation shifts; in contrast, average-case evaluations will select *DenseNet IRM*.

We observe a drastic performance deterioration on tail subpopulations. The inclusion of label information in Z significantly deteriorates worst-case performance, raising concerns about the distributional robustness of all models including changes in the label distribution. In Table 2.1, we present model performances on the out-of-distribution ("future") data collected during 2016–2018. All models suffer a significant performance drop under temporal distribution shift, particularly on images collected in Africa where predictive accuracy drops by up to 20 percentage points. *CLIP WiSE-FT* exhibits the most robustness under spatiotemporal shift than any other model, as presaged by evaluations of worst-case subpopulation performance in Figure 2.4.

A key advantage of our method is the flexibility in the choice of Z; the modeler can define granular or coarse subpopulations based on this choice. As defining subpopulations over all metadata can be conservative, we present additional results under Z =(*region*, *year*) and Z =(*region*, *year*, *label* Y) in Appendix B.2. Instead of incorporating labels as a category, it may be more informative to use the *semantic meaning of each class label*. We generate natural language description of the labels by concatenating each class label with engineered prompts, and pass it to the CLIP text encoder [121] to generate a feature representation for the label. In Appendix B.2.3, we present evaluation results where we take the feature vector in place of the label Y when defining Z.

2.5 Discussion

To ensure models perform reliably under operation, we need to *rigorously* certify their performance under distribution shift prior to deployment. We study the *worst-case subpopulation performance* of a model, a natural notion of model robustness that is easy to communicate with users, regulators, and business leaders. Our approach allows flexible modeling of subpopulations over an arbitrary variable Z and automatically accounts for complex intersectionality. We develop scalable estimation procedures for the worst-case subpopulation performance (2.1.2) and the certificate of robustness (2.2.1) of a model. Our convergence guarantees apply even when we use high-dimensional inputs (e.g. natural language) to define Z. Our diagnostic may further inform data collection and model improvement by suggesting data collection efforts and model fixes on regions of $\mathcal{Z}$ with high conditional risk (2.1.3).

The worst-case performance (2.1.2) over mixture components as subpopulations (2.1.1) provides a strong guarantee over arbitrary subpopulations, but it may be overly conservative in cases when there is a natural geometry in $Z \in \mathcal{Z}$. Incorporating such problem-specific structures in defining a tailored notion of subpopulation is a promising research direction towards operationalizing the concepts put forth in this work. As an example, Srivastava *et al.* [179] recently studied similar notions of worst-case performance defined over human annotations.

We focus on the narrow question of evaluating model robustness under distribution shift; our evaluation perspective is thus inherently limited. Data collection systems inherit socioeconomic inequities, and reinforce existing political power structures. This affects *all* aspects of the ML development pipeline, and our diagnostic is no panacea. A notable limitation of our approach is that we do not explicitly consider the power differential that often exists between those who deploy the prediction system and those for whom it gets used on. Systems must be deployed with considered analysis of its adverse impacts, and we advocate for a holistic approach towards addressing its varied implications.

References

[1] A. Abdulkadiroglu and T. Sönmez, “School choice: A mechanism design approach,” *American economic review*, vol. 93, no. 3, pp. 729–747, 2003.

[2] A. Abdulkadiroglu, Y.-K. Che, P. A. Pathak, A. E. Roth, and O. Tercieux, “Minimizing justified envy in school choice: The design of new orleans’ oneapp,” National Bureau of Economic Research, Tech. Rep., 2017.

[3] P. A. Pathak, A. Rees-Jones, and T. Sönmez, “Immigration lottery design: Engineered and coincidental consequences of h-1b reforms,” National Bureau of Economic Research, Tech. Rep., 2020.

[4] E. H. Kaplan, “Managing the demand for public housing,” Ph.D. dissertation, Massachusetts Institute of Technology, 1984.

[5] N. Arnosti and P. Shi, “Design of lotteries and wait-lists for affordable housing allocation,” *Management Science*, vol. 66, no. 6, pp. 2291–2307, 2020.

[6] Y. Azar, J. Naor, and R. Rom, “The competitiveness of on-line assignments,” *Journal of Algorithms*, vol. 18, no. 2, pp. 221–237, 1995.

[7] Y. Azar, “On-line load balancing,” *Online algorithms*, pp. 178–195, 1998.

[8] D. Bertsimas, V. F. Farias, and N. Trichakis, “The price of fairness,” *Operations research*, vol. 59, no. 1, pp. 17–31, 2011.

[9] D. Bertsimas, V. F. Farias, and N. Trichakis, “On the efficiency-fairness trade-off,” *Management Science*, vol. 58, no. 12, pp. 2234–2250, 2012.

[10] F. Ahmed, J. P. Dickerson, and M. Fuge, “Diverse weighted bipartite b-matching,” *arXiv preprint arXiv:1702.07134*, 2017.

[11] K. Donahue and J. Kleinberg, “Fairness and utilization in allocating resources with uncertain demand,” in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 2020, pp. 658–668.

[12] D. Zeng and A. Psomas, “Fairness-efficiency tradeoffs in dynamic fair division,” in *Proceedings of the 21st ACM Conference on Economics and Computation*, 2020, pp. 911–912.

[13] H. R. Varian, “Equity, envy, and efficiency,” *Journal of Economic Theory*, vol. 9, no. 1, pp. 63–91, 1974.

[14] F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia, *Handbook of computational social choice*. Cambridge University Press, 2016.

[15] S. Banerjee, V. Gkatzelis, S. Hossain, B. Jin, E. Micha, and N. Shah, “Proportionally fair online allocation of public goods with predictions,” *arXiv preprint arXiv:2209.15305*, 2022.

[16] J. Rawls, *Justice as fairness: A restatement*. Harvard University Press, 2001.

[17] J. Rawls, “A theory of justice,” in *A theory of justice*, Harvard university press, 2020.

[18] N. Bansal and M. Sviridenko, “The santa claus problem,” in *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, 2006, pp. 31–40.

[19] M. Ben-Ari, “How to solve the santa claus problem,” *Concurrency: Practice and Experience*, vol. 10, no. 6, pp. 485–496, 1998.

[20] D. Chakrabarty, J. Chuzhoy, and S. Khanna, “On allocating goods to maximize fairness,” in *2009 50th Annual IEEE Symposium on Foundations of Computer Science*, IEEE, 2009, pp. 107–116.

[21] J. F. Nash Jr, “The bargaining problem,” *Econometrica: Journal of the econometric society*, pp. 155–162, 1950.

[22] R. D. Luce and H. Raiffa, *Games and decisions: Introduction and critical survey*. Courier Corporation, 1989.

[23] F Hahn, K. Arrow, and M. Intriligator, *Handbook of mathematical economics*, 1982.

[24] R. Freeman, N. Shah, and R. Vaish, “Best of both worlds: Ex-ante and ex-post fairness in resource allocation,” in *Proceedings of the 21st ACM Conference on Economics and Computation*, ser. EC ’20, Virtual Event, Hungary: Association for Computing Machinery, 2020, 21–22, ISBN: 9781450379755.

[25] I. Caragiannis, D. Kurokawa, H. Moulin, A. D. Procaccia, N. Shah, and J. Wang, “The unreasonable fairness of maximum nash welfare,” *ACM Trans. Econ. Comput.*, vol. 7, no. 3, 2019.

[26] V. Conitzer, R. Freeman, and N. Shah, “Fair public decision making,” in *Proceedings of the 2017 ACM Conference on Economics and Computation*, ser. EC ’17, Cambridge, Massachusetts, USA: Association for Computing Machinery, 2017, 629–646, ISBN: 9781450345279.

[27] X. Wu, B. Li, and J. Gan, “Budget-feasible maximum nash social welfare is almost envy-free,” International Joint Conferences on Artificial Intelligence Organization, 2021.

[28] V. Nanda, P. Xu, K. A. Sankararaman, J. Dickerson, and A. Srinivasan, “Balancing the tradeoff between profit and fairness in rideshare platforms during high-demand hours,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, 2020, pp. 2210–2217.

[29] J. P. Dickerson, K. A. Sankararaman, A. Srinivasan, and P. Xu, “Balancing relevance and diversity in online bipartite matching via submodularity,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, 2019, pp. 1877–1884.

[30] W. C. Cheung, G. Lyu, C.-P. Teo, and H. Wang, “Online planning with offline simulation,” *Available at SSRN 3709882*, 2020.

[31] T. Q. Dinh, B. Liang, T. Q. Quek, and H. Shin, “Online resource procurement and allocation in a hybrid edge-cloud computing system,” *IEEE transactions on wireless communications*, vol. 19, no. 3, pp. 2137–2149, 2020.

[32] J. Zhang, N. Xie, X. Zhang, and W. Li, “An online auction mechanism for cloud computing resource allocation and pricing based on user evaluation and cost,” *Future Generation Computer Systems*, vol. 89, pp. 286–299, 2018.

[33] S. Cayci, S. Gupta, and A. Eryilmaz, “Group-fair online allocation in continuous time,” *Advances in Neural Information Processing Systems*, vol. 33, pp. 13 750–13 761, 2020.

[34] N. R. Devanur, K. Jain, B. Sivan, and C. A. Wilkens, “Near optimal online algorithms and fast approximation algorithms for resource allocation problems,” *Journal of the ACM (JACM)*, vol. 66, no. 1, pp. 1–41, 2019.

[35] A. Grigoryan, “Effective, fair and equitable pandemic rationing,” *Available at SSRN 3646539*, 2021.

[36] F. Alkaabneh, A. Diabat, and H. O. Gao, “A unified framework for efficient, effective, and fair resource allocation by food banks using an approximate dynamic programming approach,” *Omega*, vol. 100, p. 102 300, 2021.

[37] N. Agarwal, I. Ashlagi, M. A. Rees, P. J. Somaini, D. C. Waldinger, *et al.*, *An empirical framework for sequential assignment: The allocation of deceased donor kidneys*. National Bureau of Economic Research, 2019.

[38] M. D. Aleksandrov, H. Aziz, S. Gaspers, and T. Walsh, “Online fair division: Analysing a food bank problem,” in *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015.

[39] R. W. Lien, S. M. Iravani, and K. R. Smilowitz, "Sequential resource allocation for nonprofit operations," *Operations Research*, vol. 62, no. 2, pp. 301–317, 2014.

[40] H. Choi, C. F. Mela, S. R. Balseiro, and A. Leary, "Online display advertising markets: A literature review and future directions," *Information Systems Research*, vol. 31, no. 2, pp. 556–575, 2020.

[41] A. Watts, "Fairness and efficiency in online advertising mechanisms," *Games*, vol. 12, no. 2, p. 36, 2021.

[42] S. Gollapudi and D. Panigrahi, "Fair allocation in online markets," in *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*, 2014, pp. 1179–1188.

[43] X. Li, Y. Rong, R. Zhang, and H. Zheng, "Online advertisement allocation under customer choices and algorithmic fairness," *Available at SSRN 3538755*, 2021.

[44] M. Bateni, Y. Chen, D. F. Ciocan, and V. Mirrokni, "Fair resource allocation in a volatile marketplace," *Operations Research*, vol. 70, no. 1, pp. 288–308, 2022.

[45] N. Arnosti, M. Beck, and P. Milgrom, "Adverse selection and auction design for internet display advertising," *American Economic Review*, vol. 106, no. 10, pp. 2852–66, 2016.

[46] S. Jasin and S. Kumar, "A re-solving heuristic with bounded revenue loss for network revenue management with customer choice," *Mathematics of Operations Research*, vol. 37, no. 2, pp. 313–345, 2012.

[47] S. Balseiro, O. Besbes, and D. Pizarro, "Survey of dynamic resource constrained reward collection problems: Unified model and analysis," *Available at SSRN 3963265*, 2021.

[48] P. Bumpensanti and H. Wang, "A re-solving heuristic with uniformly bounded loss for network revenue management," *Management Science*, vol. 66, no. 7, pp. 2993–3009, 2020.

[49] M. I. Reiman and Q. Wang, "An asymptotically optimal policy for a quantity-based network revenue management problem," *Mathematics of Operations Research*, vol. 33, no. 2, pp. 257–282, 2008.

[50] S. Agrawal and N. R. Devanur, "Fast algorithms for online stochastic convex programming," in *Proceedings of the twenty-sixth annual ACM-SIAM symposium on Discrete algorithms*, SIAM, 2014, pp. 1405–1424.

[51] S. Arora, E. Hazan, and S. Kale, "The multiplicative weights update method: A meta-algorithm and applications," *Theory of computing*, vol. 8, no. 1, pp. 121–164, 2012.

[52] S. Balseiro, H. Lu, and V. Mirrokni, "Regularized online allocation problems: Fairness and beyond," in *International Conference on Machine Learning*, PMLR, 2021, pp. 630–639.

[53] Y. Kawase and H. Sumita, "Online max-min fair allocation," *arXiv preprint arXiv:2111.07235*, 2021.

[54] D. Freund and S. Banerjee, "Good prophets know when the end is near," *Available at SSRN 3479189*, 2019.

[55] S. Banerjee, V. Gkatzelis, A. Gorokh, and B. Jin, "Online nash social welfare maximization with predictions," in *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2022, pp. 1–19. eprint: `https://epubs.siam.org/doi/pdf/10.1137/1.9781611977073.1`.

[56] R. Freeman, S. M. Zahedi, and V. Conitzer, "Fair and efficient social choice in dynamic settings," in *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*, 2017, pp. 4580–4587.

[57] S. R. Sinclair, G. Jain, S. Banerjee, and C. L. Yu, "Sequential fair allocation of limited resources under stochastic demands," *arXiv preprint arXiv:2011.14382*, 2020.

[58] B. L. Deuermeyer, D. K. Friesen, and M. A. Langston, "Scheduling to maximize the minimum processor finish time in a multiprocessor system," *SIAM Journal on Algebraic Discrete Methods*, vol. 3, no. 2, pp. 190–196, 1982.

[59] Y. Azar and L. Epstein, "On-line machine covering," *Journal of Scheduling*, vol. 1, no. 2, pp. 67–77, 1998.

[60] W. Gálvez, J. A. Soto, and J. Verschae, "Symmetry exploitation for online machine covering with bounded migration," *ACM Transactions on Algorithms (TALG)*, vol. 16, no. 4, pp. 1–22, 2020.

[61] G. J. Woeginger, "A polynomial-time approximation scheme for maximizing the minimum machine completion time," *Operations Research Letters*, vol. 20, no. 4, pp. 149–154, 1997.

[62] Y. Li *et al.*, "Max-min fair allocation for resources with hybrid divisibilities," *Expert Systems with Applications*, vol. 124, pp. 325–340, 2019.

[63] A. Asadpour and A. Saberi, "An approximation algorithm for max-min fair allocation of indivisible goods," *SIAM Journal on Computing*, vol. 39, no. 7, pp. 2970–2989, 2010.

[64] Y. Kawase and H. Sumita, "On the max-min fair stochastic allocation of indivisible goods," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, 2020, pp. 2070–2078.

[65] I. Bezáková and V. Dani, “Allocating indivisible goods,” *ACM SIGecom Exchanges*, vol. 5, no. 3, pp. 11–18, 2005.

[66] D. Golovin, *Max-min fair allocation of indivisible goods*. School of Computer Science, Carnegie Mellon University, 2005.

[67] U. Feige, “On allocations that maximize fairness.,” in *SODA*, Citeseer, vol. 8, 2008, pp. 287–293.

[68] B. Haeupler, B. Saha, and A. Srinivasan, “New constructive aspects of the lovász local lemma,” *Journal of the ACM (JACM)*, vol. 58, no. 6, pp. 1–28, 2011.

[69] V. Manshadi, R. Niazadeh, and S. Rodilitz, “Fair dynamic rationing,” in *Proceedings of the 22nd ACM Conference on Economics and Computation*, 2021, pp. 694–695.

[70] W. Ma, P. Xu, and Y. Xu, “Group-level fairness maximization in online bipartite matching,” *arXiv preprint arXiv:2011.13908*, 2020.

[71] G. Benade, A. M. Kazachkov, A. D. Procaccia, and C.-A. Psomas, “How to make envy vanish over time,” in *Proceedings of the 2018 ACM Conference on Economics and Computation*, 2018, pp. 593–610.

[72] J. He, A. D. Procaccia, C. Psomas, and D. Zeng, “Achieving a fairer future by changing the past,” *IJCAI’19*, 2019.

[73] N. R. Devanur and K. Jain, “Online matching with concave returns,” in *Proceedings of the Forty-Fourth Annual ACM Symposium on Theory of Computing*, ser. STOC ’12, New York, New York, USA: Association for Computing Machinery, 2012, 137–144, ISBN: 9781450312455.

[74] M. Aleksandrov and T. Walsh, “Online fair division: A survey,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, 2020, pp. 13 557–13 562.

[75] G. Gallego and G. Van Ryzin, “A multiproduct dynamic pricing problem and its applications to network yield management,” *Operations research*, vol. 45, no. 1, pp. 24–41, 1997.

[76] K. Talluri and G. Van Ryzin, “An analysis of bid-price controls for network revenue management,” *Management science*, vol. 44, no. 11-part-1, pp. 1577–1593, 1998.

[77] A. Vera and S. Banerjee, “The bayesian prophet: A low-regret framework for online decision making,” *Management Science*, vol. 67, no. 3, pp. 1368–1391, 2021.

[78] A. Vera, S. Banerjee, and I. Gurvich, “Online allocation and pricing: Constant regret via bellman inequalities,” *Operations Research*, vol. 69, no. 3, pp. 821–840, 2021.

[79] A. Arlotto and I. Gurvich, “Uniformly bounded regret in the multisecretary problem,” *Stochastic Systems*, vol. 9, no. 3, pp. 231–260, 2019.

[80] R. L. Bray, “Logarithmic regret in multisecretary and online linear programming problems with continuous valuations,” *arXiv e-prints*, arXiv–1912, 2019.

[81] H. Moulin, *Fair division and collective welfare*. MIT press, 2004.

[82] G. H. Hardy, J. E. Littlewood, G. Pólya, G. Pólya, *et al.*, *Inequalities*. Cambridge university press, 1952.

[83] D. Bertsekas, *Dynamic programming and optimal control: Volume I*. Athena scientific, 2012, vol. 1, p. 316.

[84] D. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1999.

[85] R. T. Rockafellar, “Lipschitzian properties of multifunctions,” *Nonlinear Analysis: Theory, Methods & Applications*, vol. 9, no. 8, pp. 867–885, 1985.

[86] A. Schrijver, *Theory of linear and integer programming*. John Wiley & Sons, 1998.

[87] W. Cook, A. M. Gerards, A. Schrijver, and É. Tardos, “Sensitivity theorems in integer linear programming,” *Mathematical Programming*, vol. 34, pp. 251–264, 1986.

[88] O. L. Mangasarian and T.-H. Shiau, “Lipschitz continuity of solutions of linear inequalities, programs and complementarity problems,” *SIAM Journal on Control and Optimization*, vol. 25, no. 3, pp. 583–595, 1987.

[89] A. J. Hoffman, “On approximate solutions of systems of linear inequalities,” in *Selected Papers Of Alan J Hoffman: With Commentary*, World Scientific, 2003, pp. 174–176.

[90] H. Gfrerer and J. V. Outrata, “On lipschitzian properties of implicit multifunctions,” *SIAM Journal on Optimization*, vol. 26, no. 4, pp. 2160–2189, 2016.

[91] J.-P. Aubin, “Lipschitz behavior of solutions to convex minimization problems,” *Mathematics of Operations Research*, vol. 9, no. 1, pp. 87–111, 1984.

[92] S. M. Robinson, *Generalized equations and their solutions, Part I: Basic theory*. Springer, 1979.

[93] D. Klatte and B. Kummer, “On calmness of the argmin mapping in parametric optimization problems,” *Journal of Optimization Theory and Applications*, vol. 165, pp. 708–719, 2015.

[94] X.-B. Li and S. Li, “Hölder continuity of perturbed solution set for convex optimization problems,” *Applied Mathematics and Computation*, vol. 232, pp. 908–918, 2014.

[95] D. Klatte and B. Kummer, "On hölder calmness of minimizing sets," *Optimization*, vol. 71, no. 4, pp. 1055–1072, 2022.

[96] R. T. Rockafellar and R. J.-B. Wets, *Variational analysis*. Springer Science & Business Media, 2009, vol. 317.

[97] J. Blitzer, R. McDonald, and F. Pereira, "Domain adaptation with structural correspondence learning," in *Proceedings of the 2006 conference on empirical methods in natural language processing*, Association for Computational Linguistics, 2006, pp. 120–128.

[98] H. Daume III and D. Marcu, "Domain adaptation for statistical classifiers," *Journal of artificial Intelligence research*, vol. 26, pp. 101–126, 2006.

[99] K. Saenko, B. Kulis, M. Fritz, and T. Darrell, "Adapting visual category models to new domains," in *Proceedings of the European Conference on Computer Vision*, Springer, 2010, pp. 213–226.

[100] A. Torralba and A. A. Efros, "Unbiased look at dataset bias," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2011, pp. 1521–1528.

[101] P. W. Koh *et al.*, "Wilds: A benchmark of in-the-wild distribution shifts," *arXiv:2012.07421 [cs.LG]*, 2020.

[102] A. Koenecke *et al.*, "Racial disparities in automated speech recognition," *Proceedings of the National Academy of Sciences*, vol. 117, no. 14, pp. 7684–7689, 2020.

[103] E. Amorim, M. Cançado, and A. Veloso, "Automated essay scoring in the presence of biased ratings," in *Association for Computational Linguistics (ACL)*, 2018, pp. 229–237.

[104] S. Beery, E. Cole, and A. Gjoka, "The iwildcam 2020 competition dataset," *arXiv:2004.10340 [cs.CV]*, 2020.

[105] R. Taori, A. Dave, V. Shankar, N. Carlini, B. Recht, and L. Schmidt, "Measuring robustness to natural distribution shifts in image classification," in *Advances in Neural Information Processing Systems 20*, 2020.

[106] V. Shankar, A. Dave, R. Roelofs, D. Ramanan, B. Recht, and L. Schmidt, "Do image classifiers generalize across time?" *arXiv:1906.02168 [cs.LG]*, 2019.

[107] J. Miller, K. Krauth, B. Recht, and L. Schmidt, "The effect of natural distribution shift on question answering models," in *International Conference on Machine Learning*, PMLR, 2020, pp. 6905–6916.

[108] B. Recht, R. Roelofs, L. Schmidt, and V. Shankar, "Do ImageNet classifiers generalize to ImageNet?" In *Proceedings of the 36th International Conference on Machine Learning*, 2019.

[109] J. Buolamwini and T. Gebru, "Gender shades: Intersectional accuracy disparities in commercial gender classification," in *Conference on Fairness, Accountability and Transparency*, 2018, pp. 77–91.

[110] K. Goel *et al.*, "Robustness gym: Unifying the nlp evaluation landscape," *arXiv:2101.04840 [cs.CL]*, 2021.

[111] E. Denton, A. Hanna, R. Amironesei, A. Smart, H. Nicole, and M. K. Scheuerman, "Bringing the people back in: Contesting benchmark machine learning datasets," *arXiv:2007.07399 [cs.CY]*, 2020.

[112] M. S. Chen, P. N. Lara, J. H. Dang, D. A. Paterniti, and K. Kelly, "Twenty years post-NIH revitalization act: Enhancing minority participation in clinical trials (EMPaCT): Laying the groundwork for improving minority clinical trial accrual: Renewing the case for enhancing minority participation in cancer clinical trials," *Cancer*, vol. 120, pp. 1091–1096, 2014.

[113] D. Amodei *et al.*, "Deep speech 2: End-to-end speech recognition in English and Mandarin," in *Proceedings of the 33rd International Conference on Machine Learning*, 2016, pp. 173–182.

[114] P. J. Grother, G. W. Quinn, and P. J. Phillips, "Report on the evaluation of 2d still-image face recognition algorithms," *NIST Interagency/Internal Reports (NISTIR)*, vol. 7709, 2010.

[115] D. Hovy and A. Søgaard, "Tagging performance correlates with author age," in *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics (Short Papers)*, vol. 2, 2015, pp. 483–488.

[116] S. L. Blodgett, L. Green, and B. O'Connor, "Demographic dialectal variation in social media: A case study of African-American English," in *Proceedings of Empirical Methods for Natural Language Processing*, 2016, pp. 1119–1130.

[117] P. Sapiezynski, V. Kassarnig, and C. Wilson, "Academic performance prediction in a gender-imbalanced environment," in *Proceedings of the Eleventh ACM Conference on Recommender Systems*, vol. 1, 2017, pp. 48–51.

[118] R. Tatman, "Gender and dialect bias in YouTube's automatic captions," in *First Workshop on Ethics in Natural Langauge Processing*, vol. 1, 2017, pp. 53–59.

[119] P. L. Bartlett, D. J. Foster, and M. J. Telgarsky, "Spectrally-normalized margin bounds for neural networks," in *Advances in Neural Information Processing Systems*, 2017, pp. 6241–6250.

[120] C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals, "Understanding deep learning requires rethinking generalization," in *Proceedings of the Fifth International Conference on Learning Representations*, 2017.

[121] A. Radford *et al.*, "Learning transferable visual models from natural language supervision," in *Proceedings of the 38th International Conference on Machine Learning*, 2021.

[122] E. Delage and Y. Ye, "Distributionally robust optimization under moment uncertainty with application to data-driven problems," *Operations Research*, vol. 58, no. 3, pp. 595–612, 2010.

[123] J. Goh and M. Sim, "Distributionally robust optimization and its tractable approximations," *Operations research*, vol. 58, no. 4-part-1, pp. 902–917, 2010.

[124] Anonymous, "Distributionally robust neural networks," in *Submitted to International Conference on Learning Representations*, under review, 2020.

[125] A. Ben-Tal, D. den Hertog, A. D. Waegenaere, B. Melenberg, and G. Rennen, "Robust solutions of optimization problems affected by uncertain probabilities," *Management Science*, vol. 59, no. 2, pp. 341–357, 2013.

[126] D. Bertsimas, V. Gupta, and N. Kallus, "Data-driven robust optimization," *Mathematical Programming, Series A*, vol. 167, no. 2, pp. 235–292, 2018.

[127] H. Lam and E. Zhou, "Quantifying input uncertainty in stochastic optimization," in *Proceedings of the 2015 Winter Simulation Conference*, IEEE, 2015.

[128] H. Lam, "Recovering best statistical guarantees via the empirical divergence-based distributionally robust optimization," *Operations Research*, vol. 67, no. 4, pp. 1090–1105, 2019.

[129] T. Miyato, S.-i. Maeda, M. Koyama, K. Nakae, and S. Ishii, "Distributional smoothing with virtual adversarial training," *arXiv:1507.00677 [stat.ML]*, 2015.

[130] J. C. Duchi, P. W. Glynn, and H. Namkoong, "Statistics of robust optimization: A generalized empirical likelihood approach," *Mathematics of Operations Research*, vol. 46, pp. 946–969, 2021.

[131] J. C. Duchi and H. Namkoong, "Learning models with uniform performance via distributionally robust optimization," *Annals of Statistics*, vol. 49, no. 3, pp. 1378–1406, 2021.

[132] E. Erdoğan and G. Iyengar, "Ambiguous chance constrained problems and robust optimization," *Mathematical Programming*, vol. 107, no. 1-2, pp. 37–61, 2006.

[133] P. M. Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations," *Mathematical Programming, Series A*, vol. 171, no. 1–2, pp. 115–166, 2018.

[134] S. Shafieezadeh-Abadeh, P. M. Esfahani, and D. Kuhn, "Distributionally robust logistic regression," in *Advances in Neural Information Processing Systems 28*, 2015, pp. 1576–1584.

[135] J. Blanchet, Y. Kang, and K. Murthy, "Robust Wasserstein profile inference and applications to machine learning," *Journal of Applied Probability*, vol. 56, no. 3, pp. 830–857, 2019.

[136] R. Gao and A. J. Kleywegt, "Distributionally robust stochastic optimization with Wasserstein distance," *Mathematics of Operations Research*, 2022.

[137] J. Blanchet, Y. Kang, F. Zhang, and K. Murthy, "Data-driven optimal transport cost selection for distributionally robust optimizatio," *arXiv:1705.07152 [stat.ML]*, 2017.

[138] R. Volpi, H. Namkoong, O. Sener, J. Duchi, V. Murino, and S. Savarese, "Generalizing to unseen domains via adversarial data augmentation," in *Advances in Neural Information Processing Systems 31*, 2018.

[139] M. Staib and S. Jegelka, "Distributionally robust optimization and generalization in kernel methods," in *Advances in Neural Information Processing Systems*, 2019, pp. 9131–9141.

[140] J.-J. Zhu, W. Jitkrittum, M. Diehl, and B. Schölkopf, "Kernel distributionally robust optimization: Generalized duality theorem and stochastic approximation," in *International Conference on Artificial Intelligence and Statistics*, PMLR, 2021, pp. 280–288.

[141] J. Duchi, T. Hashimoto, and H. Namkoong, "Distributionally robust losses for latent covariate mixtures," *arXiv:2007.13982 [stat.ML]*, 2020.

[142] S. Sagawa, P. W. Koh, T. B. Hashimoto, and P. Liang, "Distributionally robust neural networks for group shifts: On the importance of regularization for worst-case generalization," in *Proceedings of the Seventh International Conference on Learning Representations*, 2019.

[143] S. Jeong and H. Namkoong, "Robust causal inference under covariate shift via worst-case subpopulation treatment effect," in *Proceedings of the Thirty Third Annual Conference on Computational Learning Theory*, 2020.

[144] S. Barocas, M. Hardt, and A. Narayanan, *Fairness and Machine Learning*. fairmlbook.org, 2019.

[145] S. Corbett-Davies and S. Goel, "The measure and mismeasure of fairness: A critical review of fair machine learning," *arXiv:1808.00023 [cs.CY]*, 2018.

[146] A. Chouldechova, "A study of bias in recidivism prediciton instruments," *Big Data*, pp. 153–163, 2017.

[147] M. Feldman, S. A. Friedler, J. Moeller, C. Scheidegger, and S. Venkatasubramanian, "Certifying and removing disparate impact," in *proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, 2015, pp. 259–268.

[148] S. Barocas and A. D. Selbst, "Big data's disparate impact," *104 California Law Review*, vol. 3, pp. 671–732, 2016.

[149] M. Hardt, E. Price, and N. Srebro, "Equality of opportunity in supervised learning," in *Advances in Neural Information Processing Systems 29*, 2016.

[150] J. Kleinberg, S. Mullainathan, and M. Raghavan, "Inherent trade-offs in the fair determination of risk scores," in *Proceedings of the 8th Conference on Innovations in Theoretical Computer Science (ITCS)*, 2016.

[151] B. Woodworth, S. Gunasekar, M. I. Ohannessian, and N. Srebro, "Learning non-discriminatory predictors," in *Conference on Learning Theory*, PMLR, 2017, pp. 1920–1953.

[152] M. Kearns, S. Neel, A. Roth, and Z. S. Wu, "Preventing fairness gerrymandering: Auditing and learning for subgroup fairness," *arXiv:1711.05144 [cs.LG]*, 2018.

[153] M. Kearns, S. Neel, A. Roth, and Z. S. Wu, "An empirical study of rich subgroup fairness for machine learning," in *Proceedings of the Conference on Fairness, Accountability, and Transparency*, ACM, 2019, pp. 100–109.

[154] Ú. Hébert-Johnson, M. P. Kim, O. Reingold, and G. N. Rothblum, "Calibration for the (computationally-identifiable) masses," *arXiv:1711.08513 [cs.LG]*, 2017.

[155] J. Rawls, *A theory of justice*. Harvard university press, 2009.

[156] T. Hashimoto, M. Srivastava, H. Namkoong, and P. Liang, "Fairness without demographics in repeated loss minimization," in *Proceedings of the 35th International Conference on Machine Learning*, 2018.

[157] S. Santurkar, D. Tsipras, and A. Madry, "Breeds: Benchmarks for subpopulation shift," *arXiv:2008.04859 [cs.CV]*, 2020.

[158] M. Mitchell *et al.*, "Model cards for model reporting," in *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 2019, pp. 220–229.

[159] I. W. P. Consortium, "Estimation of the warfarin dose with clinical and pharmacogenetic data," *New England Journal of Medicine*, vol. 360, no. 8, pp. 753–764, 2009.

[160] A. Shapiro, D. Dentcheva, and A. Ruszczyński, *Lectures on Stochastic Programming: Modeling and Theory*. SIAM and Mathematical Programming Society, 2009.

[161] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *Journal of Risk*, vol. 2, pp. 21–42, 2000.

[162] R. T. Rockafellar and S. Uryasev, "Conditional value-at-risk for general loss distributions," *Journal of Banking & Finance*, vol. 26, no. 7, pp. 1443–1471, 2002.

[163] P. L. Bartlett, O. Bousquet, and S. Mendelson, "Local Rademacher complexities," *Annals of Statistics*, vol. 33, no. 4, pp. 1497–1537, 2005.

[164] V. Koltchinskii, "Local Rademacher complexities and oracle inequalities in risk minimization," *Annals of Statistics*, vol. 34, no. 6, pp. 2593–2656, 2006.

[165] A. W. van der Vaart and J. A. Wellner, *Weak Convergence and Empirical Processes: With Applications to Statistics*. New York: Springer, 1996.

[166] G. Christie, N. Fendley, J. Wilson, and R. Mukherjee, "Functional map of the world," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6172–6180.

[167] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794.

[168] D. S. Char, N. H. Shah, and D. Magnus, "Implementing machine learning in health care—addressing ethical challenges," *New England Journal of Medicine*, vol. 378, no. 11, p. 981, 2018.

[169] A. Rajkomar, M. Hardt, M. D. Howell, G. Corrado, and M. H. Chin, "Ensuring fairness in machine learning to advance health equity," *Annals of Internal Medicine*, 2018.

[170] S. N. Goodman, S. Goel, and M. R. Cullen, "Machine learning, health disparities, and causal reasoning," *Annals of Internal Medicine*, 2018.

[171] American Medical Association, *AMA passes first policy recommendations on augmented intelligence.* 2018.

[172] B. Abelson, R. Kush, and J. Sun, "Targeting direct cash transfers to the extremely poor," *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2014.

[173] S. Han *et al.*, "Learning to score economic development from satellite imagery," in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2020, 2970–2979.

[174] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the International Conference on Computer Vision*, 2017, pp. 4700–4708.

[175] M. Arjovsky, L. Bottou, I. Gulrajani, and D. Lopez-Paz, "Invariant risk minimization," in *Proceedings of the Eighth International Conference on Learning Representations*, 2020.

[176] Y. Chen, J. Li, H. Xiao, X. Jin, S. Yan, and J. Feng, "Dual path networks," *arXiv:1707.01629 [cs.CV]*, 2017.

[177] J. Miller *et al.*, "Accuracy on the line: On the strong correlation between out-of-distribution and in-distribution generalization," in *Proceedings of the 38th International Conference on Machine Learning*, 2021.

[178] M. Wortsman *et al.*, "Robust fine-tuning of zero-shot models," *arXiv:2109.01903 [cs.CV]*, 2021.

[179] M. Srivastava, T. Hashimoto, and P. Liang, "Robustness to spurious correlations via human annotations," in *Proceedings of the 37th International Conference on Machine Learning*, PMLR, 2020, pp. 9109–9119.

[180] I. Pinelis, "Optimum bounds for the distributions of martingales in banach spaces," *The Annals of Probability*, pp. 1679–1706, 1994.

[181] D. B. Brown, "Large deviations bounds for estimating conditional value-at-risk," *Operations Research Letters*, vol. 35, no. 6, pp. 722–730, 2007.

[182] L. Prashanth, K. Jagannathan, and R. K. Kolla, "Concentration bounds for cvar estimation: The cases of light-tailed and heavy-tailed distributions," in *Proceedings of the 37th International Conference on Machine Learning*, 2020.

[183] M. J. Wainwright, *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*. Cambridge University Press, 2019.

Appendix A: Dynamic Fair Allocation

A.1 Proofs

A.1.1 Proof of Lemma 2

Fix $P \in \mathcal{P}$. By Lemma 25, P-almost surely, there is a type-based policy $\boldsymbol{x}$ that is hindsight optimal, i.e., $\mathsf{OPT} = w(\sum_{t\in[T]} u^i_{\theta_t}(x^i_{\theta_t}))$. Hence,

$$\mathbb{E}[\mathsf{OPT}] = \mathbb{E}\left[w\left(\sum_{t\in[T]} u^i_{\theta_t}(x^i_{\theta_t})\right)\right] \leq w\left(\sum_{t\in[T]} \mathbb{E}[u^i_{\theta_t}(x^i_{\theta_t})]\right) = w\left(\sum_{t\in[T]} \mathbb{E}[u^i_{\theta}(x^i_{\theta})]\right) \leq \mathsf{FLU},$$

where the first equality follows by definition of $\boldsymbol{x}$, the first inequality by concavity of the welfare metric w, the second equality because the input is i.i.d., and the second inequality by definition of FLU.

A.1.2 Proof of Lemma 3

Fix type-based policy $\boldsymbol{x}$ and welfare metric $w \equiv w_q$—we omit the notation for convenience in the proof. Clearly, we can assume $\mathbb{E}[y^i_\theta] = \mathbb{E}[u^i_\theta(x^i_\theta)] > 0$ for every agent i because otherwise $w(\mathbb{E}[U_T]) = Tw(\mathbb{E}[y_\theta]) = 0$ and the lemma is trivial. It suffices to show $\mathbb{E}[w(U_T/T)] - w(\mathbb{E}[U_T]/T) = O(T^{-1/2})$ because the welfare function $w(\cdot)$ is positively homogeneous. First notice

$$w\left(\frac{U_T}{T}\right) \geq w\left(\frac{\mathbb{E}[U_T]}{T} \min_i \frac{U^i_T}{\mathbb{E}[U^i_T]}\right) = w\left(\frac{\mathbb{E}[U_T]}{T}\right) \min_i \frac{U^i_T}{\mathbb{E}[U^i_T]} = w\left(\frac{\mathbb{E}[U_T]}{T}\right) \min_i \left\{1 + \frac{\Delta U^i_T}{\mathbb{E}[U^i_T]}\right\},$$

where the inequality follows because w is nonnegative and monotone, the first equality because w is positively homogeneous, and the second equality by definition of $\Delta U_T = U_T - \mathbb{E}[U_T]$.

Hence,

$$1 - \frac{\mathbb{E}[w(U_T/T)]}{w(\mathbb{E}[U_T]/T)} \le \mathbb{E}\left[\max_{i\in[n]}\left(-\frac{\Delta U_T^i}{\mathbb{E}[U_T^i]}\right)\right]$$
$$\le \sqrt{\mathbb{E}\left(\max_{i\in[n]}\left(-\frac{\Delta U_T^i}{\mathbb{E}[U_T^i]}\right)\right)^2}$$
$$\le \sqrt{\sum_{i\in[n]}\mathbb{E}\left(-\frac{\Delta U_T^i}{\mathbb{E}[U_T^i]}\right)^2}$$
$$= \sqrt{\frac{1}{T}\sum_{i\in[n]}\frac{\mathsf{Var}(y_\theta^i)}{\mathbb{E}[y_\theta^i]^2}},$$

where the first inequality follows because $w(\mathbb{E}[U_T]/T) > 0$, the second inequality by Jensen's inequality, the third inequality by nonnegativity of squares, and the equality because $U_T = \sum_{t\in[T]} y_{\theta_t}$. The proof is thereby concluded.

A.1.3 Proof of Theorem 1

By Lemma 23, $\mathbb{E}[y_\theta^i] = \mathbb{E}[u_\theta^i(x_\theta^i)] > 0$. We present and show the following claim, to which the theorem statement is an immediate corollary, as any Hölder-mean welfare metric w_q with $q \in (-\infty, 1]$ is concave and positively homogenous on $\mathbb{R}_+^n$ and twice continuously differentiable in $\mathbb{R}_{++}^n$.

Claim: Under welfare metric $w : \mathbb{R}_+^n \to \mathbb{R}_+$ that is concave and positively homogeneous on $\mathbb{R}_+^n$ and twice continuously differentiable in $\mathbb{R}_{++}^n$, for any type-based policy $\boldsymbol{x}$ such that $\mathbb{E}[y_\theta] = \mathbb{E}[u_\theta(x_\theta)] > 0$, we have $w(\mathbb{E}[U_T]) - \mathbb{E}[w(U_T)] = O(1)$ uniformly bounded over time horizon length T.

Proof To prove the claim, it suffices to show $w(\mathbb{E}[U_T]/T) - \mathbb{E}[w(U_T/T)] = O(T^{-1})$, as w is positively homogeneous.

We first examine the perturbed welfare around $\mathbb{E}[y_\theta] = \mathbb{E}[U_T]/T$. Consider the high-probability event $E_{1/2} = \{\mathbb{E}[U_T]/2 \le U_T \le 3\mathbb{E}[U_T]/2\}$, whose probability according to Lemma 26 is at least

$\mathbb{P}(E_{1/2}) \geq 1 - \sum_{i\in[n]} 2\exp\left(-\mathbb{E}[y_\theta^i]^2 T/8\right)$. On the event $E_{1/2}$, we defined a hyper-rectangular region $R_{1/2} := \prod_{i\in[n]}[\frac{1}{2}\mathbb{E}y_\theta^i, \frac{3}{2}\mathbb{E}y_\theta^i] \subseteq \mathbb{R}_{++}^n$, so that both $U_T/T, \mathbb{E}[U_T]/T \in R_{1/2}$. Recall $\Delta U_T := U_T - \mathbb{E}[U_T]$. Note $w(U_T/T)$ is twice continuously differentiable on the event $E_{1/2}$, so for some $\alpha \in [0,1]$,

$$\begin{aligned} w\left(\frac{U_T}{T}\right) &= w\left(\frac{\mathbb{E}[U_T]}{T}\right) + \nabla w\left(\frac{\mathbb{E}[U_T]}{T}\right)\left(\frac{\Delta U_T}{T}\right) + \frac{1}{2}\left(\frac{\Delta U_T}{T}\right)^\top \nabla^2 w\left(\frac{\mathbb{E}[U_T]+\alpha\Delta U_T}{T}\right)\left(\frac{\Delta U_T}{T}\right), \\ &\geq w\left(\frac{\mathbb{E}[U_T]}{T}\right) + \nabla w\left(\frac{\mathbb{E}[U_T]}{T}\right)\left(\frac{\Delta U_T}{T}\right) + \frac{\Lambda}{2}\left\|\frac{\Delta U_T}{T}\right\|_2^2, \end{aligned}$$

where the equality follows by the mean value theorem, and the inequality because $\mathbb{E}[U_T]+\alpha\Delta U_T \in R_{1/2}$, and we define the constant

$$\Lambda := \min\left\{\lambda_{\min}\left(\nabla^2 w(\hat{y})\right) : \hat{y}^i \in R_{1/2} = \left[\frac{1}{2}\mathbb{E}[y_\theta^i], \frac{3}{2}\mathbb{E}[y_\theta^i]\right]\right\},$$

an lower bound on all eigenvalues of the Hessian matrix $\nabla^2 w$ in the hyper-rectangule $R_{1/2}$. Clearly Λ is well-defined as w is twice continuously differentiable in $\mathbb{R}_{++}^n$, $\Lambda \leq 0$ because the welfare metric is concave, and $\Lambda > -\infty$ because the hyper-rectangle $R_{1/2}$ is compact.

We have now constructed a local strong smoothness condition around $\mathbb{E}[y_\theta]$, and we are to leverage this property on the entire sample space. We then obtain

$$\begin{aligned} w\left(\frac{\mathbb{E}[U_T]}{T}\right) - \mathbb{E}\left[w\left(\frac{U_T}{T}\right)\right] &\leq w(\mathbb{E}[y_\theta]) - \mathbb{E}\left[w\left(\frac{U_T}{T}\right); E_{1/2}\right] \\ &\leq w(\mathbb{E}[y_\theta]) - \mathbb{E}\left[w(\mathbb{E}[y_\theta]) + \nabla w(\mathbb{E}[y_\theta])\frac{\Delta U_T}{T} + \frac{\Lambda}{2}\left\|\frac{\Delta U_T}{T}\right\|_2^2; E_{1/2}\right] \\ &\leq \mathbb{E}\left[w(\mathbb{E}[y_\theta]) + \nabla w(\mathbb{E}[y_\theta])\frac{\Delta U_T}{T}; E_{1/2}^\complement\right] + \frac{-\Lambda}{2}\mathbb{E}\left\|\frac{\Delta U_T}{T}\right\|_2^2 \\ &\leq \left(w(\mathbb{E}[y_\theta]) + \|\nabla w(\mathbb{E}[y_\theta])\|_1\right)\mathbb{P}(E_{1/2}^\complement) + \frac{-\Lambda \mathsf{Var}(y_\theta^i)}{2T} \\ &\leq 2\left(w(\mathbb{E}[y_\theta]) + \|\nabla w(\mathbb{E}[y_\theta])\|_1\right)\sum_{i\in[n]}\exp\left(-\frac{\mathbb{E}[y_\theta^i]^2}{8}T\right) + \frac{-\Lambda \mathsf{Var}(y_\theta^i)}{2T} \\ &\leq O(T^{-1}), \end{aligned}$$

where the first inequality follows because $\mathbb{E}[U_T] = T\mathbb{E}[y_\theta]$ and $w(\cdot) \geq 0$, the second by the local strong smoothness condition above, the third because $\mathbb{E}[\Delta U_T] = 0$ and $\|\Delta U_T/T\|_2^2 \geq 0$, the fourth because $\|\Delta U_T/T\|_\infty \leq T$ and $\Delta U_T = \sum_{t\in[T]}(y^i_{\theta_t} - \mathbb{E}[y^i_{\theta_t}])$, and the fifth by Lemma 26. The proof is thereby concluded.

□

A.1.4 Proof of Lemma 4

Consider the arrival distribution P where $u^i(x) = \beta^i x$ and $\beta \sim \text{Unif}\{(1, 1/2), (1/2, 1)\}$. Denoting the two types by $\Theta_P = \{1, 2\}$ and the arrival counts by N_1 and N_2, the hindsight optimum is $\mathsf{OPT} = T/2 - |N_1 - N_2|/6$. WLOG suppose $N_1 \leq N_2$, then this is attained by the offline policy $x_1 = (1, 0)$ and $x_2 = (4N_2 - 2T, 2T - N_2)/3N_2$. The fluid policy is $\mathsf{F}_1 = (1, 0)$ and $\mathsf{F}_2 = (0, 1)$, which achieves only $\mathsf{ALG}(\mathsf{F}) = T/2 - |N_1 - N_2|/2$. Hence, Lemma 29 implies

$$\mathcal{R}_T(\mathsf{F}) = \mathbb{E}[\mathsf{OPT} - \mathsf{ALG}(\mathsf{F})] = \frac{2}{3}\mathbb{E}\left|N_1 - \frac{T}{2}\right| = \Theta(\sqrt{T}).$$

A.1.5 Proof of Lemma 5

Denote the dual variables corresponding to the constraints in the primal problem (1.5) by $\nu \in \mathbb{R}^n_+$, $\lambda \in \mathbb{R}^{|\Theta|}$ and $\boldsymbol{\mu} \in \mathbb{R}^{n\times|\Theta|}_+$ so that the Lagrangian function is

$$\begin{aligned} L_\rho(\boldsymbol{x}, U; \nu, \lambda, \boldsymbol{\mu}) &= U + \sum_i \nu^i \left(\rho^i + \mathbb{E}_\theta[u^i_\theta(x^i_\theta)] - U\right) + \sum_\theta \lambda_\theta \left(1 - \sum_i x^i_\theta\right) + \sum_\theta \sum_i \mu^i_\theta x^i_\theta, \\ &= \nu^\top \rho + \mathbb{E}_\theta\left[\nu^\top u_\theta(x_\theta)\right] + U(1 - \mathbf{1}^\top \nu) + \sum_\theta \lambda_\theta \left(1 - \sum_i x^i_\theta\right) + \sum_\theta \sum_i \mu^i_\theta x^i_\theta, \end{aligned} \tag{A.1.1}$$

where we abuse notation to allow $u^i_\theta(\cdot)$ to take values equal to $-\infty$ for arguments outside $[0, 1]$, so that it is still concave. Define the dual function $q_\rho(\nu, \lambda, \boldsymbol{\mu}) := \sup\{L_\rho(\boldsymbol{x}, U; \nu, \lambda, \boldsymbol{\mu}) : \boldsymbol{x} \in$

$\mathbb{R}^{n\times|\Theta|}, U \in \mathbb{R}\}$, so that the optimum of the dual problem can be written as

$$
\begin{aligned}
\inf_{\substack{\nu\geq 0\\ \lambda\in\mathbb{R}^{|\Theta|}\\ \boldsymbol{\mu}\geq 0}} q_\rho(\nu,\lambda,\boldsymbol{\mu}) &= \inf_{\substack{\nu\in\Delta_n\\ \lambda\in\mathbb{R}^{|\Theta|}\\ \boldsymbol{\mu}\geq 0}} \sup_{\substack{\boldsymbol{x}\in\mathbb{R}^{n\times|\Theta|}\\ U\in\mathbb{R}}} L_\rho(\boldsymbol{x},U;\nu,\lambda,\boldsymbol{\mu}) \\
&= \inf_{\nu\in\Delta_n} \inf_{\substack{\lambda\in\mathbb{R}^{|\Theta|}\\ \boldsymbol{\mu}\geq 0}} \sup_{\boldsymbol{x}\in\mathbb{R}^{n\times|\Theta|}} \left\{\nu^\top\rho + \mathbb{E}_\theta\left[\nu^\top u_\theta(x_\theta)\right] + \sum_\theta \lambda_\theta\left(1-\sum_i x_\theta^i\right) + \sum_\theta\sum_i \mu_\theta^i x_\theta^i\right\} \\
&= \inf_{\nu\in\Delta_n} \sup_{\boldsymbol{x}\in\mathbb{R}^{n\times|\Theta|}} \inf_{\substack{\lambda\in\mathbb{R}^{|\Theta|}\\ \boldsymbol{\mu}\geq 0}} \left\{\nu^\top\rho + \mathbb{E}_\theta\left[\nu^\top u_\theta(x_\theta)\right] + \sum_\theta \lambda_\theta\left(1-\sum_i x_\theta^i\right) + \sum_\theta\sum_i \mu_\theta^i x_\theta^i\right\} \\
&= \inf_{\nu\in\Delta_n} \sup_{\boldsymbol{x}\in\mathbb{R}^{n\times|\Theta|}} \left\{\nu^\top\rho + \mathbb{E}_\theta\left[\nu^\top u_\theta(x_\theta)\right] + \sum_\theta \inf_{\lambda_\theta\in\mathbb{R}} \lambda_\theta\left(1-\sum_i x_\theta^i\right) + \sum_\theta\sum_i \inf_{\mu_\theta^i\geq 0} \mu_\theta^i x_\theta^i\right\} \\
&= \inf_{\nu\in\Delta_n} \sup_{\boldsymbol{x}\in\Delta_n^{|\Theta|}} \left\{\nu^\top\rho + \mathbb{E}_\theta\left[\nu^\top u_\theta(x_\theta)\right]\right\} \\
&= D_J(\rho),
\end{aligned}
$$

where the first equality follows because $q_\rho(\nu,\lambda,\boldsymbol{\mu}) = +\infty$ if $\mathbf{1}^\top\nu \neq 1$, the second because the dependence on U vanishes when $\nu \in \Delta_n$, the third because the inner objective is affine in $(\lambda, \boldsymbol{\mu})$ and concave in $\boldsymbol{x}$, the fourth because the objective is separable, the fifth because the inner infima are zero, and the last because the problem is separable across θ.

Strong duality follows by Slater's condition.

A.1.6 Proof of Lemma 6

We know strong duality holds by Lemma 5, so $\boldsymbol{x}_0 \in \arg\max L(\boldsymbol{x}, U; \nu_0, \lambda, \boldsymbol{\mu})$, and hence (a) holds. We can see (c) implies (b) because of the complementary slackness condition $\nu_0^i\left(\mathbb{E}_\theta[u_\theta^i((x_0)_\theta^i)] - J(0)\right) = 0$ (A.2.1). Hence, it suffices to show (c). Suppose otherwise, i.e., there is some agent j where $\nu_0^j = 0$.

We first show $(x_0)_\theta^j = 0$ for almost every θ. If $(x_0)_\theta^j > 0$ for some θ with positive $p_\theta > 0$, then $(\mu_0)_\theta^j = 0$ by complementary slackness (A.2.1); since $\nu_0^j = 0$, the first-order optimality condition implies $(\lambda_0)_\theta = 0$. This means $\nu_0^i p_\theta (u_\theta^i)'((x_0)_\theta^i) = -(\mu_0)_\theta^i \leq 0$ for all i, so $\nu_0^i = 0$ for all i because $p_\theta > 0$ and $(u_\theta^i)'(\cdot) > 0$. This implies $\sum_{i\in[n]} \nu^i = 0 < 1$, which is absurd. Hence, $(x_0)_\theta^j = 0$ for almost every θ, implying $J(0) \leq \mathbb{E}_\theta[u^j_\theta((x_0)^j_\theta)] = 0$.

Now if $\nu^i_0 > 0$ for some agent i, then its complementary slackness condition implies $\mathbb{E}_\theta[u^i_\theta((x_0)^i_\theta)] = J(0) = 0$, so $(x_0)^i_\theta = 0$ for almost every θ. Therefore, $(x_0)^i_\theta = 0$ for almost every θ and every agent i, no matter $\nu^i_0 = 0$ or $\nu^i_0 > 0$. This violates the primal feasibility condition $\sum_i x^i_\theta = 1$ for almost every type θ. The proof of (c) is thereby concluded.

A.1.7 Proof of Theorem 2

As a surrogate for estimating the performance of the BIR policy, we consider a lower bound given by its performance up to τ, the stopping time when $\|\tilde{\rho}_t\|_\infty \geq \delta$ and the guarantee for the well-behavior conditions in Assumption 1 fails to continue to hold. To be more precise, the performance of the BIR policy can be written as

$$\mathsf{ALG}(\mathsf{BIR}) = \min_i U^i_T \geq \min_i U^i_\tau = \overline{U}_\tau - \left\|(\tilde{U}_\tau)_-\right\|_\infty \geq \overline{U}_\tau - \left\|\tilde{U}_\tau\right\|_\infty, \tag{A.1.2}$$

where the first equality follows by definition of the egalitarian welfare, the first inequality because $\tau \leq T$ P-a.s., the second equality because $\left\|(\tilde{U}_\tau)_-\right\|_\infty = \max_i(\overline{U}_\tau - U^i_\tau)$, and the second inequality because $|(\tilde{U}_\tau)_-| \leq |\tilde{U}_\tau|$.

We bound the two terms separately in the following lemmata. The proofs can be found in Sections A.1.8 and A.1.9.

Lemma 12. *Fix $P \in \mathcal{P}$. Let Assumption 1 hold. Under the* BIR *policy with re-solving schedule set by f,*

$$\mathbb{E}[\overline{U}_\tau] \geq \mathbb{E}[\tau]J(0) - \frac{\kappa_J}{2}\sum_{t=2}^{T}\frac{1}{f(t)}.$$

Lemma 13. *Fix $P \in \mathcal{P}$. Under the* BIR *policy, $\left\|\tilde{U}_\tau\right\|_\infty \leq \delta(T-\tau) + (1+\delta)$ P-a.s.*

Hence, we know

$$\mathcal{R}_T(\mathsf{BIR}) \leq TJ(0) - \mathbb{E}\left[\overline{U}_\tau - \left\|\tilde{U}_\tau\right\|_\infty\right] \leq (J(0)+\delta)\mathbb{E}[T-\tau] + (1+\delta) + \frac{\kappa_J}{2}\sum_{t=2}^{T}\frac{1}{f(t)},$$

where the first inequality follows because because $\mathbb{E}[\mathsf{OPT}] \leq \mathsf{FLU} = TJ(0)$ by Lemma 2 and $\mathsf{ALG}(\mathsf{BIR}) \geq \overline{U}_\tau - \|\tilde{U}_\tau\|_\infty$ by Equation (A.1.2), and the second by Lemmata 12 and 13. The proof is concluded by invoking Proposition 2 to show $\mathbb{E}[T-\tau] = 1$ as $f(t) = \Omega(\sqrt{t \log^{2+\varepsilon} t})$.

A.1.8 Proof of Lemma 12

We first notice

$$\mathbb{E}\left[\overline{U}_\tau\right] = \mathbb{E}\left[\sum_{t\in[\tau]} \overline{y}_t\right] = \mathbb{E}\left[\sum_{t\in[\tau]} \mathbb{E}[\overline{y}_t | \mathsf{F}_{t-1}]\right] = \mathbb{E}\left[\sum_{t\in[\tau]} J(\tilde{\rho}_{L(t)})\right]$$

where the first equality follows by definition of $U_t = \sum_{s\in[t]} y_s$, the second by the Optional Stopping Theorem because $\sum_{s\in[t]}(\overline{y}_s - \mathbb{E}[\overline{y}_s|\mathsf{F}_{s-1}])$ is a martingale, and the last because $\mathsf{F}_{t-1} = \mathsf{F}_{L(t)} \in \mathcal{F}_{L(t)}$ and $\mathbb{E}[\overline{y}_t|\mathsf{F}_{L(t)}] = J(\rho_{L(t)}) - \overline{\rho}_{L(t)} = J(\tilde{\rho}_{L(t)})$ for $t \leq \tau$ by Assumption 1. Since $L(t) < t \leq \tau$, Assumption 1 also implies

$$\begin{aligned} J(\tilde{\rho}_{L(t)}) &\geq J(0) + \nabla J(0)^\top \tilde{\rho}_{L(t)} - \frac{\kappa_J}{2}\left\|\tilde{\rho}_{L(t)}\right\|_2^2 \\ &= J(0) + \nabla J(0)^\top M_{L(t)} - \frac{\kappa_J}{2}\left\|M_{L(t)}\right\|_2^2, \end{aligned}$$

where the equality follows because $\tilde{\rho}_{L(t)} = M_{L(t)}$ by Equation (A.1.3). Hence,

$$\begin{aligned} \mathbb{E}\left[\overline{U}_\tau\right] &\geq \mathbb{E}[\tau]J(0) + \nabla J(0)^\top \mathbb{E}\left[\sum_{t=1}^{\tau} M_{L(t)}\right] - \frac{\kappa_J}{2}\mathbb{E}\left[\sum_{t=1}^{\tau}\left\|M_{L(t)}\right\|_2^2\right] \\ &\geq \mathbb{E}[\tau]J(0) - \frac{\kappa_J}{2}\sum_{t=1}^{T}\mathbb{E}\left\|M_{L(t)}\right\|_2^2, \end{aligned}$$

because the Optional Stopping Theorem implies $\mathbb{E}\sum_{t=1}^{\tau-1} M_{L(t)} = \mathbb{E}\sum_{t\in[\tau]}\mathbb{E}[M_{L(t)}] = 0$ as $\tau \in [T]$ is almost surely bounded, and the L^2-norm is non-negative.

Now it suffices to argue $\sum_{t=1}^{T}\mathbb{E}\left\|M_{L(t)}\right\|_2^2 \leq \sum_{t=2}^{T} 1/f(t)$. First notice since $M_0 = 0$ and martingale differences are orthogonal,

$$\mathbb{E}\left\|M_{L(t)}\right\|_2^2 = \sum_{s=1}^{L(t)} \mathbb{E}\left\|M_s - M_{s-1}\right\|_2^2 \leq \sum_{s=1}^{L(t)} \frac{1}{(T-N(s))^2},$$

as $M_s - M_{s-1} = \Delta\tilde{y}_s/(T-N(s))$ and $\mathbb{E}\left\|\Delta\tilde{y}_s\right\|_2^2 \leq \mathbb{E}\left\|\Delta y_s\right\|_2^2 \leq \mathbb{E}\left\|y_s\right\|_2^2 \leq \mathbb{E}\left\|y_s\right\|_1 \leq \mathbb{E}\left\|\mathsf{F}_{s-1}\right\|_1 = 1$ because $y_s^i = u_{\theta_s}^i(\mathsf{F}_{s-1}^i) \leq \mathsf{F}_{s-1}^i \leq 1$ and $\mathsf{F}_{s-1} \in \Delta_n$ by definition of admissible utility functions. Hence,

$$\begin{aligned}
\sum_{t=1}^{T} \mathbb{E}\left\|M_{L(t)}\right\|_2^2 &\leq \sum_{t=1}^{T}\sum_{s=1}^{L(t)} \frac{1}{(T-N(s))^2} \\
&= \sum_{s=1}^{T-1}\sum_{t=N(s)+1}^{T} \frac{1}{(T-N(s))^2} \\
&= \sum_{s=1}^{T-1} \frac{1}{T-N(s)} \\
&= \sum_{s=1}^{T-1} \frac{1}{f(T-L(s))} \\
&\leq \sum_{s=1}^{T-1} \frac{1}{f(T-s+1)} \\
&= \sum_{t=2}^{T} \frac{1}{f(t)},
\end{aligned}$$

where the first inequality follows from above, the first equality by Fubini, the second equality by summation, the third equality because $T-N(s) = f(T-L(s))$ for $s \in [T-1]$, the second inequality because $L(s) < s$, and the last equality by change of variables. The proof is thereby concluded.

A.1.9 Proof of Lemma 13

By definition $\tilde{U}_\tau = \tilde{U}_{\tau-1} + \tilde{y}_\tau = (T-\tau+1)\rho_{\tau-1} + \tilde{y}_\tau$, so

$$\left\|\tilde{U}_\tau\right\| \leq (T-\tau+1)\left\|\tilde{\rho}_{\tau-1}\right\| + \left\|\tilde{y}_\tau\right\| \leq (T-\tau+1)\delta + 1,$$

where the first inequality follows by the triangle inequality and the second because $\|\tilde{\rho}_{\tau-1}\| < \delta$ by definition of τ and $\|y\| \leq 1$. The proof is thereby concluded.

A.1.10 Proof of Proposition 2

We begin by noticing $1 \leq \tau \leq T$ P-a.s., implying $\mathbb{E}[\tau] = 1 + \sum_{t=1}^{T-1} \mathbb{P}\{\tau > t\} = T - \sum_{t=1}^{T-1} \mathbb{P}\{\tau \leq t\}$, so $\mathbb{E}[T-\tau] = \sum_{t=1}^{T-1} \mathbb{P}\{\tau \leq t\}$. However, it is impossible to bound each probability term without a thorough understanding of the evolution dynamics of the $(\tilde{\rho}_t : t = 0, \ldots, T)$ process.

Notice under a type-based policy, the cumulative utility vector U_t by time t is a random walk in $\mathbb{R}^n_+$, whose t-th step has a drift of $\mathbb{E}[y_t|\mathsf{F}_{t-1}] = \mathbb{E}[y_t|\mathsf{F}_{L(t)}]$. Furthermore, if the fluid policy is computed before τ, Assumption 1 implies a relation between $\rho_{L(t)}$ and $\mathbb{E}[y_t|\mathsf{F}_{L(t)}]$. Hence, we present the following lemma, which states a recursive relation showing that the evolution dynamics of $\tilde{\rho}_t$ up until τ resembles that of a martingale. The proof can be found in Section A.1.12.

Lemma 14. *If $L(t) < \tau$, then*

$$\tilde{\rho}_t = \sum_{s=1}^{t} \frac{\Delta \tilde{y}_s}{T - t \wedge N(s)}.$$

Inspired by the martingale-like evolution dynamics of $\tilde{\rho}_t$, we introduce the following auxiliary martingale process $(M_t : t = 0, \ldots, T)$ adapated to the natural filtration $\mathcal{F}$ and give a relation between the two processes to help bound $\mathbb{E}[\tau]$.

$$\begin{aligned} M_t :=& \sum_{s=1}^{t} \frac{\Delta \tilde{y}_s}{T - N(s)} \\ =& \frac{1}{T - N(t)} \sum_{s=L(t)+1}^{t} \Delta \tilde{y}_s + \sum_{k=1}^{k(t-1)} \frac{1}{T - t_k} \sum_{s=t_k+1}^{t_k} \Delta \tilde{y}_s. \end{aligned}$$

In particular, we let $M_0 = 0$. Clearly,

$$M_{t_k} = \tilde{\rho}_{t_k} = \sum_{\ell=1}^{k} \frac{1}{T - t_\ell} \sum_{s=t_{\ell-1}+1}^{t_\ell} \Delta \tilde{y}_s \text{ if } t_{k-1} \leq \tau. \tag{A.1.3}$$

In other words, M_{t_k} and $\tilde{\rho}_{t_k}$ agree on every re-solving epoch up to the first one after τ. We observe that the terms comprising M_t and $\tilde{\rho}_t$ have identical numerators but denominators differ. Intuitively speaking, the last term in $\tilde{\rho}_t$ is always a weakly larger step than that of M_t, but both processes are calibrated against each other at re-solving epochs. This observation leads to the following result on their evolution dynamics.

Lemma 15. *For any $\delta > 0$, $\tau \geq \tau_M$ P-a.s., where $\tau_M := \inf\{t \in [T] : \|M_t\|_\infty \geq \delta\}$.*

It is immediate from Lemma 15 that $\mathbb{P}\{\tau \leq t\} \leq \mathbb{P}\{\tau_M \leq t\}$ and hence

$$\mathbb{E}[T - \tau] \leq \sum_{t=1}^{T-1} \mathbb{P}\{\tau_M \leq t\}.$$

Since $(M_t)_t$ is a martingale adapted to $\mathcal{F}$, we then invoke Theorem 3.5 in Pinelis [180] to obtain

$$\mathbb{P}\{\tau_M \leq t\} = \mathbb{P}\left\{\max_{s\in[t]} \|M_s\|_\infty \geq \delta\right\} \leq 2\exp\left(-\frac{\delta^2}{2\sum_{s=1}^t \|M_s - M_{s-1}\|_\infty^2}\right),$$

because the Euclidean space $(\mathbb{R}^n, \|\cdot\|_2)$ is a Hilbert space and hence a separable Banach space that is $(2,1)$-smooth. The denominator in the exponential term can be bound as follows.

$$\begin{aligned}
\sum_{s=1}^t \|M_s - M_{s-1}\|_\infty^2 &= \sum_{s=1}^t \frac{\|\Delta\tilde{y}_s\|_\infty^2}{(T - N(s))^2} \\
&\leq \sum_{s=1}^t \frac{1}{(T - N(s))^2} \\
&= \sum_{s=1}^t \frac{1}{f^2(T - L(s))} \\
&\leq \sum_{s=1}^t \frac{1}{f^2(T - s + 1)},
\end{aligned}$$

where the first equality follows by definition of M_s, the first inequality because $N(\cdot)$ is non-decreasing and $y_s \in [0,1]$ a.s., the second equality because $f(T - L(s)) = T - N(s)$, and the second inequality because $L(s) < s$ and f is non-decreasing. Then

$$\begin{aligned}
\mathbb{E}[T-\tau] &\le 2\sum_{t=1}^{T-1} \exp\left(-\frac{\delta^2}{2\sum_{s=1}^{t} f^{-2}(T-s+1)}\right) \\
&= 2\sum_{t=1}^{T-1} \exp\left(-\frac{\delta^2}{2\sum_{s=T-t+1}^{T} f^{-2}(s)}\right) \\
&= 2\sum_{t=1}^{T-1} \exp\left(-\frac{\delta^2}{2\sum_{s=t+1}^{T-1} f^{-2}(s)}\right) \\
&\le 2\sum_{t=1}^{\infty} \exp\left(-\frac{\delta^2}{2\sum_{s=t+1}^{\infty} f^{-2}(s)}\right),
\end{aligned}$$

where the first inequality follows from the last three equations above, the two equalities by reindexing, and the last inequality by non-negativity.

Last, we leverage the asymptotic behavior of f to bound the last series. By assumption, there exists some positive constant $m > 0$ such that $f(t) > m\sqrt{t \log^{2+\varepsilon} t}$ for $t \ge 1$. This means for any $t \ge 1$,

$$\sum_{s=t+1}^{\infty} f^{-2}(s) \le \frac{1}{m^2} \sum_{s=t+1}^{\infty} \frac{1}{s \log^{2+\varepsilon} s} \le \frac{1}{m^2} \int_t^{\infty} \frac{1}{s \log^{2+\varepsilon} s} \, \mathrm{d}s = \frac{1}{(1+\varepsilon) m^2 \log^{1+\varepsilon} t},$$

so we cite Lemma 30 to conclude

$$\mathbb{E}[T-\tau] \le 2\sum_{t=1}^{\infty} \exp\left(-m^2 \delta^2 \log^{1+\varepsilon} t\right) \le O(1).$$

A.1.11 Proof of Proposition 3

Recall $v_0^i > 0$ for every agent i by Lemma 6. Define $r_v := 4\sqrt{n}\kappa_u \delta / \varphi^2 = \min(\kappa_u \gamma, \min_i v_0^i)$ for $\kappa_u > 0$.

1. We first provide some basic facts for use in the sequel.

   By definition of v_0, $J(0) = \phi(v_0)$. Since v_0 is unique by Lemma 7, the Envelope Theorem implies $J(\cdot)$ is differentiable at 0 and the gradient is given by $\nabla J(0) = v_0$ because v_0 is unique, Δ_n is compact, and $\nu^\top \rho + \phi(\nu)$ is continuously differentiable.

2. We proceed to show two local lower bounds for $J(\rho)$, respectively for the cases $\kappa_u = 0$ and $\kappa_u > 0$, and to unify them to a single bound. Recall the optimal dual solution $\nu_\rho \in \arg\min_{\nu \in \Delta_n} \{\nu^\top \rho + \phi(\nu)\}$ for ρ with $\tilde{\rho} \in D_\delta(0)$.

   On the one hand, for any ρ,

$$
\begin{aligned}
J(\rho) - J(0) - \nabla J(0)^\top \rho &= [\nu_\rho^\top \rho + \phi(\nu_\rho)] - \phi(\nu_0) - \nu_0^\top \rho \\
&= (\nu_\rho - \nu_0)^\top \tilde{\rho} + [\phi(\nu_\rho) - \phi(\nu_0)] \\
&\geq (\nu_\rho - \nu_0)^\top \tilde{\rho} + \ell(\left\|\nu_\rho - \nu_0\right\|_2 ; \kappa_u) \\
&\geq \underbrace{-\sqrt{n} \left\|\nu_\rho - \nu_0\right\|_2 \|\tilde{\rho}\|_\infty + \ell(\left\|\nu_\rho - \nu_0\right\|_2 ; \kappa_u)}_{=:\spadesuit},
\end{aligned}
$$

   where the first equality follows because strong duality holds for $J(\cdot)$ by Lemma 5 and $\nabla J(0) = \nu_0$, the second equality because $\rho = \overline{\rho}\mathbf{1} + \tilde{\rho}$ and $\mathbf{1}^\top(\nu - \nu_0) = 0$, the first inequality by Lemma 8, and the second inequality by Cauchy-Schwarz and $\|\cdot\|_2 \leq \sqrt{n}\|\cdot\|_\infty$.

   On the other hand, strong duality holds for $J(\cdot)$ by Lemma 5, i.e., $J(\rho) = \inf_{\nu \in \Delta_n}\{\nu^\top \rho + \phi(\nu)\}$, so $J(\cdot)$ is concave as it is a pointwise infimum. This means $J(\rho) - J(0) - \nabla J(0)^\top \rho \leq 0$ for any ρ. Hence, we know $\spadesuit \leq 0$.

   - For linear utilities ($\kappa_u = 0$), we claim $\nu_\rho = \nu_0$, i.e., the optimal dual variable remains the same. In fact, otherwise $\spadesuit > -\delta \left\|\nu_\rho - \nu_0\right\|_2 + \varphi\gamma \left\|\nu_\rho - \nu_0\right\|_2 > 0$, which is absurd. As a result, $\spadesuit = 0$. Note in this case, we have shown $J(\cdot)$ is in fact locally linear around 0.

   - For concave utilities ($\kappa_u > 0$), we claim $\left\|\nu_\rho - \nu_0\right\|_2 < r_\nu = 4\sqrt{n}\kappa_u\delta/\varphi^2$. In fact, otherwise $\spadesuit > -\sqrt{n}\delta \left\|\nu_\rho - \nu_0\right\|_2 + \frac{\varphi^2}{2\kappa_u} r_\nu (\left\|\nu_\rho - \nu_0\right\|_2 - r_\nu/2) \geq -\sqrt{n}\delta \left\|\nu_\rho - \nu_0\right\|_2 + \frac{\varphi^2}{2\kappa_u} r_\nu \left\|\nu_\rho - \nu_0\right\|_2 /2 = 0$, which is absurd. As a result, $\|\nu - \nu_0\|_2 < r_\nu$, and

$$
\spadesuit = -\left\|\nu_\rho - \nu_0\right\|_2 \|\tilde{\rho}\|_2 + \frac{\varphi^2}{4\kappa_u} \left\|\nu_\rho - \nu_0\right\|_2^2 \geq -\frac{\kappa_u}{\varphi^2} \|\tilde{\rho}\|_2^2 .
$$

To conclude, we have shown in both cases that for any ρ with $\|\tilde{\rho}\|_2 < \delta$,

$$J(\rho) \geq J(0) + \nabla J(0)^\top \rho - \frac{\kappa_u}{\varphi^2} \|\tilde{\rho}\|_2^2 .$$

3. Lastly, we show $J(\rho) = \rho^i + \mathbb{E}_\theta[y_\theta^i]$ for all agent i if $\|\tilde{\rho}\|_2 < \delta$.

   From part (b), we know $\left\|\nu_\rho - \nu_0\right\| < r_\nu \leq \min_i \nu_0^i$ so $\nu_\rho > \mathbf{0}$. The proof is thus concluded by noting the complementary slackness condition (A.2.1), i.e., $\nu_\rho^i(\rho^i+\mathbb{E}_\theta[u_\theta^i((x_\rho)_\theta^i)]-J(\rho)) = 0$ for all agent i.

A.1.12 Proof of Lemma 14

For any t with $L(t) < \tau$,

$$\begin{aligned}
(T-t)\rho_t &= U_t \\
&= \sum_{s=L(t)+1}^{t} y_s + U_{L(t)} \\
&= \sum_{s=L(t)+1}^{t} (\Delta y_s + \mathbb{E}[y_s|\mathcal{F}_{s-1}]) + (T - L(t))\rho_{L(t)} \\
&= \sum_{s=L(t)+1}^{t} \Delta y_s + (t - L(t))J(\rho_{L(t)}) + (T-t)\rho_{L(t)},
\end{aligned}$$

where the first equality follows by definition of ρ_t, the second by the evolution of $U_\cdot$, the third by definitions of Δy_s and $\rho_{L(t)}$, ans the last because $\mathbb{E}[y_s|\mathcal{F}_s] = \mathbb{E}[u_{\theta_s}(x_{L(s)})|\mathcal{F}_s] = \mathbb{E}[u_{\theta_s}(x_{L(s)})|\mathcal{F}_{L(s)}] = J(\rho_{L(s)}) - \rho_{L(s)}$ by $L(t) < \tau$. Hence,

$$\rho_t = \rho_{L(t)} + \frac{t - L(t)}{T-t} J(\rho_{L(t)}) + \frac{1}{T-t} \sum_{s=L(t)+1}^{t} \Delta y_s,$$

so

$$
\begin{aligned}
\tilde{\rho}_t &= \tilde{\rho}_{L(t)} + \frac{1}{T-t} \sum_{s=L(t)+1}^{t} \Delta \tilde{y}_s \\
&= \frac{1}{T-t} \sum_{s=L(t)+1}^{t} \Delta \tilde{y}_s + \sum_{k=1}^{k(t)} \frac{1}{T-t_k} \sum_{s=t_{k-1}+1}^{t_k} \Delta \tilde{y}_s \\
&= \sum_{s=1}^{t} \frac{\Delta \tilde{y}_s}{T - t \wedge N(s)},
\end{aligned}
$$

where the first equality follows by taking the tilde operation on both sides, the second by induction, and the last because $N(s) = t_k$ if $t_{k-1} + 1 \leq s \leq t_k$.

A.1.13 Proof of Lemma 15

It suffices to show $\|M_\tau\| \geq \delta$.

Since $L(\tau) < \tau$, we know $\tilde{\rho}_{L(\tau)} = M_{L(\tau)}$. By definition of the $(\tilde{\rho}_t : t \in [T])$ and $(M_t : t \in [T])$ processes,

$$
\begin{aligned}
M_\tau &= M_{L(\tau)} + \frac{1}{T - N(\tau)} \sum_{s=L(\tau)+1}^{\tau} \Delta y_s \\
&= M_{L(\tau)} + \frac{T-\tau}{T-N(\tau)} (\tilde{\rho}_\tau - \tilde{\rho}_{L(\tau)}) \\
&= \frac{(T-\tau)\tilde{\rho}_\tau - [N(\tau)-\tau]\tilde{\rho}_{L(\tau)}}{T-N(\tau)}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\|M_\tau\| &\geq \frac{(T-\tau)\|\tilde{\rho}_\tau\| - [N(\tau)-\tau]\left\|\tilde{\rho}_{L(\tau)}\right\|}{T-N(\tau)} \\
&\geq \frac{(T-\tau)\delta - [N(\tau)-\tau]\delta}{T-N(\tau)} \\
&= \delta,
\end{aligned}
$$

where the first inequality follows by the triangle inequality and $N(\tau) \geq \tau$, and the second because $\left\|\tilde{\rho}_{L(\tau)}\right\| < \delta$ and $\|\tilde{\rho}_\tau\| \geq \delta$ by definition of τ.

A.1.14 Proof of Lemma 8

Assume in the sequel $\nu \neq \nu_0$. Denote $I_\theta := \{i \in [n] : (x_0)^i_\theta > 0\}$ so that $I_\theta \neq \emptyset$ for almost every type $\theta \in \Theta_P$. Denote $\beta^i_\theta := (u^i_\theta)'((x_0)^i_\theta)$ so that $\beta^i_\theta \in (0, 1]$ by definition of admissible utility functions (Definition 1) and $\nabla u_\theta((x_0)_\theta) = \mathrm{diag}(\beta_\theta)$.

1. We first show using the duality theory that $\nu_0^\top \nabla u_\theta((x_0)_\theta)(\hat{x} - x_0)_\theta = 0$ if $\hat{x}^i_\theta = 0$ for $i \notin I_\theta$.

   From the KKT necessary conditions (A.2.1),

$$
\begin{aligned}
\nu_0^\top \nabla u_\theta((x_0)_\theta)(\hat{x} - x_0)_\theta &= \sum_i \nu_0^i \beta^i_\theta (\hat{x} - x_0)^i_\theta \\
&= p_\theta^{-1} \sum_i ((\lambda_0)_\theta - (\mu_0)^i_\theta)(\hat{x} - x_0)^i_\theta \\
&= p_\theta^{-1} \sum_i -(\mu_0)^i_\theta (\hat{x} - x_0)^i_\theta \\
&= p_\theta^{-1} \sum_i -(\mu_0)^i_\theta \hat{x}^i_\theta \\
&= p_\theta^{-1} \sum_{i \notin I_\theta} -(\mu_0)^i_\theta \hat{x}^i_\theta,
\end{aligned}
$$

   where the first equality follows because $\mathrm{diag}(\beta_\theta) = \nabla u_\theta((x_0)_\theta)$, the second by the first-order stationarity condition, the third because $\mathbf{1}^\top (\hat{x} - x_0)_\theta = 0$, the fourth by the complementary slackness condition, and the last because for $i \in I_\theta$, $(x_0)^i_\theta > 0$ and so $(\mu_0)^i_\theta = 0$ by the complementary slackness condition. The result follows.

2. We next characterize the curvature of $\phi_\theta(\cdot)$ by lower bounding $\phi_\theta(\nu) - \phi_\theta(\nu_0) - \nabla \phi_\theta(\nu_0)^\top (\nu - \nu_0)$ for each type θ based on a subset of feasible primal solutions, where $\nabla \phi_\theta(\nu_0)$ is a subgradient of $\phi_\theta(\cdot)$ at ν_0.

Fix any $\nu \in \Delta_n$. If $\hat{x}^i_\theta = 0$ for $i \notin I_\theta$, then

$$\begin{aligned}
\phi_\theta(\nu) - \phi_\theta(\nu_0) - \nabla\phi_\theta(\nu_0)^\top(\nu - \nu_0) &\geq \nu^\top u_\theta(\hat{x}_\theta) - \nu^\top u_\theta((x_0)_\theta) \\
&\geq \nu^\top \left(\nabla u_\theta((x_0)_\theta)(\hat{x} - x_0)_\theta - \frac{\kappa_u \mathbf{1}}{2} \|(\hat{x} - x_0)_\theta\|_2^2 \right) \\
&= \nu^\top \nabla u_\theta((x_0)_\theta)(\hat{x} - x_0)_\theta - \frac{\kappa_u}{2} \|(\hat{x} - x_0)_\theta\|_2^2 \\
&= \underbrace{(\nu - \nu_0)^\top \mathrm{diag}(\beta_\theta)(\hat{x} - x_0)_\theta - \frac{\kappa_u}{2} \|(\hat{x} - x_0)_\theta\|_2^2,}_{\clubsuit}
\end{aligned}$$

where the first inequality follows by definition of $\phi_\theta(\cdot)$ and Lemma 6, the second inequality by Lemma 22, the first equality because $\mathbf{1}^\top \nu = 1$, and the second equality by part (a) and $\mathrm{diag}(\beta_\theta) = \nabla u_\theta((x_0)_\theta)$.

3. We proceed to pin down the particular choices of $\hat{x}_\theta$, according to ν and κ_u.

   Denote the directional (unitary) vector $d_\nu := (\nu - \nu_0)/\|\nu - \nu_0\|_2$ so that $\mathbf{1}^\top d_\nu = 0$ and $\|d_\nu\|_2 = 1$. Take $\hat{x}_\theta = (x_0)_\theta + r_x(\mathsf{e}^{i_\theta} - \mathsf{e}^{j_\theta})$ where $i_\theta \in \arg\max_{i \in I_\theta} \beta^i_\theta d^i_\nu$ and $j_\theta \in \arg\min_{i \in I_\theta} \beta^i_\theta d^i_\nu$ and $r_x \geq 0$ to be determined later.

   We claim $\hat{x}_\theta$ is feasible as long as $r_x \leq \gamma$, where we recall $\gamma > 0$ is the margin of $\boldsymbol{x}_0$ (Definition 4). Clearly $\mathbf{1}^\top x_\theta = 1$. If $|I_\theta| = 1$ then $i_\theta = j_\theta$ and $\hat{x}_\theta = x^0_\theta \in \Delta_n$. Otherwise, $(x_0)^i_\theta \in [\gamma, 1 - \gamma]$ for any $i \in I_\theta$, so $\hat{x}^i_\theta \in [0, 1]$.

   Now define $f_\theta(d_\nu) := \mathrm{range}_{i \in I_\theta} \beta^i_\theta d^i_\nu$, so that $0 \leq f_\theta(d_\nu) \leq 2$, and

$$\clubsuit = r_x \|\nu - \nu_0\|_2 \cdot f_\theta(d_\nu) - \kappa_u r_x^2.$$

   - Linear utilities ($\kappa_u = 0$). Take $r_x = \gamma$, and

$$\clubsuit = \gamma f_\theta(d_\nu) \|\nu - \nu_0\|_2 .$$

   - Concave utilities ($\kappa_u > 0$). Pick any $r_\nu \in [0, \kappa_u \gamma]$ and take $r_x = \frac{f_\theta(d_\nu)}{2\kappa_u} \min(\|\nu - \nu_0\|_2, r_\nu)$

so that $r_x \leq \gamma f_\theta(d_\nu)/2 \leq \gamma$. Then

$$\clubsuit = \begin{cases} \frac{f_\theta^2(d_\nu)}{4\kappa_u} \|\nu - \nu_0\|_2^2, & \|\nu - \nu_0\|_2 < r_\nu, \\ \frac{f_\theta^2(d_\nu)}{2\kappa_u} r_\nu \left(\|\nu - \nu_0\|_2 - \frac{r_\nu}{2} \right), & \|\nu - \nu_0\|_2 \geq r_\nu. \end{cases}$$

4. We take expectations to obtain the lower envelopes for $\phi(\cdot)$, i.e., $\phi(\nu) - \phi(\nu_0) \geq \ell(\|\nu - \nu_0\|_2 ; \kappa_u)$.

$$\begin{aligned} \phi(\nu) - \phi(\nu_0) &\geq \phi(\nu) - \phi(\nu_0) - \nabla\phi(\nu_0)^\top (\nu - \nu_0) \\ &= \mathbb{E}_\theta \left[\phi_\theta(\nu) - \phi_\theta(\nu_0) - \nabla\phi_\theta(\nu_0)^\top (\nu - \nu_0) \right] \\ &\geq \mathbb{E}_\theta [\clubsuit], \end{aligned}$$

where the first inequality follows by the first-order optimality condition of $\nu_0 \in \arg\min \phi(\cdot)$, the equality because subgradients are additive, and the second inequality by part (b). Now define $\varphi := \inf\{\mathbb{E}_\theta[f_\theta(d_\nu)] : d_\nu \in \mathbb{R}^n, \mathbf{1}^\top d_\nu = 0, \|d_\nu\|_2 = 1\}$. Then we know $\varphi \geq 0$ because the range function is non-negative. We conclude by noticing $\mathbb{E}[f_\theta(d_\nu)] \geq \varphi$ and $\mathbb{E}[f_\theta^2(d_\nu)] \geq \mathbb{E}[f_\theta(d_\nu)]^2 \geq \varphi^2$, so $\mathbb{E}_\theta[\clubsuit] \geq \ell(\|\nu - \nu_0\|_2 ; \kappa_u)$.

5. Lastly, it remains to show $\varphi > 0$ to render this lower envelope nontrivial. Suppose on the contrary $\varphi = 0$.

   First notice the infimum is attained, because $\{d_\nu \in \mathbb{R}^n : \mathbf{1}^\top d_\nu = 0, \|d_\nu\|_2 = 1\}$ is compact. Then $\mathbb{E}_\theta[f_\theta(d_\nu)] = 0$ for some $d_\nu \in \mathbb{R}^n$ with $\mathbf{1}^\top d_\nu = 0$ and $\|d_\nu\|_2 = 1$.

   Since $f_\theta(d_\nu) = \mathrm{range}_{i \in I_\theta} \beta_\theta^i d_\nu^i \geq 0$ by definition, we know for almost every θ that $f_\theta(d_\nu) = 0$, i.e., there is a constant scalar $c_\theta \in \mathbb{R}$ such that $\beta_\theta^i d_\nu^i = c_\theta$ for every i. In other words, for every agent i and almost every type θ,

$$(\beta_\theta^i d_\nu^i - c_\theta) \mathbf{1}\{i \in I_\theta\} = 0.$$

   Hence, $(d, \lambda, \mu) \in \mathbb{R}^n \times \mathbb{R}^\Theta \times \mathbb{R}^{n \times |\Theta|}$ given by $d = d_\nu$, $\lambda_\theta = -p_\theta c_\theta$ and $\mu_\theta^i = -p_\theta \beta_\theta^i d_\nu^i + p_\theta c_\theta$

are a set of multipliers on the gradients of active constraints of the primal problem $J(0)$ (Equation (1.5)) that gives a zero sum. By the regularity condition (Assumption 2), the multipliers should all equal zero. In particular, $d_v = 0$, which is absurd. Therefore, we conclude $\varphi > 0$.

A.1.15 Proof of Lemma 10

Fix distinct $N, \hat{N} \in \mathbb{Z}_+^n$. We first argue that it suffices to show there exists an optimal solution $\hat{\boldsymbol{x}}$ to the offline problem $\mathsf{OPT}(\hat{N}, U_0)$ such that

$$\max_{\theta \in \Theta_P} \left\| N_\theta u_\theta(x_\theta) - \hat{N}_\theta u_\theta(\hat{x}_\theta) \right\|_\infty \leq 2 \left\| \hat{N} - N \right\|_1 . \tag{A.1.4}$$

This is because $\left\|\hat{N} - N\right\|_1 \leq |\Theta_P| \left\|\hat{N} - N\right\|_\infty$, and for any $\hat{\boldsymbol{x}}$,

$$\begin{aligned} |N_\theta x_\theta^i - \hat{N}_\theta \hat{x}_\theta^i| &\leq N_\theta |x_\theta^i - \hat{x}_\theta^i| + |N_\theta - \hat{N}_\theta| \hat{x}_\theta^i \\ &\leq \underline{\beta}^{-1} N_\theta |u_\theta^i(x_\theta^i) - u_\theta^i(\hat{x}_\theta^i)| + |N_\theta - \hat{N}_\theta| \hat{x}_\theta^i \\ &\leq \underline{\beta}^{-1} \left(|N_\theta u_\theta^i(x_\theta^i) - \hat{N}_\theta u_\theta^i(\hat{x}_\theta^i)| + |\hat{N}_\theta - N_\theta| u_\theta^i(x_\theta^i) \right) + |N_\theta - \hat{N}_\theta| \\ &\leq \underline{\beta}^{-1} \left(|N_\theta u_\theta^i(x_\theta^i) - \hat{N}_\theta u_\theta^i(\hat{x}_\theta^i)| + 2 \left\|\hat{N} - N\right\|_\infty \right), \end{aligned}$$

where the first inequality follows by the triangle inequality, the second by the mean value theorem and $(u_\theta^i)'(\cdot) \geq (u_\theta^i)'(1) \geq \underline{\beta}$ because $u_\theta^i(\cdot)$ is admissible, the third again by the triangle inequality and $\hat{x}_\theta^i \in [0, 1]$, and the last because $u_\theta^i(\cdot) \in [0, 1]$ and $\underline{\beta} \in (0, 1]$.

In the sequel, for convenience, denote $\Delta := 2 \left\|\hat{N} - N\right\|_1$ and $\hat{\Theta}_P := \{\theta \in \Theta_P : \hat{N}_\theta > 0\} \subseteq \Theta_P$, which we assume to be nonempty w.l.o.g., i.e., $\hat{N} \neq 0$. In the proof we omit U_0 and denote $\mathsf{OPT}(\cdot, U_0)$ by $\mathsf{OPT}(\cdot)$. We also relax the second constraint from equality to a weak inequality $\sum_{i \in [n]} x_\theta^i \leq 1$ for every type $\theta \in \Theta_P$ and note the formulation is equivalent by Lemma 21.

Suppose the statement (A.1.4) is false, then the following system on $\hat{\boldsymbol{y}} \in \prod_{\theta \in \hat{\Theta}_P} \overline{\mathcal{Y}}_\theta$ is infeasible, where we recall $y_\theta^i = u_\theta^i(x_\theta^i)$ for every agent $i \in [n]$ and every type $\theta \in \Theta_P$.

$$\begin{cases} \hat{N}_\theta \hat{y}_\theta^i \le N_\theta y_\theta^i + \Delta, & \theta \in \hat{\Theta}_P, i \in [n], \\ -\hat{N}_\theta \hat{y}_\theta^i \le -N_\theta y_\theta^i + \Delta, & \theta \in \hat{\Theta}_P, i \in [n], \\ -\sum_{\theta \in \hat{\Theta}_P} \hat{N}_\theta \hat{y}_\theta^i \le U_0^i - \mathsf{OPT}(\hat{N}), & i \in [n], \\ \hat{y}_\theta \in \overline{\mathcal{Y}}_\theta = \{(u_\theta^i(x^i) : i \in [n]) : x \ge 0, \mathbf{1}^\top x \le 1\}, & \theta \in \hat{\Theta}_P \end{cases} \tag{A.1.5}$$

Since the first three sets of constraints are affine, they map the closed and convex set $\overline{\mathcal{Y}}_\theta$ to a closed and convex set. By the generalized Farkas' lemma (Lemma 31), there exist a constant $c \in \mathbb{R}$ and vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}_+^{n \times \hat{\Theta}_P}$ and $\nu \in \mathbb{R}_+^n$ such that

$$\begin{cases} \sum_{\theta \in \hat{\Theta}_P} \sum_{i \in [n]} (a_\theta^i - b_\theta^i - \nu^i) \hat{N}_\theta \hat{y}_\theta^i \ge c, & \forall \hat{\boldsymbol{y}} \in \prod_{\theta \in \Theta_P} \overline{\mathcal{Y}}_\theta, \\ \sum_{\theta \in \hat{\Theta}_P} \sum_{i \in [n]} \left[a_\theta^i (N_\theta y_\theta^i + \Delta) + b_\theta^i (-N_\theta y_\theta^i + \Delta) \right] + \sum_{i \in [n]} \nu^i \left(U_0^i - \mathsf{OPT}(\hat{N}) \right) < c. \end{cases} \tag{A.1.6}$$

First notice $(\boldsymbol{a}, \boldsymbol{b}) \neq 0$ because the last two sets of constraints in (A.1.5) are feasible by definition of $\mathsf{OPT}(\hat{N})$. Next we know $c \le 0$ from the first set of inequalities in (A.1.6) by taking $\hat{\boldsymbol{y}} = 0$. They also imply $a_\theta^i - b_\theta^i - \nu^i \ge 0$ for every agent $i \in [n]$ and every type $\theta \in \hat{\Theta}_P$ with positive arrivals. Plugging these two results into the second inequality in (A.1.6), we obtain

$$\sum_{i \in [n]} \nu^i \sum_{\theta \in \hat{\Theta}_P} N_\theta y_\theta^i + \|\boldsymbol{a} + \boldsymbol{b}\|_1 \Delta + \sum_{i \in [n]} \nu^i \left(U_0^i - \mathsf{OPT}(\hat{N}) \right) < 0.$$

Hence,

$$
\begin{aligned}
\|\boldsymbol{a}+\boldsymbol{b}\|_1 \Delta &< \sum_{i\in[n]} \nu^i \left(\mathsf{OPT}(\hat{N}) - \mathsf{OPT}(N) + \sum_{\theta\in\Theta_P\setminus\hat{\Theta}_P} N_\theta y_\theta^i \right) \\
&\leq \sum_{i\in[n]} \nu^i \left(|\mathsf{OPT}(\hat{N}) - \mathsf{OPT}(N)| + \sum_{\theta\in\Theta_P\setminus\hat{\Theta}_P} |N_\theta - \hat{N}_\theta| \right) \\
&\leq 2\,\|\nu\|_1 \left\|N - \hat{N}\right\|_1 \\
&= \|\nu\|_1 \Delta,
\end{aligned}
$$

where the first inequality follows because $\boldsymbol{x}$ is an optimal solution to $\mathsf{OPT}(N)$ implies $\mathsf{OPT}(N) \leq U_0^i + \sum_{\theta\in\Theta_P} N_\theta y_\theta^i$ for every agent i, the second because $\hat{N}_\theta = 0$ for $\theta \in \Theta_P \setminus \hat{\Theta}_P$, $\nu \geq 0$ and $y_\theta^i \in [0,1]$, and the last by the definition of the L_1-norm and Lemma 16 on the Lipschitzness of the hindsight optimum against arrivals, whose proof can be found in Section A.1.16.

Lemma 16. *Fix finite support $\Theta_P \subseteq \Theta > 0$, i.e., $|\Theta_P| < \infty$. For any $N, \hat{N} \in \mathbb{Z}_+^{\Theta_P}$,*

$$\mathsf{OPT}(N, U_0) - \mathsf{OPT}(\hat{N}, U_0) \leq \left\|N - \hat{N}\right\|_1.$$

Canceling out $\Delta = 2\left\|N - \hat{N}\right\|_1$, we obtain $\|\boldsymbol{a}+\boldsymbol{b}\|_1 < \|\nu\|_1$, which violates

$$\|\boldsymbol{a}+\boldsymbol{b}\|_1 = \sum_{\theta\in\hat{\Theta}_P}\sum_{i\in[n]} (a_\theta^i + b_\theta^i) \geq \sum_{\theta\in\hat{\Theta}_P}\sum_{i\in[n]} (a_\theta^i - b_\theta^i) \geq \sum_{\theta\in\hat{\Theta}_P}\sum_{i\in[n]} \nu^i = |\hat{\Theta}_P|\,\|\nu\|_1 \geq \|\nu\|_1,$$

because $\boldsymbol{b} \geq 0$ and $a_\theta^i - b_\theta^i - \nu^i \geq 0$ for any $i \in [n]$ and $\theta \in \hat{\Theta}_P$. The proof is thereby concluded.

A.1.16 Proof of Lemma 16

By strong duality, $\mathsf{OPT}(N, U_0) = \inf_{\nu\in\Delta_n}\{\nu^\top U_0 + \sum_{\theta\in\Theta_P} N_\theta \phi_{\mathcal{Y}_\theta}(\nu)\}$ for any $N \in \mathbb{Z}_+^n$, so $\mathsf{OPT}(N, U_0) - \mathsf{OPT}(\hat{N}, U_0) \leq \sup_{\nu\in\Delta_n} \sum_{\theta\in\Theta_P} (N_\theta - \hat{N}_\theta)\phi_{\mathcal{Y}_\theta}(\nu) \leq \left\|N - \hat{N}\right\|_1 \sup_{\nu\in\Delta_n} \left\|\phi_{\mathcal{Y}_{(\cdot)}}(\nu)\right\|_\infty$. It suffices to show $\phi_{\mathcal{Y}_\theta}(\nu) = \sup_{y_\theta\in\mathcal{Y}_\theta} \nu^\top y_\theta \leq 1$ for any $\nu \in \Delta_n$ and any type $\theta \in \Theta_P$. This results from $\nu^\top y_\theta \leq \|\nu\|_1 \|y_\theta\|_\infty \leq 1$ because $\nu \in \Delta_n$ and $y_\theta \in [0,1]^n$. The proof is thereby concluded.

A.1.17 Proof of Lemma 9

We denote $\vec{\phi}(\nu) := (\phi_\theta(\nu) : \theta \in \Theta_P)$, so that $\mathbb{E}_p[\phi_\theta(\nu)] = \langle \vec{\phi}(\nu), p \rangle$ for any $p \in \Delta_{\Theta_P}$. We denote by $\Phi(\cdot) := \inf_{\nu \in \Delta_n} \langle \vec{\phi}(\nu), \cdot \rangle$ the fluid benchmark parameterized by the probability distribution. For example, $\mathsf{FLU} = T\Phi(p)$, and $\mathsf{OPT} = T\Phi(\hat{p})$ where $\hat{p} := N/T$ denotes the empirical distribution. Clearly $\Phi(p)$ is concave.

1. We first show if $\nu \in \arg\min_{\Delta_n} \langle \vec{\phi}(\cdot), p \rangle$, then the vector $\vec{\phi}(\nu)$ is a supergradient of Φ at p, i.e., $\vec{\phi}(\nu) \in \partial\Phi(p)$. This is because $\Phi(p') \le \langle \vec{\phi}(\nu), p' \rangle = \langle \vec{\phi}(\nu), p \rangle + \langle \vec{\phi}(\nu), p' - p \rangle = \Phi(p) + \langle \vec{\phi}(\nu), p' - p \rangle$ for any $p' \in \Delta_{\Theta_P}$.

2. Next, for $\nu, \hat{\nu} \in \arg\min_{\Delta_n} \mathbb{E}_p[\phi_\theta(\cdot)]$, if $\nu \neq \hat{\nu}$, then $\vec{\phi}(\nu) \neq \vec{\phi}(\hat{\nu})$.

   Fix an optimal primal solution $\boldsymbol{x}$ and denote the other dual variables corresponding to ν (resp., $\hat{\nu}$) by λ and $\boldsymbol{\mu}$ (resp., $\hat{\lambda}$ and $\hat{\boldsymbol{\mu}}$) according to (A.2.2). Suppose $\vec{\phi}(\nu) = \vec{\phi}(\hat{\nu})$. Notice the KKT conditions (A.2.1) implies $\nu^i p_\theta (u^i_\theta)'(x^i_\theta) = \lambda_\theta - \mu^i_\theta$ for every $i \in [n]$ and almost every $\theta \in \Theta_P$, so
$$\nu^i u^i_\theta(x^i_\theta) = (\lambda_\theta - \mu^i_\theta) \frac{u^i_\theta(x^i_\theta)}{p_\theta (u^i_\theta)'(x^i_\theta)} = \lambda_\theta \frac{u^i_\theta(x^i_\theta)}{p_\theta (u^i_\theta)'(x^i_\theta)}$$
   by complementary slackness on $\boldsymbol{\mu}$. This means $\phi_\theta(\nu) = \nu^\top u_\theta(x_\theta) = \lambda_\theta \sum_{i \in [n]} u^i_\theta(x^i_\theta)/p_\theta(u^i_\theta)'(x^i_\theta)$; likewise for $\hat{\nu}$ and $\hat{\lambda}$. Hence, $\lambda_\theta = \hat{\lambda}_\theta$ because $\phi_\theta(\nu) = \phi_\theta(\hat{\nu}) > 0$ for almost every $\theta \in \Theta_P$.

   Now that $\nu \neq \hat{\nu}$, there must be an agent i where $\nu^i \neq \hat{\nu}^j$. Since $\mathsf{FLU} > 0$ by Lemma 23, there must be a type $\theta \in \Theta_P$ where $x^i_\theta > 0$. By the KKT conditions (A.2.1), $\lambda_\theta = \nu^i p_\theta (u^i_\theta)'(x^i_\theta)$; likewise, $\hat{\lambda}_\theta = \hat{\nu}^i p_\theta (u^i_\theta)'(x^i_\theta)$, which is absurd because $\nu^i \neq \hat{\nu}^i$ but $\lambda_\theta = \hat{\lambda}_\theta$.

3. We then claim $\mathsf{FLU} - \mathbb{E}[\mathsf{OPT}] = \Omega(\varepsilon^\top \Sigma \varepsilon \sqrt{T})$ for some $\varepsilon \in \mathbb{R}^{\Theta_P} \setminus \{0\}$, where $\Sigma := \mathrm{diag}(p) - pp^\top = \mathrm{Cov}(N)/T$.

   We have shown $\Phi(\cdot)$ has two distinct gradients $\phi(\nu), \phi(\hat{\nu}) \in \partial\Phi(p)$ at p given by two distinct optimal dual solutions ν and $\hat{\nu}$. Define $g_0 = (\phi(\nu) + \phi(\hat{\nu}))/2$ and $\varepsilon = (\phi(\nu) - \phi(\hat{\nu}))/2$ so that $\varepsilon \neq 0$ and $g_0 \pm \varepsilon \in \partial\Phi(p)$.

For any supergradient g of $\Phi(\cdot)$ at p, we know $\mathsf{FLU}-\mathsf{OPT} = T(\Phi(p)-\Phi(\hat{p})) \geq T\langle g, p-\hat{p}\rangle$. In particular, we can take $g = g_0 + \varepsilon\,\mathrm{sgn}(\langle \varepsilon, p-\hat{p}\rangle) \in \partial\Phi(p)$, so that

$$\frac{\mathsf{FLU}-\mathbb{E}[\mathsf{OPT}]}{T} \geq \langle g_0, \mathbb{E}[p-\hat{p}]\rangle + \mathbb{E}\,|\langle \varepsilon, p-\hat{p}\rangle| = \frac{\mathbb{E}\,|\varepsilon^\top(N-\mathbb{E}[N])|}{T}$$

because $\mathbb{E}[\hat{p}] = p$ and $\hat{p} = N/T$. Since Lemma 28 implies $\mathbb{E}\,|\varepsilon^\top(N-\mathbb{E}[N])|\,/\sqrt{T} \to \varepsilon^\top\Sigma\varepsilon/\sqrt{2\pi}$, we know $\mathsf{FLU}-\mathbb{E}[\mathsf{OPT}] = \Omega(\varepsilon^\top\Sigma\varepsilon\sqrt{T})$.

4. Last, it remains to argue $\varepsilon^\top\Sigma\varepsilon > 0$.

   Suppose otherwise, then $\mathbb{E}|\varepsilon^\top(p-\hat{p})|^2 = \varepsilon^\top\Sigma\varepsilon/T = 0$, implying $\varepsilon^\top(p-\hat{p}) = 0$ P-a.s. This means $\varepsilon = \alpha\mathbf{1}$ for some scalar $\alpha \in \mathbb{R}\setminus\{0\}$. Recall $\phi(v) = \phi(\hat{v}) + 2\varepsilon = \phi(\hat{v}) + 2\alpha\mathbf{1}$, so $\Phi(p) = \phi(v)^\top p = \phi(\hat{v})^\top p + 2\alpha = \Phi(p) + 2\alpha$, which is absurd. Therefore, $\varepsilon^\top\Sigma\varepsilon > 0$, and the proof is concluded.

A.1.18 Proof of Theorem 3

We notice the last term in the telescopic sum (1.10) is bounded by $\mathsf{OPT}^K - \mathsf{ALG}(\mathsf{BIRT}) \leq T - t_K = 1$ because $t_K = T-1$. Then it suffices to bound each welfare loss term in the finite series.

We show the regret bound separately for linear utilities ($\kappa_u = 0$) and general concave utilities ($\kappa_u > 0$), with the re-solving schedule specified in the theorem statement. In both cases, we find appropriate positive constants $m, \varepsilon > 0$ and take the thresholds

$$\gamma_{t_k} = \frac{2m^2}{3n^2}\frac{\sqrt{(T-t_k)\log^{1+\varepsilon}(T-t_k)}}{T-t_{k+1}} \tag{A.1.7}$$

for the thresholding adjustment at t_k.

1. Suppose all utilities are linear ($\kappa_u = 0$) and the scheduling function f satisfies

$$\Omega(\sqrt[4]{\log^{1+\varepsilon} t/t}) \leq f(t)/t < 1.$$

By the assumption on f, we can take the constant $m > 0$ such that $f(t) > m\sqrt[4]{t^3 \log^{1+\varepsilon} t}$ for any $t \geq 1$, so that the threshold γ_{t_k} used during thresholding adjustment at t_k satisfies the condition in Proposition 4, i.e., $\gamma_{t_k} \leq 2(T - t_{k+1})/3n^2(T - t_k)$. Proposition 4 implies for any $k \in [K]$,

$$\begin{aligned}
\mathbb{E}\left[\mathsf{OPT}^{k-1} - \mathsf{OPT}^{k}\right] &\leq 4|\Theta_P|(t_k - t_{k-1}) \exp\left(-\frac{8\underline{p}^2 m^4}{81L^2n^4} \log^{1+\varepsilon}(T - t_{k-1})\right) \\
&\leq 4|\Theta_P| \sum_{t=t_{k-1}+1}^{t_k} \exp\left(-C \log^{1+\varepsilon}(T - t)\right),
\end{aligned}$$

where the second inequality follows because $\log(\cdot)$ is increasing, and we denote $C := 8\underline{p}^2 m^4/81L^2n^4$ for notational convenience. By recursion we argue

$$\begin{aligned}
\mathbb{E}\left[\mathsf{OPT} - \mathsf{OPT}^{K}\right] &= \sum_{k=1}^{K} \mathbb{E}\left[\mathsf{OPT}^{k-1} - \mathsf{OPT}^{k}\right] \\
&\leq 4|\Theta_P| \sum_{t=1}^{t_K} \exp\left(-C \log^{1+\varepsilon}(T - t)\right) \\
&= 4|\Theta_P| \sum_{t=T-t_K}^{T-1} \exp\left(-C \log^{1+\varepsilon} t\right) \\
&\leq 4|\Theta_P| \sum_{t=1}^{\infty} \exp\left(-C \log^{1+\varepsilon} t\right) \\
&\leq O(1),
\end{aligned}$$

where the first equality follows from telescoping, the first inequality from above, the second equality by reordering indices, the second inequality because $t_K = T - 1$, and the last inequality by Lemma 30.

We therefore conclude $\mathcal{R}_T(\mathsf{BIRT}) \leq= O(1)$.

2. Now consider the general case where utilities can be nonlinear ($\kappa_u > 0$), and assume $f(t)/t \in (\underline{M}, \overline{M})$ for $t \geq 1$.

   By the assumption on f, we can take any constant $m \in (0, \underline{M})$ such that for any $t > 1$, $f(t) >$

mt; w.l.o.g., suppose $\varepsilon \in (0,1)$ so that $\log^{1+\varepsilon} t < t$ for any $t \geq 1$. Hence, the threshold γ_{t_k} used during thresholding adjustment at t_k satisfies the condition in Proposition 4, i.e., $\gamma_{t_k} \leq 2(T - t_{k+1})/3n^2(T - t_k)$. Proposition 4 implies for any $k \in [K]$,

$$\begin{aligned}
\mathbb{E}\left[\mathsf{OPT}^{k-1} - \mathsf{OPT}^{k}\right] &\leq 4|\Theta_P|(t_k - t_{k-1}) \exp\left(-C \log^{1+\varepsilon}(T - t_k)\right) \\
&\qquad + \frac{\kappa_u m^4}{3n^2}\left(\frac{T - t_{k-1}}{T - t_k}\right)^3 \log^{1+\varepsilon}(T - t_{k-1}) \\
&\leq 4|\Theta_P| \sum_{t=t_{k-1}+1}^{t_k} \exp\left(-C \log^{1+\varepsilon}(T - t)\right) + \frac{\kappa_u m^4}{3n^2 \underline{M}^3} \log^{1+\varepsilon} T,
\end{aligned}$$

where the first inequality follows from Proposition 4, and the second because $T - t_k = f(T - t_{k-1}) > \underline{M}(T - t_{k-1})$. Recall the constant $C = 8\underline{p}^2 m / 81 L^2 n^4$. Similarly to the linear case, we argue recursively

$$\begin{aligned}
\mathbb{E}\left[\mathsf{OPT} - \mathsf{OPT}^{K}\right] &= \sum_{k=1}^{K} \mathbb{E}\left[\mathsf{OPT}^{k-1} - \mathsf{OPT}^{k}\right] \\
&\leq 4|\Theta_P| \sum_{t=1}^{t_K} \exp\left(-C \log^{1+\varepsilon}(T - t)\right) + \sum_{k=1}^{K} \frac{\kappa_u m^4}{3n^2 \underline{M}^3} \log^{1+\varepsilon}(T - t_{k-1}) \\
&\leq O(1) + \frac{\kappa_u m^4}{3n^2 \underline{M}^3} K \log^{1+\varepsilon} T \\
&\leq O(1) + \frac{\kappa_u m^4}{3n^2 \underline{M}^3 \log(1/\overline{M})} \log^{2+\varepsilon} T,
\end{aligned}$$

where the equality follows from telescoping, the first inequality from the one-epoch welfare loss bound above, the second inequality from the series convergence result above, and the last inequality because $1 = T - t_K = f^{(K)}(T - t_0) < \overline{M}^K T$ implies the number of re-solving is bounded by $K \leq \log T / \log(1/\overline{M})$.

Therefore, we conclude $\mathcal{R}_T(\mathsf{BIRT}) \leq O(1) + \kappa_u O(\log^{2+\varepsilon} T)$.

A.1.19 Proof of Proposition 4

Denote an alias $\boldsymbol{x}_0$ for the initial thresholded policy $\hat{\mathsf{F}}$ that is adopted for resources $t = 1, \ldots, t_1$. To bound the expected welfare loss $\mathbb{E}[\mathsf{OPT} - \mathsf{OPT}^1]$, we first present the following almost sure bound on the welfare loss $\mathsf{OPT}-\mathsf{OPT}^1$, so that it remains to formalize the intuition that the welfare loss $\mathsf{OPT} - \mathsf{OPT}^1$ is small with a high probability, i.e., in most cases after making possibly suboptimal allocations by time t_1, the hindsight optimal policy can still make up for most of the errors. The proof of Lemma 17 is deferred to Section A.1.20.

Lemma 17. *Suppose* OPT^1 *is the hindsight optimal egalitarian welfare starting with utilities* U_{t_1} *at time* $t_1 \in [T]$*, where* $U_{t_1} \geq U_0$*. Then its egalitarian welfare loss against the hindsight optimum* OPT *starting with utilities* U_0 *at time* 0 *is at most*

$$\mathsf{OPT} - \mathsf{OPT}^1 \leq t_1.$$

While we aim at finding an appropriate event on which the welfare loss $\mathsf{OPT} - \mathsf{OPT}^1$ is small, it is nontrivial to characterize the later hindsight optimal policy OPT^1. Instead, we consider a surrogate policy to be adopted after t_1, which would necessarily lead to a final welfare not exceeding OPT^1 by definition. More specifically, we consider the possible existence of an offline policy after t_1 that yields the same per-type per-agent aggregate allocations as the initial hindsight optimal policy OPT, i.e.,

$$N_\theta^{\leq}(x_0)_\theta + N_\theta^{>}\hat{x}_\theta = N_\theta(x_*)_\theta, \forall\theta \text{ for some } \hat{\boldsymbol{x}} \in \Delta_n^{\Theta_P} \text{ and hindsight optimal } \boldsymbol{x}_*, \tag{A.1.8}$$

where we denote for each type θ the number of type-θ arrivals by t_1 by $N_\theta^{\leq} := \sum_{t=1}^{t_1} \mathbf{1}\{\theta_t = \theta\}$ and that of those after t_1 by $N_\theta^{>} := \sum_{t=t_1+1}^{T} \mathbf{1}\{\theta_t = \theta\}$.

Under condition (A.1.8), it is still feasible after t_1 to match the per-type per-agent aggregate allocations as the initial hindsight optimal policy. When all utilities are linear, this implies the per-type per-agent aggregate utilities can be matched too, leading to a zero welfare loss

$\mathsf{OPT} - \mathsf{OPT}^1 = 0$. However, if utilities are nonlinear and concave, although per-type per-agent aggregate allocations may match, corresponding utilities might not, resulting in a positive welfare loss. Fortunately, even in the latter case, we are able to bound the welfare loss by constraining our attention to a high-probability event on which Equation (A.1.8) holds.

Now that we have a possible candidate policy, the remaining parts of the proof can be outlined as the following three steps: 1) providing a sufficient condition on the arrivals under which Equation (A.1.8) holds, 2) bounding the welfare loss under the sufficient condition by leveraging the shape of the admissible utility functions (Definition 1), and 3) proving the sufficient condition holds with high probability, so that the welfare loss is small when it does not hold.

In developing the sufficient condition, Lemma 10 will be crucial in developing the sufficient condition, which states a Lipschitz continuity property of the optimal solution multi-function against arrivals. This assumption guarantees there exists a hindsight optimal policy that is close to the fluid policy. Since the BIRT policy is constructed based on the fluid policy, this further implies the BIRT policy is close to the hindsight optimal policy, i.e., stochastic fluctuations are small compared to the leeway that the prophet OPT^1 has after time t_1 to make up for previous suboptimal allocations. Hence, we will be able to show that OPT^1 can recover the exact same per-type per-agent aggregate allocations for all agents as OPT does with high probability.

Using the Lipschitz continuity property (Lemma 10), we provide a sufficient condition for Equation (A.1.8) to hold with the thresholded policy $\boldsymbol{x}_0$. Recall the deviations of random variables from their expected values are denoted with the Δ operator (e.g., $\Delta N_\theta := N_\theta - \mathbb{E}[N_\theta]$). The proof can be found in Section A.1.21.

Lemma 18. *Suppose the online policy $\hat{\mathbf{F}}$ is constructed by thresholding a fluid policy $\mathbf{F}$ with $\gamma \in [0, 2(T - t_1)/3n^2T]$. Then Equation* (A.1.8) *holds on event*

$$E := \left\{ \max\left(\left\|\Delta N^{\leq}\right\|_\infty, \left\|\Delta N^{>}\right\|_\infty \right) \leq \underline{p}\frac{T - t_1}{3}\frac{\gamma}{L} \right\}, \tag{A.1.9}$$

where we recall $\underline{p} = \min_\theta p_\theta$.

Lemma 18 provides a sufficient condition for Equation (A.1.8) on the deviations of arrival counts from their means. More specifically, the condition states that deviations are uniformly bounded by a quantity dependent on the minimum probability of all arrival types.

Now with the help of Lemma 18, we can examine the performance of $\hat{\boldsymbol{x}}$ and obtain the following bound on the welfare loss on event E. The proof can be found in Section A.1.22.

Lemma 19. *Suppose the online policy* $\hat{\mathsf{F}}$ *is constructed by thresholding a fluid policy* F *with* $\gamma \in [0, 2(T-t_1)/3n^2T]$. *Then on event* E,

$$\mathsf{OPT} - \mathsf{OPT}^1 \le \frac{3\kappa_u n^2}{4}\frac{\gamma^2 T^2}{T-t_1}.$$

The last piece is the welfare loss outside of event E, which Lemma 17 implies is bounded by

$$\mathbb{E}\left[\left(\mathsf{OPT} - \mathsf{OPT}^1\right)\mathbf{1}\left(E^{\complement}\right)|U_0\right] \le t_1 \mathbb{P}\left(E^{\complement}|U_0\right).$$

We can bound the probability of event E as follows.

$$\begin{aligned}
\mathbb{P}\left(E^{\complement}|U_0\right) &\le \sum_{\theta\in\Theta}\mathbb{P}\left\{\left|\Delta N_\theta^{\le}\right| > \underline{p}\frac{T-t_1}{3}\frac{\gamma}{L}\right\} + \sum_{\theta\in\Theta}\mathbb{P}\left\{\left|\Delta N_\theta^{>}\right| > \underline{p}\frac{T-t_1}{3}\frac{\gamma}{L}\right\} \\
&\le 2|\Theta|\exp\left(-\frac{2\underline{p}^2(T-t_1)^2\gamma^2}{9L^2t_1}\right) + 2|\Theta|\exp\left(-\frac{2\underline{p}^2(T-t_1)^2\gamma^2}{9L^2(T-t_1)}\right) \qquad \text{(A.1.10)} \\
&\le 4|\Theta|\exp\left(-\frac{2\underline{p}^2(T-t_1)^2\gamma^2}{9L^2T}\right),
\end{aligned}$$

where the first inequality follows from the union bound, the second from Hoeffding's inequality as $N_\theta^{\le} \sim \text{Bin}(t_1, p_\theta)$ and $N_\theta^{>} \sim \text{Bin}(T-t_1, p_\theta)$ for every type θ, and the last because $1 \le t_1 \le T-1$.

Therefore, we conclude

$$\begin{aligned}
\mathbb{E}\left[\mathsf{OPT} - \mathsf{OPT}^1|U_0\right] &= \mathbb{E}\left[(\mathsf{OPT} - \mathsf{OPT}^1)\mathbf{1}(E^{\complement})|U_0\right] + \mathbb{E}\left[(\mathsf{OPT} - \mathsf{OPT}^1)\mathbf{1}(E)|U_0\right] \\
&\le 4|\Theta|t_1\exp\left(-\frac{2\underline{p}^2(T-t_1)^2\gamma^2}{9L^2T}\right) + \frac{3\kappa_u n^2}{4}\frac{\gamma^2T^2}{T-t_1},
\end{aligned}$$

where the first term is bounded by Lemma 17 and Equation (A.1.10), and the second term by Lemma 19 and $\mathsf{OPT} - \mathsf{OPT}^1 \geq 0$ a.s.

A.1.20 Proof of Lemma 17

Suppose a hindsight optimal policy allocates $\boldsymbol{x} \in \Delta_n^T$. Then

$$
\begin{aligned}
\mathsf{OPT} - \mathsf{OPT}^1 &= \min_{j \in [n]} \left(U_0^j + \sum_{t=1}^{T} u_{\theta_t}^j(x_t^j) \right) - \max_{\hat{\boldsymbol{x}} \in \Delta_n^T} \min_i \left(U_{t_1}^i + \sum_{t=t_1+1}^{T} u_{\theta_t}^i(\hat{x}_t^i) \right) \\
&\leq \min_{j \in [n]} \left(U_0^j + \sum_{t=1}^{T} u_{\theta_t}^j(x_t^j) \right) - \min_{i \in [n]} \left(U_{t_1}^i + \sum_{t=t_1+1}^{T} u_{\theta_t}^i(x_t^i) \right) \\
&\leq \max_{i \in [n]} \left(U_0^i - B_{t_1}^i + \sum_{t=1}^{t_1} u_{\theta_t}^i(x_t^i) \right) \\
&\leq \max_{i \in [n]} \sum_{t=1}^{t_1} u_{\theta_t}^i(x_t^i) \\
&\leq t_1,
\end{aligned}
$$

where the equality follows by definition of the hindsight optimal allocation $\boldsymbol{x}$, the first inequality because $\boldsymbol{x} \in \Delta_n^T$, the second inequality by taking $j = i$, the third inequality because $U_0 \leq B_{t_1}$, and the last inequality because $u_\theta^i(\cdot) \leq 1$ for any θ and i.

A.1.21 Proof of Lemma 18

We first present the following lemma that gives a sufficient condition on the existence of $\boldsymbol{x}_*$ in Equation (A.1.8). The proof is deferred to Section A.1.23.

Lemma 20. *Suppose the central planner acts according to a type-based policy $\boldsymbol{x}_0 \in \Delta_n^{\Theta_P}$ for $t = 1, 2, \ldots, t_1$. Then* (A.1.8) *holds if there is a hindsight optimal static solution $\boldsymbol{x}_* \in \Delta_n^{\Theta_P}$ such that*

$$
N_\theta(x_*)_\theta^i \geq N_\theta^{\leq}(x_0)_\theta^i, \quad \forall \theta, i. \tag{A.1.11}
$$

Lemma 20 states Equation (A.1.8) holds if a hindsight optimal policy $\boldsymbol{x}_*$ allocates more of type-θ resource to agent i than the online policy $\boldsymbol{x}_0$ does by t_1, for every type θ and every agent i. In other words, $\boldsymbol{x}_0$ leaves room for future allocation by some hindsight optimal policy $\boldsymbol{x}_*$. Hence, it suffices to show $N_\theta(x_*)^i_\theta \geq N^{\leq}_\theta(x_0)^i_\theta$ for all i and θ.

To compare allocations by the two policies, we first compare the hindsight optimal policy $\boldsymbol{x}_*$ and the fluid policy F. Notice by Lemma 10, there exists a hindsight optimal type-based policy $\boldsymbol{x}_* \in \Delta_n^{|\Theta|}$ such that $\max_{i,\theta} \left|N_\theta(x_*)^i_\theta - Tp_\theta \mathsf{F}^i_\theta\right| \leq L \|\Delta N\|_\infty$ for a constant $L \geq 2$. Hence for all i and θ,

$$\begin{aligned} N_\theta(x_*)^i_\theta - N^{\leq}_\theta \mathsf{F}^i_\theta &\geq \left[Tp_\theta \mathsf{F}^i_\theta - L\|\Delta N\|_\infty\right] - \left[t_1 p_\theta + \Delta N^{\leq}_\theta\right] \mathsf{F}^i_\theta \\ &= \left[(T-t_1)p_\theta - \Delta N^{\leq}_\theta\right] \mathsf{F}^i_\theta - L\|\Delta N\|_\infty \\ &\geq \left[(T-t_1)p_\theta - \underline{p}\frac{T-t_1}{3}\frac{\gamma}{L}\right] \mathsf{F}^i_\theta - \frac{2}{3}(T-t_1)\underline{p}\gamma, \\ &\geq \frac{2}{3}(T-t_1)p_\theta \left[\mathsf{F}^i_\theta - \gamma\right], \end{aligned} \tag{A.1.12}$$

where the second inequality follows from the assumption (A.1.9) and $\|\Delta N\|_\infty \leq \left\|\Delta N^{\leq}\right\|_\infty + \|\Delta N^{>}\|_\infty \leq 2(T-t_1)\underline{p}\gamma/3L$ by the triangle inequality, and the last inequality because $\underline{p} \leq p_\theta$, $\gamma \leq 1$ and $L \geq 1$.

Now we check condition (A.1.11) element-wise. Recall from the thresholding rule that $\hat{\mathsf{F}}^i_\theta \in \{0\} \cup [\gamma, 1]$ for all i and θ.

- If $\hat{\mathsf{F}}^i_\theta = 0$, then $N_\theta(x_*)^i_\theta \geq 0 = N^{\leq}_\theta \hat{\mathsf{F}}^i_\theta$.
- If $\hat{\mathsf{F}}^i_\theta \geq \gamma$ and $i \neq j$, then agent i is not adjusted, so $\mathsf{F}^i_\theta = \hat{\mathsf{F}}^i_\theta \geq \gamma$, and Equation (A.1.12) implies
$$N_\theta(x_*)^i_\theta - N^{\leq}_\theta \hat{\mathsf{F}}^i_\theta = N_\theta(x_*)^i_\theta - N^{\leq}_\theta \mathsf{F}^i_\theta \geq \frac{2}{3}(T-t_1)p_\theta\left[\mathsf{F}^i_\theta - \gamma\right] \geq 0.$$
- Agent $j \in \arg\max_i \mathsf{F}^i_\theta$ is the only one possibly receiving positive adjustment, and the adjust-

ment is at most $(\hat{\mathsf{F}} - \mathsf{F})^j_\theta = \sum_{i \neq j} (\mathsf{F} - \hat{\mathsf{F}})^i_\theta \leq (n-1)\gamma$. For agent j,

$$\begin{aligned} N_\theta(x_*)^j_\theta - N^{\leq}_\theta \hat{\mathsf{F}}^j_\theta &= N_\theta(x_*)^j_\theta - N^{\leq}_\theta \mathsf{F}^j_\theta - N^{\leq}_\theta (\hat{\mathsf{F}} - \mathsf{F})^j_\theta \\ &\geq \frac{2}{3}(T - t_1) p_\theta \left[\mathsf{F}^i_\theta - \gamma\right] - (n-1)\gamma T p_\theta \\ &\geq \frac{2}{3}(T - t_1) p_\theta \cdot \frac{1}{n} - n\gamma T p_\theta \\ &\geq 0. \end{aligned}$$

where the first inequality by Equation (A.1.12) and $N^{\leq}_\theta = t_1 p_\theta + \Delta N^{\leq}_\theta \leq t_1 p_\theta + (T - t_1)\underline{p}\gamma/3L \leq T p_\theta$ by the assumption (A.1.9), the second inequality because $\mathsf{F}^j_\theta = \max_i \mathsf{F}^i_\theta \geq 1/n$, and the last inequality by the threshold rule $0 \leq \gamma \leq 2(T - t_1)/3n^2 T$.

Therefore, we conclude $N_\theta(x_*)^i_\theta \geq N^{\leq}_\theta \hat{\mathsf{F}}^i_\theta$ for all i and θ.

A.1.22 Proof of Lemma 19

On event E, since $N^{>}_\theta \geq (T - t_1)p_\theta - (T - t_1)\underline{p}\gamma/3L > 0$, Lemma 18 implies we can define the type-based policy $\hat{\boldsymbol{x}} \in \Delta_n^{|\Theta|}$ given by

$$\hat{x}_\theta = \frac{N_\theta(x_*)_\theta - N^{\leq}_\theta (x_0)_\theta}{N^{>}_\theta}. \tag{A.1.13}$$

To bound the welfare loss $\mathsf{OPT} - \mathsf{OPT}^1$, we consider the differences in each agent's aggregate utilities by T separately. We denote by $U^i_T(\boldsymbol{x}_0, \hat{\boldsymbol{x}})$ agent i's aggregate utility by T if the central planner acts according to $\boldsymbol{x}_0$ up to t_1 and $\hat{\boldsymbol{x}}$ afterwards; similarly, we denote by $U^i_T(\boldsymbol{x}_*)$ agent i's aggregate utility by T under the hindsight optimal policy $\boldsymbol{x}_*$. Notice for any agent i,

$$\begin{aligned}
U_T^i(\boldsymbol{x}_0, \hat{\boldsymbol{x}}) - U_T^i(\boldsymbol{x}_*) &= \left[U_0^i + \sum_\theta N_\theta^{\leq} u_\theta^i((x_0)_\theta^i) + \sum_\theta N_\theta^{>} u_\theta^i(\hat{x}_\theta^i)\right] - \left[U_0^i + \sum_\theta N_\theta u_\theta^i((x_*)_\theta^i)\right] \\
&= \sum_\theta N_\theta^{\leq} \left[u_\theta^i((x_0)_\theta^i) - u_\theta^i((x_*)_\theta^i)\right] + \sum_\theta N_\theta^{>} \left[u_\theta^i(\hat{x}_\theta^i) - u_\theta^i((x_*)_\theta^i)\right] \\
&\geq \sum_\theta N_\theta^{\leq} \left[(u_\theta^i)'((x_*)_\theta^i)(x_0 - x_*)_\theta^i - \frac{\kappa_u}{2}\left|(x_0 - x_*)_\theta^i\right|^2\right] \\
&\quad + \sum_\theta N_\theta^{>} \left[(u_\theta^i)'((x_*)_\theta^i)(\hat{x} - x_*)_\theta^i - \frac{\kappa_u}{2}\left|(\hat{x} - x_*)_\theta^i\right|^2\right] \\
&= -\frac{\kappa_u}{2}\sum_\theta \left[N_\theta^{\leq}|(x_0 - x_*)_\theta^i|^2 + N_\theta^{>}|(\hat{x} - x_*)_\theta^i|^2\right] \\
&= -\frac{\kappa_u}{2}\sum_\theta \frac{N_\theta^{\leq} N_\theta}{N_\theta^{>}}|(x_0 - x_*)_\theta^i|^2,
\end{aligned}$$

where the first two equalities follow because utilities are additive across time, the inequality because admissible utility functions are strongly smooth by Definition 1, and the last two equalities by definition of $\hat{\boldsymbol{x}}$. Then we can bound the welfare loss on event E based on the policies of $\boldsymbol{x}_0$ and $\boldsymbol{x}_*$.

$$\begin{aligned}
\mathsf{OPT} - \mathsf{OPT}^1 &\leq \min_j U_T^j(\boldsymbol{x}_*) - \min_i U_T^i(\boldsymbol{x}_0, \hat{\boldsymbol{x}}) \\
&= \min_j U_T^j(\boldsymbol{x}_*) + \max_i \left[-U_T^i(\boldsymbol{x}_0, \hat{\boldsymbol{x}})\right] \\
&\leq \max_i \left[U_T^i(\boldsymbol{x}_*) - U_T^i(\boldsymbol{x}_0, \hat{\boldsymbol{x}})\right] \\
&\leq \frac{\kappa_u}{2}\sum_\theta \frac{N_\theta^{\leq} N_\theta}{N_\theta^{>}} \max_i |(x_0 - x_*)_\theta^i|^2,
\end{aligned}$$

where the first inequality follows because $\hat{\boldsymbol{x}} \in \Delta_n^{|\Theta|}$ on event E by Lemma 18, the second by taking $j = i$, and the last by Jensen's inequality. Now it suffices to bound $N_\theta^{\leq} N_\theta / N_\theta^{>}$ and $|(x_0 - x_*)_\theta^i|^2$.

For the former, notice on event E,

$$\sum_\theta \frac{N_\theta^{\leq} N_\theta}{N_\theta^{>}} \leq \sum_\theta \frac{t_1 p_\theta + (T - t_1)\underline{p}/3}{(T - t_1)p_\theta - (T - t_1)\underline{p}/3} N_\theta \leq \frac{3T^2}{2(T - t_1)},$$

where the first inequality follows from Equation (A.1.9), and the second because $\underline{p} \leq p_\theta$ and $\sum_\theta N_\theta = T$.

For the latter, recall $\boldsymbol{x}_0 \equiv \hat{\mathsf{F}}$. We claim $|(\hat{\mathsf{F}} - x_*)^i_\theta| \leq |(\hat{\mathsf{F}} - \mathsf{F})^i_\theta| + |(\mathsf{F} - x_*)^i_\theta| \leq n\gamma$ because the thresholding adjustment is bounded by $|(\hat{\mathsf{F}} - \mathsf{F})^i_\theta| \leq (n-1)\gamma$, and

$$
\begin{aligned}
|(\mathsf{F} - x_*)^i_\theta| &= \frac{|\mathbb{E}[N_\theta]\mathsf{F}^i_\theta - N_\theta(x_*)^i_\theta + (N_\theta - \mathbb{E}[N_\theta])(x_*)^i_\theta|}{\mathbb{E}[N_\theta]} \\
&\leq \frac{|\mathbb{E}[N_\theta]\mathsf{F}^i_\theta - N_\theta(x_*)^i_\theta| + |N_\theta - \mathbb{E}[N_\theta]|}{\mathbb{E}[N_\theta]} \\
&\leq \frac{L\,\|\Delta N\|_\infty + \|\Delta N\|_\infty}{\mathbb{E}[N_\theta]} \\
&\leq \frac{2(L+1)(T-t_1)\gamma}{3LT}\frac{\underline{p}}{p_\theta} \\
&\leq \gamma,
\end{aligned}
$$

where the first follows by the triangle inequality and $\boldsymbol{x}_* \in \Delta_n^\Theta$, the second by the Lipschitz continuity condition (Lemma 10), the third because $\|\Delta N\|_\infty \leq \left\|\Delta N^{\leq}\right\|_\infty + \|\Delta N^{>}\|_\infty \leq 2(T-t_1)\underline{p}\gamma/3L$ by the triangle inequality on event E (A.1.9), and the last because $L \geq 2$ and $\underline{p} \leq p_\theta$.

Therefore, we conclude on event E that

$$
\mathsf{OPT} - \mathsf{OPT}^1 \leq \frac{3\kappa_u n^2}{4}\frac{\gamma^2 T^2}{T - t_1}.
$$

A.1.23 Proof of Lemma 20

We aim to show there exists $\boldsymbol{x} \in \Delta_n^{\Theta_P}$ such that $N_\theta(x_*)^i_\theta = N^{\leq}_\theta(x_0)^i_\theta + N^{>}_\theta x^i_\theta$.

Fix arrival type θ. If $N^{>}_\theta = 0$, then we claim Equation (A.1.11) holds with equality. Otherwise, $N_\theta = \sum_i N_\theta(x_*)^i_\theta > \sum_i N^{\leq}_\theta(x_0)^i_\theta = N^{\leq}_\theta = N_\theta$, which is absurd. Now we have shown $N_\theta(x_*)^i_\theta = N^{\leq}_\theta(x_0)^i_\theta$ for all i, i.e., all agents have accumulated the exact welfare needed from type-θ arrivals by time t_1. An arbitrary allocation for type θ would satisfy the condition, since there will be no such arrivals.

If $N_\theta^> > 0$, the equality is satisfied by simply taking the type-based policy $\boldsymbol{x} \in \Delta_n^{\Theta_P}$ given by

$$x_\theta = \frac{N_\theta (x_*)_\theta - N_\theta^\leq (x_0)_\theta}{N_\theta^>}, \quad \forall \theta \in \Theta.$$

A.2 KKT Optimality Conditions

We present in this section the classical Karush-Kukn-Tucker optimality conditions of the fluid problem (1.5) based on the Lagrangian function (A.1.1). Since the fluid problem is convex by concavity of admissible utility functions, and strong duality holds by Slater's condition (Lemma 5), the KKT conditions are sufficient and necessary for optimality of the primal-dual pair $(\boldsymbol{x}_\rho; \nu_\rho, \lambda_\rho, \boldsymbol{\mu}_\rho)$ (see Section 3.3.4 in [84]). The first two conditions are first-order stationarity conditions for $\boldsymbol{x}$ and Υ, and the last three are complementary slackness conditions on ν, λ and $\boldsymbol{\mu}$.

$$\begin{cases} \nu_\rho^i p_\theta (u_\theta^i)'((x_\rho)_\theta^i) - (\lambda_\rho)_\theta + (\mu_\rho)_\theta^i = 0, & \text{for all } i \in [n], \theta \in \Theta_P, \\ \sum_{i=1}^n \nu_\rho^i = 1, & \\ \nu_\rho^i \left(\rho^i + \mathbb{E}_\theta[u_\theta^i((x_\rho)_\theta^i)] - J(\rho)\right) = 0, & \text{for all } i \in [n], \\ (\lambda_\rho)_\theta \left(\sum_i (x_\rho)_\theta^i - 1\right) = 0, & \text{for all } \theta \in \Theta_P, \\ (\mu_\rho)_\theta^i (x_\rho)_\theta^i = 0, & \text{for all } i \in [n], \theta \in \Theta_P. \end{cases} \tag{A.2.1}$$

The complementary slackness condition on $\boldsymbol{\mu}_\rho$ implies $(\mu_\rho)_\theta^i = 0$ if $(x_\rho)_\theta^i > 0$, which is true for some i for almost every type θ. Hence, the first-order stationarity condition implies $(\lambda_\rho)_\theta = \max_i \nu_\rho^i p_\theta (u_\theta^i)'((x_\rho)_\theta^i)$. Then clearly $(\mu_\rho)_\theta^i = (\lambda_\rho)_\theta - \nu_\rho^i p_\theta (u_\theta^i)'((x_\rho)_\theta^i)$. We thus obtain the following equivalent conditions based on the KKT conditions (A.2.1).

$$
\begin{cases}
(x_\rho)^i_\theta = 0, & \text{if } i \notin \arg\max_j \nu^j_\rho p_\theta (u^j_\theta)'((x_\rho)^j_\theta), \\
\nu^i_\rho = 0, & \text{if } \rho^i + \mathbb{E}_\theta[u^i_\theta((x_\rho)^i_\theta)] - J(\rho) > 0, \\
\sum_{i=1}^n \nu^i_\rho = 1, & \\
\sum_{i=1}^n (x_\rho)^i_\theta = 1, & \text{for all } \theta \in \Theta_P, \\
(\lambda_\rho)_\theta = \max_j \nu^j_\rho p_\theta (u^j_\theta)'((x_\rho)^j_\theta), & \text{for all } \theta \in \Theta_P, \\
(\mu_\rho)^i_\theta = (\lambda_\rho)_\theta - \nu^i_\rho p_\theta (u^i_\theta)'((x_\rho)^i_\theta), & \text{for all } i \in [n], \theta \in \Theta_P.
\end{cases}
\tag{A.2.2}
$$

It is thus clear that λ_ρ and $\boldsymbol{\mu}_\rho$ can be uniquely determined by the $(\boldsymbol{x}_\rho, \nu_\rho)$ pair.

Lemma 21. *The fluid problem* (1.5) *has the exact same optimal solutions and optimum as if the second constraint is replace with the inequality* $\sum_{i\in[n]} x^i_\theta \leq 1$ *for any* $\theta \in \Theta_P$.

Proof We first show any optimal solution $\hat{\boldsymbol{x}}_\rho$ in the new formulation is an optimal solution to the original formulation. When the constraint is replaced with the inequality, the exact same KKT optimality conditions (A.2.1) and (A.2.2) hold. In particular, $(\hat{\lambda}_\rho)_\theta = \max_j \hat{\nu}^j_\rho p_\theta (u^j_\theta)'((\hat{x}_\rho)^j_\theta) > 0$ because $p_\theta > 0$ for some $\theta \in \Theta_P$, $(u^j_\theta)'(\cdot) > 0$ by Definition 1, and $\hat{\nu}^j_\rho > 0$ for some $j \in [n]$ as $\hat{\nu}_\rho \in \Delta_n$. Hence, complementary slackness (A.2.1) implies $\sum_{i\in[n]} (\hat{x}_\rho)^i_\theta = 1$, so $\hat{x}_\rho$ is feasible in the original formulation. Since the new formulation is relaxed, its optimal solution must also be optimal in the original formulation.

On the contrary, any optimal solution $\boldsymbol{x}_\rho$ in the original formulation is clearly feasible in the new formulation. Since we have shown any optimal solution to the new formulation is feasible in the original formulation, $\boldsymbol{x}_\rho$ must also be optimal in the new formulation. □

A.3 Additional results

Lemma 22. *Any vector of admissible utility functions* $u(x) \equiv (u^i(x^i) : i \in [n])$ *admits a global lower quadratic envelope in* $[0,1]^n$ *with curvature* κ_u*, i.e.,*

$$u(\hat{x}) \geq u(x) + \nabla u(x)(\hat{x} - x) - \frac{\kappa_u \mathbf{1}}{2} \|\hat{x} - x\|_2^2, \quad \forall \hat{x}, x \in \Delta_n.$$

Proof Since $\nabla u(x) = \operatorname{diag}((u^i)'(x^i) : i \in [n])$, it suffices to show $u^i(\cdot)$ admits a global lower quadratic envelope in $[0,1]$ with curvature κ_u, i.e., $u^i(\hat{x}) \geq u^i(x) + (u^i)'(x)(\hat{x} - x) - \frac{\kappa_u}{2}(\hat{x} - x)^2$ for any $\hat{x}, x \in [0,1]$. This is immediate from the multivariate mean value theorem because $(u^i)''(\cdot) \geq -\kappa_u$ on $[0,1]$. The proof is thereby concluded. □

Lemma 23. *If for some resource* t*, the utility functions* $u_t^i(\cdot)$ *to all agents are admissible, then the hindsight optimal utilities are all positive; as a result, the hindsight optimum is positive, i.e.,* $\mathsf{OPT} > 0$.

Proof Consider the policy $x_t = \mathbf{1}/n > 0$ that equally divides resource t across the agents. By definition of admissible utility functions, $u_t(x_t) > 0$, so $U_T = \sum_{t'} u_{t'}(x_{t'}) > 0$. As a result, $\mathsf{OPT} \geq w(U_T) > 0$. □

Lemma 24 (Futility of randomization)**.** *Consider the online convex optimization problem*

$$\begin{aligned} \text{maximize} \quad & w\left(\sum_{t=1}^{T} u_t(x_t)\right) \\ \text{subject to} \quad & x_t \in \mathcal{X}_t, \\ & x_t \perp \sigma(u_{t+1}, \ldots, u_T) \end{aligned}$$

where $w : \mathbb{R}^m \to \mathbb{R}$ *is a concave increasing function,* $u_t : \mathcal{X}_t \to \mathbb{R}^m$ *are fixed concave utility* *functions, and $\mathcal{X}_t$ is a convex set for every t. Then in expectation, any randomized online policy is weakly dominated by a deterministic online policy.*

Proof Fix any randomized online policy π, and suppose it outputs random actions $(x_t^\pi : t \in [T])$. We argue the deterministic mean policy $\bar{\pi}$ given by actions $x_t^{\bar{\pi}} := \mathbb{E}_\pi \left[x_t^\pi\right]$ is feasible and dominates the randomized policy π. Its feasibility follows because $x_t^{\bar{\pi}} \in \mathcal{X}_t$ by the convexity of the action domains $\mathcal{X}_t$, so it is an online policy. Notice

$$\mathbb{E}_\pi \left[w \left(\sum_{t=1}^{T} u_t(x_t^\pi) \right) \right] \leq w \left(\sum_{t=1}^{T} \mathbb{E}[u_t(x_t^\pi)] \right) \leq w \left(\sum_{t=1}^{T} u_t \left(\mathbb{E}_\pi [x_t^\pi] \right) \right),$$

where the first inequality follows by Jensen's inequality because $w(\cdot)$ is concave, and the second because $w(\cdot)$ is increasing and $u_t(\cdot)$ is concave for all t. The proof is concluded as the RHS is the objective obtained by the mean policy $\bar{\pi}$. □

Lemma 25 (Type-based policies dominate)**.** *If all resources are admissible, any offline policy is weakly dominated by a type-based policy.*

Proof Fix an optimal offline policy $\boldsymbol{x}$ and consider agent i's cumulative utility $U_T^i = \sum_{t \in [T]} u_{\theta_t}^i(x_t^i)$. We define an alternative type-based policy $\hat{\boldsymbol{x}}$ given by

$$\hat{x}_\theta^i = \frac{\sum_{t:\theta_t = \theta} x_t^i}{|\{t : \theta_t = \theta\}|} \quad \text{for all } \theta \in \{\theta_t : t \in [T]\}$$

and show it weakly dominates $\boldsymbol{x}$. First notice $x_\theta \in \Delta_n$ for all $\theta \in \{\theta_t : t \in [T]\}$. Next,

$$\begin{aligned}
\sum_{t\in[T]} u^i_{\theta_t}(x^i_t) &= \sum_{\theta\in\{\theta_t:t\in[T]\}} \sum_{t:\theta_t=\theta} u^i_\theta(x^i_t) \\
&\le \sum_{\theta\in\{\theta_t:t\in[T]\}} u^i_\theta\left(\frac{\sum_{t:\theta_t=\theta} x^i_t}{|\{t:\theta_t=\theta\}|}\right)|\{t:\theta_t=\theta\}| \\
&= \sum_{\theta\in\{\theta_t:t\in[T]\}} \sum_{t:\theta_t=\theta} u^i_\theta\left(\hat{x}^i_\theta\right) \\
&= \sum_{t\in[T]} u^i_{\theta_t}(x^i_{\theta_t}),
\end{aligned}$$

where the inequality follows by Jensen's inequality because $u^i_{\theta_t}$ is concave for any agent i and resource t. The proof is hereby concluded. □

Lemma 26. *Define the event $E_\delta := \{|U^i_T - \mathbb{E}[U^i_T]| \le \delta\mathbb{E}[U^i_T], \forall i\}$ for all $\delta > 0$. Then*

$$\mathbb{P}\left(E^\complement_\delta\right) \le \sum_{i\in[n]} 2\exp\left(-\frac{\delta^2\mathbb{E}[y^i_\theta]^2}{2}T\right)\boldsymbol{1}\{\mathbb{E}[y^i_\theta] > 0\}.$$

Proof

$$\begin{aligned}
\mathbb{P}(E^\complement_\delta) &= \mathbb{P}\left\{\left|U^i_T - \mathbb{E}[U^i_T]\right| > \delta\mathbb{E}[U^i_T] \text{ for some } i\right\} \\
&\le \sum_{i\in[n]} \mathbb{P}\{\left|U^i_T - \mathbb{E}[U^i_T]\right| > \delta\mathbb{E}[U^i_T]\} \\
&= \sum_{i\in[n]} \mathbb{P}\left\{\left|\sum_{t\in[T]}(y^i_{\theta_t} - \mathbb{E}[y^i_{\theta_t}])\right| > \delta\mathbb{E}[y^i_\theta]T\right\} \\
&\le \sum_{i\in[n]} 2\exp\left(-\frac{\delta^2\mathbb{E}[y^i_\theta]^2}{2}T\right)\mathbf{1}\{\mathbb{E}[y^i_\theta] > 0\},
\end{aligned}$$

where the first equality follows by the definition of $Z^i_T = U^i_T/\mathbb{E}[U^i_T] - 1$, the first inequality by the union bound, the second equality by the definition of $U^i_T = \sum_{t\in[T]} y^i_{\theta_t}$ and the second by Hoeffding's inequality and the fact that $y^i_\theta = 0$ almost surely if $\mathbb{E}[y^i_\theta] = 0$. □

Lemma 27. *The class of infrequent re-solving policies is identical to the class of Backward Infrequent Re-solving (*BIR*) policies, i.e., for any $K \in \mathbb{N}$ and $(t_0 = 0, t_1, \ldots, t_K = T - 1) \in \mathbb{Z}_+^{K+1}$, the following two conditions are equivalent.*

- $t_0 < t_1 < \cdots < t_K$;
- *there is an admissible scheduling function f such that $T - t_k = f(T - t_{k-1})$ for any $k \in [K]$, and $t_k < T - 1$ for any $k \in [K-1]$.*

Proof We first assume the first condition. Then we claim the scheduling function f given by $f(1) = 1$ and $f(t) = T - N((T-t)_+ + 1)$ for $t \geq 2$ is admissible and satisfies the second condition. Notice $f(2) = T - N(T-1) = 1 = f(1)$; for $t \geq 2$, $f(t)$ is non-decreasing because $N(\cdot)$ and $(\cdot)_+$ are non-decreasing functions; for $t > T$, $f(t) = T - N(1) = T - t_1$ is well-defined. Hence, f maps $\mathbb{N}$ to $\mathbb{N}$ and is non-decreasing. For $t \geq 2$, $f(t) \leq T - N(T - t + 1) \leq T - (T - t + 1) = t - 1 < t$ because $N(\cdot) \geq \cdot$, so f is strictly contracting. Finally, f satisfies the second condition because $f(T - t_{k-1}) = T - N(t_{k-1} + 1) = T - t_k$ for $k \in [K]$, and $t_k < t_K = T - 1$ for any $k \in [K-1]$.

Next, we show the second condition implies the first. First notice that by definition $f(2) = 1$, so $f(1) = 1$, implying $f(t) \leq t$ for any $t \in \mathbb{N}$. Notice in particular $t_{K-1} < T - 1 = t_K$. As a result, $T - t_{K-1} \geq 2$; since $T - t_{K-1} = f(T - t_{K-2})$, this means $T - t_{K-2} \geq T_{K-1} \geq 2$. We can argue by induction that $T - t_{k-1} \geq 2$ for any $k \in [K]$. This means $T - t_k = f(T - t_{k-1}) < T - t_{k-1}$ for any $k \in [K]$ because f is strictly contracting except at 1. The proof is thereby concluded. □

A.4 Additional proofs

Lemma 28. *Fix a finite set of types Θ_P (namely $|\Theta_P| < \infty$) and a probability vector $p \in \Delta_{|\Theta_P|}$ with $p > 0$. If $N \sim \text{Multinomial}(T, p)$ for every $T \geq 1$, then for any $g \in \mathbb{R}^{\Theta_P}$,*

$$\frac{\mathbb{E}|g^\top (N - \mathbb{E}[N])|}{\sqrt{T}} \to \frac{g^\top \Sigma g}{\sqrt{2\pi}} \quad as \quad T \to \infty,$$

where $\Sigma := \mathrm{diag}(p) - pp^\top = \mathrm{Cov}(N)/T$.

Proof Denote the empirical probability vector by $\hat{p} := N/T$ for every $T \geq 1$ so that $\mathbb{E}[\hat{p}] = p$ and the covariance matrix of $\hat{p}$ is $\mathrm{Cov}(\hat{p}) = \Sigma/T$. Define the random process $Z_T := \sqrt{T} g^\top (\hat{p} - p)$ for $T \geq 1$. Then it suffices to show that $\mathbb{E}|Z_T| \to g^\top \Sigma g \sqrt{2/\pi}$ as $T \to \infty$.

Observe that $\mathbb{E}[Z_T] = 0$ and $\mathbb{E}[Z_T^2] = Tg^\top \mathrm{Cov}(\hat{p}) g = g^\top \Sigma g$ for all T, so $(Z_T : T \geq 1)$ is uniformly integrable, and that the Central Limit Theorem implies $Z_T \Rightarrow \mathsf{N}(0, g^\top \Sigma g)$. By Skorokhod's Representation Theorem, there exists a sequence of random variables $(Z_T' : T \geq 1)$ such that $Z_T \stackrel{d}{=} Z_T'$ and $Z_T' \to Z$ almost surely, where $Z \sim \mathsf{N}(0, g^\top \Sigma g)$. This means the class $(Z_T' : T \geq 1)$ is also uniformly integrable, so $Z_T' \to Z$ in L^1, implying in particular $\mathbb{E}|Z_T'| \to \mathbb{E}|Z| = g^\top \Sigma g \sqrt{2/\pi}$. Finally, we conclude the proof by noticing $\mathbb{E}|Z_T| = \mathbb{E}|Z_T'|$. □

Lemma 29. *Suppose* $N_T \sim \mathsf{Bin}(T, 1/2)$ *for* $T \geq 1$. *Then*

$$\frac{\mathbb{E}|N_T - T/2|}{\sqrt{T}} \to \frac{1}{\sqrt{2\pi}} \quad \text{as} \quad T \to \infty.$$

Proof Define the sequence $M_T := (2N_T - T)/\sqrt{T}$ for $T \geq 1$. Then it suffices to show that $\mathbb{E}|M_T| \to \sqrt{2/\pi}$ as $T \to \infty$.

Observe that $\mathbb{E}[M_T] = 0$ and $\mathbb{E}[M_T^2] = 4\mathrm{Var}(N_T)/T = 1$ for all T, so $\{M_T\}$ is uniformly integrable, and that the Central Limit Theorem implies $M_T \Rightarrow \mathsf{N}(0, 1)$. By Skorokhod's Representation Theorem, there exists a sequence of random variables $(Z_T : T \geq 1)$ such that $M_T \stackrel{d}{=} Z_T$ and $Z_T \to Z$ almost surely, where $Z \sim \mathsf{N}(0, 1)$. This means the class $\{Z_T : T \geq 1\}$ is also uniformly integrable, so $Z_T \to Z$ in L^1, implying in particular $\mathbb{E}|Z_T| \to \mathbb{E}|Z| = \sqrt{2/\pi}$. Hence, we conclude that $\mathbb{E}|M_T| = \mathbb{E}|Z_T| \to \sqrt{2/\pi}$. □

Lemma 30. *For any positive constants* $C > 0$ *and* $\varepsilon > 0$, $\sum_{t=1}^{\infty} \exp(-C \log^{1+\varepsilon} t) < \infty$.

Proof For $t > \exp((2/C)^{1/\varepsilon})$, $\exp(-C \log^{1+\varepsilon} t) \le t^{-2}$, so

$$\sum_{t=\lceil \exp(2/C)^{1/\varepsilon} \rceil}^{\infty} \exp(-C \log^{1+\varepsilon} t) \le \sum_{t=\lceil \exp(2/C)^{1/\varepsilon} \rceil}^{\infty} \frac{1}{t^2} \le \sum_{t=1}^{\infty} \frac{1}{t^2} = \frac{\pi^2}{6}.$$

□

Lemma 31 (Generalized Farkas' Lemma)**.** *Suppose $\mathcal{X} \subseteq \mathbb{R}^n$ is a closed and convex set, $g : \mathbb{R}^n \to \mathbb{R}^m$ is an affine function, and $b \in \mathbb{R}^m$. If there is no $x \in \mathcal{X}$ for which $g(x) \le b$, then there are constants $\mu \in \mathbb{R}_+^m$ and $c \in \mathbb{R}$ such that*

$$\begin{cases} \mu^\top g(x) \ge c, & \forall x \in \mathcal{X}, \\ \mu^\top b < c. \end{cases}$$

Proof By the assumption, $b \notin \{g(x) + y : y \in \mathbb{R}_+^m, x \in \mathcal{X}\}$, where the latter is a closed and convex set because $\mathcal{X}$ is closed and convex and g is affine. By the strict separating theorem, there are constants $\mu \in \mathbb{R}^m$ and $c \in \mathbb{R}$ such that $\mu^\top b < c$ and $\mu^\top (g(x) + y) \ge c$ for any $x \in \mathcal{X}$ and $y \in \mathbb{R}_+^m$. Since y can be arbitrarily large, we know $\mu \in \mathbb{R}_+^m$; taking $y = 0$, we conclude $\mu^\top g(x) \ge c$ for any $x \in \mathcal{X}$. □

Appendix B: Evaluating Model Performance Under Worst-case Sub-populations

B.1 Proof of finite-sample concentration results

Our results are based on a general concentration guarantee for estimating the dual reformulation (2.2.2) for any given $\mu(Z)$. We give this result in Appendix B.1.1, and build on it in subsequent proofs of key results. In the following, we use $\lesssim$ to denote inequality up to a numerical constant that may change line by line.

B.1.1 Concentration bounds for worst-case subpopulation performance

Since $\ell(\widehat{y}; y) \geq 0$ for losses used in most machine learning problems, we assume that $\mathcal{H}$ consists of nonnegative functions. To show exponential concentration guarantees, we consider sub-Gaussian conditional risk models $\mu(Z)$. Note the concentration results here are more general than needed for the purpose of proving the main results, because any random variable bounded in $[0, B]$ is inherently sub-Gaussian with paramater $B^2/4$.

Definition 6. *A function $h : \mathcal{Z} \to \mathbb{R}$ with $\mathbb{E}|h(Z)| < \infty$ is sub-Gaussian with parameter σ^2 if*

$$\mathbb{E}\left[\exp\left(\lambda(\mu(Z) - \mathbb{E}[\mu(Z)])\right)\right] \leq \exp\left(\frac{\sigma^2\lambda^2}{2}\right) \text{ for all } \lambda \in \mathbb{R}.$$

The sub-Gaussian assumption can be relaxed to sub-exponential random variables, with minor and standard modifications to subsequent results. We omit these results for brevity.

Define a dual plug-in estimator for the worst-case subpopulation performance of $h(Z)$ on S_2

$$\widehat{\mathsf{W}}_{\alpha,k}(h) = \inf_{\eta} \left\{ \frac{1}{\alpha|S_2|} \sum_{i \in S_2} (\widehat{\mu}_1(Z_i) - \eta)_+ + \eta \right\}. \tag{B.1.1}$$

The following result shows that for any sub-Gaussian μ that is bounded from below, the plug-in estimator (3) converges at the rate $O_p(|S_2|^{-1/2})$.

Proposition 5. *There is a universal constant $C > 0$ such that for all $h \geq 0$ that is sub-Gaussian with parameter σ^2,*

$$|\widehat{\mathsf{W}}_{\alpha,k}(\mu) - \mathsf{W}_\alpha(\mu)| \leq \frac{C\sigma}{\alpha}\sqrt{\frac{\log(2/\delta)}{|S_2|}} \text{ with probability at least } 1-\delta.$$

We prove the proposition in the rest of the subsection. By a judicious application of the empirical process theory, our bounds—which apply to nonnegative random variables—are simpler than existing concentration guarantees for conditional value-at-risk [181, 182].

Our starting point is the following claim, which bounds $|\widehat{\mathsf{W}}_{\alpha,k}(\mu) - \mathsf{W}_\alpha(\mu)|$ in terms of the suprema of empirical process on $\{z \mapsto (\mu(z) - \eta)_+ : \eta \geq 0\}$.

Claim 32.

$$\left|\widehat{\mathsf{W}}_{\alpha,k}(\mu) - \mathsf{W}_\alpha(\mu)\right| \leq \frac{1}{\alpha}\sup_{\eta \geq 0}\left|\frac{1}{|S_2|}\sum_{i \in S_2}(\mu(Z_i) - \eta)_+ - \mathbb{E}\left(\mu(Z) - \eta\right)_+\right| \tag{B.1.2}$$

The crux of this claim is that η does not range over $\mathbb{R}$, but rather has a lower bound; the value 0 can be replaced with any almost sure lower bound on $\mu(Z)$. Deferring the proof of Claim 32 to the end of the subsection, we proceed by bounding the suprema of the empirical process in the preceding display.

We begin by introducing requisite concepts in empirical process theory, which we use in the rest of the proof; we refer readers to Vaart and Wellner [165] for a comprehensive treatment. Recall the definition of Orlicz norms, which allows controlling the tail behavior of random variables.

Definition 7 (Orlicz norms)**.** *Let $\psi : \mathbb{R} \to \mathbb{R}$ be a non-decreasing, convex function with $\psi(0) = 0$. For any random variable W, its Orlicz norm $\|W\|_\psi$ is*

$$\|W\|_\psi := \inf\left\{t > 0 : \mathbb{E}\left[\psi\left(\frac{|W|}{t}\right)\right] \leq 1\right\}.$$

Remark 1. *From Markov's inequality, we have*

$$\mathbb{P}(|W| > t) \le \mathbb{P}\left(\psi\left(\frac{|W|}{\|W\|_\psi}\right) \ge \psi\left(\frac{t}{\|W\|_\psi}\right)\right) \le \psi\left(\frac{t}{\|W\|_\psi}\right)^{-1}.$$

For $\psi_p(s) = e^{s^p} - 1$, a similar argument yields

$$\mathbb{P}(|W| > t) \le 2\exp\left(-t^p/\|W\|_{\psi_p}^p\right). \tag{B.1.3}$$

A sub-Gaussian random variable $\mu(Z)$ with parameter σ^2 has bounded Orlicz norm $\|\mu(Z)\|_{\psi_2} \le 2\sigma$ (see, for example, Wainwright [183, Section 2.4] and Vaart and Wellner [165, Lemma 2.2.1]).

Remark 2. *The converse also holds: for W such that $\mathbb{P}(|W| > t)] \le c_1 \exp(-c_2 t^p)$ for all t, and constants $c_1, c_2 > 0$ and $p \ge 1$, Fubini gives*

$$\mathbb{E}\left[\exp\left(\frac{|W|^p}{t^p}\right) - 1\right] = \mathbb{E}\left[\int_0^{|W|^p} t^{-1/p}\exp(t^{-1/p}s)ds\right] = \int_0^\infty \mathbb{P}(|W|^p > s)t^{-1/p}\exp\left(t^{-1/p}s\right)ds.$$

Using the tail probability bound, the preceding display is bounded by

$$c_1\int_0^\infty \exp(-c_2 s)t^{-1/p}\exp(t^{-1/p}s)ds = \frac{c_1 t^{-1/p}}{c_2 - t^{-1/p}}.$$

So the Orlicz norm $\|W\|_{\psi_p}$ is bounded by $\left(\frac{1+c_1}{c_2}\right)^{1/p}$.

In the following, we let W be the right hand side of the bound (B.1.2), and control its Orlicz norm $\|W\|_{\psi_2}$ using Dudley's entropy integral [165]. We use the standard notion of the covering number. For a vector space $\mathcal{V}$, let $V \subset \mathcal{V}$ be a collection of vectors. Letting $\|\cdot\|$ be a norm on $\mathcal{V}$, a collection $\{v_1, \dots, v_N\} \subset \mathcal{V}$ is an *ϵ-cover* of $\mathcal{V}$ if for each $v \in \mathcal{V}$, there is a v_i satisfying $\|v - v_i\| \le \epsilon$. The *covering number* of V with respect to $\|\cdot\|$ is

$$N(\epsilon, V, \|\cdot\|) := \inf\{N \in \mathbb{N} : \text{ there is an } \epsilon\text{-cover of } V \text{ with respect to } \|\cdot\|\}.$$

For a collection $\mathcal{H}$ of functions $f : \mathcal{Z} \to \mathbb{R}$, let F be its envelope function such that $|f(z)| \leq F(z)$ for all $z \in \mathcal{Z}$. The following result controls the suprema of empirical processes using the (uniform) metric entropy. The result is based on involved chaining arguments [165, Section 2.14].

Lemma 33 (Vaart and Wellner [165, Theorem 2.14.1 and 2.14.5])**.**

$$\sqrt{|S_2|}\left\|\sup_{f\in\mathcal{H}}\left|\frac{1}{|S_2|}\sum_{i\in S_2} f(Z_i) - \mathbb{E}f(Z)\right|\right\|_{\psi_2}$$
$$\lesssim \|F\|_{\psi_2} + \|F\|_{L^2(P)} \sup_Q \int_0^1 \sqrt{1+\log N(\epsilon \|F\|_{L^2(Q)}, \mathcal{H}, L^2(Q))}d\epsilon,$$

where the supremum is over all discrete probability measures Q such that $\|F\|_{L^2(Q)} > 0$.

Evidently, $F(z) = (\mu(z))_+ = \mu(z)$ is an envelope function for the following class of functions

$$\mathcal{F} = \{z \mapsto (\mu(z) - \eta)_+ : \eta \geq 0\}.$$

Using the tail probability bound (B.1.3), we conclude

$$\sup_{\eta\geq 0}\left|\frac{1}{|S_2|}\sum_{i\in S_2}(\mu(Z_i)-\eta)_+ - \mathbb{E}\,(\mu(Z)-\eta)_+\right|$$
$$\lesssim \sqrt{\frac{\log\left(\frac{2}{\delta}\right)}{|S_2|}}\left(\|F\|_{\psi_2} + \|F\|_2 \sup_Q \int_0^1 \sqrt{1+\log N(\epsilon \|F\|_{L^2(Q)}, \mathcal{H}, L^2(Q))}d\epsilon\right),$$

with probability at least $1-\delta$.

Since we have $\|F\|_{L^2(P)} \leq \|F\|_{\psi_2} \lesssim \sigma$, it now suffices to show that the above uniform metric entropy is bounded by a universal constant. We use the standard notion of VC-dimension [165, Chapter 2.6, page 135].

Lemma 34 (Vaart and Wellner [165, Theorem 2.6.7])**.** *Let $\mathsf{VC}(\mathcal{H})$ be the VC-dimension of the colelction of subsets $\{(z,t) : t < f(x)\}$ for $f \in \mathcal{H}$. For any probability measure Q such that*

$\|F\|_{L^2(Q)} > 0$ *and* $0 < \epsilon < 1$, *we have*

$$N(\epsilon \|F\|_{L^2(Q)}, \mathcal{H}, L^2(Q)) \lesssim \mathsf{VC}(\mathcal{H})(16e)^{\mathsf{VC}(\mathcal{H})} \left(\frac{1}{\epsilon}\right)^{2(\mathsf{VC}(\mathcal{H})-1)}.$$

Translations of a monotone function on $\mathbb{R}$ has VC-dimension 2.

Lemma 35 (Vaart and Wellner [165, Theorem 2.6.16]). *The class of functions* $\mathcal{F}' = \{z \mapsto (\mu(z) - \eta)_+ : \eta \in \mathbb{R}\}$ *has VC-dimension* $\mathsf{VC}(\mathcal{H}') = 2$.

From Lemmas 34 and 35, we conclude that for the function class $\mathcal{F} = \{z \mapsto (\mu(z) - \eta)_+ : \eta \geq 0\}$, the uniform metric entropy

$$\sup_Q \int_0^1 \sqrt{1 + \log N(\epsilon \|F\|_{L^2(Q)}, \mathcal{H}, L^2(Q))} d\epsilon$$

is bounded by a universal constant. This gives our desired result.

Proof of Claim 32 To show the bound (B.1.2), we use the dual reformulation for both $\mathsf{W}_\alpha(\mu)$ and its empirical approximation $\widehat{\mathsf{W}}_{\alpha,k}(\mu)$ on S_2. For any probability measure P, recall two different definitions of the quantile of $\mu(Z)$

$$P^{-1}_{1-\alpha}(\mu(Z)) := \inf\{t : \mathcal{P}_Z(\mu(Z) \leq t) \geq 1 - \alpha\}$$
$$P^{-1}_{1-\alpha,+}(\mu(Z)) := \inf\{t : \mathcal{P}_Z(\mu(Z) \leq t) > 1 - \alpha\}.$$

We call $P^{-1}_{1-\alpha,+}(\mu(Z))$ the upper $(1-\alpha)$-quantile. The two values characterize the optimal solution set of the dual problem (2.2.2); they are identical when $\mu(Z)$ has a positive density at $P^{-1}_{1-\alpha}(\mu(Z))$.

Lemma 36 (Rockafellar and Uryasev [162, Theorem 10]). *For any probability measure P such that $\mu(Z) \geq 0$ P-a.s. and $\mathbb{E}_P[\mu(Z)_+] < \infty$, we have*

$$[P^{-1}_{1-\alpha}(\mu(Z)), P^{-1}_{1-\alpha,+}(\mu(Z))] = \operatorname*{argmin}_{\eta \in \mathbb{R}} \left\{ \frac{1}{\alpha} \mathbb{E}_P \left(\mu(Z) - \eta\right)_+ + \eta \right\}.$$

Since P was an arbitrary measure in Lemmas 11 and 36, identical results follow for the empirical distribution on S_2. Hence, we have

$$\begin{aligned}\left|\widehat{\mathsf{W}}_{\alpha,k}(\mu) - \mathsf{W}_\alpha(\mu)\right| &= \left|\inf_{\eta\in\mathbb{R}}\left\{\frac{1}{\alpha|S_2|}\sum_{i\in S_2}(\mu(Z_i)-\eta)_+ + \eta\right\} - \inf_{\eta\in\mathbb{R}}\left\{\frac{1}{\alpha}\mathbb{E}\,(\mu(Z)-\eta)_+ + \eta\right\}\right| \\ &= \left|\inf_{\eta\geq 0}\left\{\frac{1}{\alpha|S_2|}\sum_{i\in S_2}(\mu(Z_i)-\eta)_+ + \eta\right\} - \inf_{\eta\geq 0}\left\{\frac{1}{\alpha}\mathbb{E}\,(\mu(Z)-\eta)_+ + \eta\right\}\right|\end{aligned}$$

where we used Lemma 36 to restrict the feasible region in the last equality. The preceding display is then bounded by

$$\sup_{\eta\geq 0}\left|\frac{1}{\alpha|S_2|}\sum_{i\in S_2}(\mu(Z_i)-\eta)_+ + \eta - \frac{1}{\alpha}\mathbb{E}\,(\mu(Z)-\eta)_+ - \eta\right|.$$

B.1.2 Proof of Theorem 4

We abuse notation and use C for a numerical constant that may change line to line. From the decomposition (2.3.1), it suffices to bound term (a) and term (b) separately. Since $\widehat{\mu}_1(\cdot)$ is trained on a sample S_1 independent from S_2 used to estimate the worst-case subpopulation performance (Eq. (B.1.1)), we can directly apply Proposition 5 to bound term (b). Recalling that any bounded random variable random variable taking values in $[0, B]$ is sub-Gaussian with parameter $B^2/4$, Proposition 5 implies

$$\left|\widehat{\mathsf{W}}_{\alpha,k}(\widehat{\mu}_1) - \mathsf{W}_\alpha(\widehat{\mu}_1)\right| \leq \frac{CB}{\alpha}\sqrt{\frac{\log(2/\delta)}{|S_2|}} \text{ with probability at least } 1-\delta.$$

To bound term (a) in the decomposition (2.3.1), we note

$$
\begin{aligned}
\left|\mathsf{W}_\alpha(\widehat{\mu}_1) - \mathsf{W}_\alpha(\mu^\star)\right| &\le \frac{1}{\alpha}\sup_\eta \left|\mathbb{E}\left[(\widehat{\mu}_1(Z)-\eta)_+ \mid S_1\right] - \mathbb{E}\left(\mu^\star(Z)-\eta\right)_+\right| \\
&\le \frac{1}{\alpha}\mathbb{E}\left[\left|\widehat{\mu}_1(Z)-\mu^\star(Z)\right| \mid S_1\right] \\
&\le \frac{1}{\alpha}\sqrt{\mathbb{E}\left[\left(\widehat{\mu}_1(Z)-\mu^\star(Z)\right)^2 \mid S_1\right]} = \frac{1}{\alpha}\sqrt{\mathsf{err}(\mathcal{H}, S_1)},
\end{aligned}
$$

where the first inequality follows from the dual (2.2.2), the second inequality follows from the non-expansiveness of the function $(\cdot)_+$, the last inequality uses Holder inequality, and we define the generalization error for the first-stage estimation problem (2.2.3) based on S_1,

$$
\begin{aligned}
\mathsf{err}(\mathcal{H}, S_1) :=& \mathbb{E}\left[\left(\mu^\star(Z)-\widehat{\mu}_1(Z)\right)^2 \mid S_1\right] \\
=& \mathbb{E}\left[(\ell(\theta(X);Y)-\widehat{\mu}_1(Z))^2 \mid S_1\right] - \mathbb{E}(\ell(\theta(X);Y)-\mu^\star(Z))^2 \\
=& \mathbb{E}\left[(\ell(\theta(X);Y)-\widehat{\mu}_1(Z))^2 \mid S_1\right] - \mathbb{E}(\ell(\theta(X);Y)-\bar{\mu}(Z))^2 + \mathbb{E}\left(\mu^\star(Z)-\bar{\mu}(Z)\right)^2
\end{aligned}
$$

We use the following concentration result based on the localized Rademacher complexity [163].

Lemma 37 (Bartlett *et al.* [163, Corollary 5.3])**.** *Let Assumption C hold. Then, with probability at least* $1-\delta$*,*

$$
\mathbb{E}\left[(\ell(\theta(X);Y)-\widehat{\mu}_1(Z))^2 \mid S_1\right] - \mathbb{E}(\ell(\theta(X);Y)-\bar{\mu}(Z))^2 \le CB^2\left(r^\star_{|S_1|} + \frac{\log(1/\delta)}{|S_1|}\right).
$$

Using $\sqrt{a+b+c} \le \sqrt{a}+\sqrt{b}+\sqrt{c}$ for $a, b, c \ge 0$, we have the desired result.

B.1.3 Proof of Theorem 5

Instead of the decomposition (2.3.1) we use for Theorem 4, we use an alterantive form

$$
\widehat{\mathsf{W}}_{\alpha,k}(\widehat{\mu}_1) - \mathsf{W}_\alpha(\mu^\star) = \underbrace{\widehat{\mathsf{W}}_{\alpha,k}(\widehat{\mu}_1) - \widehat{\mathsf{W}}_{\alpha,k}(\mu^\star)}_{(a):\text{first stage}} + \underbrace{\widehat{\mathsf{W}}_{\alpha,k}(\mu^\star) - \mathsf{W}_\alpha(\mu^\star)}_{(b):\text{second stage}} \tag{B.1.4}
$$

Term (b) can be bounded using Proposition 5 as before. Without assuming $\mu^\star \in \mathcal{H}$, recall that any bounded random variable taking values in $[0, B]$ is sub-Gaussian with parameter $B^2/4$, so Proposition 5 yields

$$\left|\widehat{\mathsf{W}}_{\alpha,k}(\mu^\star) - \mathsf{W}_\alpha(\mu^\star)\right| \leq C\frac{B}{\alpha}\sqrt{\frac{\log(2/\delta)}{|S_2|}} \text{ with probability at least } 1-\delta.$$

It remains to bound term (a). Our starting point is the bound

$$\begin{aligned}\left|\widehat{\mathsf{W}}_{\alpha,k}(\widehat{\mu}_1) - \widehat{\mathsf{W}}_{\alpha,k}(\mu^\star)\right| &\leq \frac{1}{\alpha}\sup_\eta \left|\frac{1}{|S_2|}\sum_{i\in S_2}\left((\widehat{\mu}_1(Z_i)-\eta)_+ - \left(\mu^\star(Z_i)-\eta\right)_+\right)\right| \\ &\leq \frac{1}{\alpha|S_2|}\sum_{i\in S_2}|\widehat{\mu}_1(Z_i)-\mu^\star(Z_i)| \leq \frac{1}{\alpha}\left(\frac{1}{|S_2|}\sum_{i\in S_2}\left(\widehat{\mu}_1(Z_i)-\mu^\star(Z_i)\right)^2\right)^{1/2}.\end{aligned} \tag{B.1.5}$$

Denoting the residuals by $\zeta := \ell(\theta(X);Y) - \mu^\star(Z)$ and $\zeta_i := \ell(\theta(X_i);Y_i) - \mu^\star(Z_i)$ for all $i \in S_2$, we have the identity

$$\begin{aligned}\frac{1}{|S_2|}\sum_{i\in S_2}\left(\widehat{\mu}_1(Z_i)-\mu^\star(Z_i)\right)^2 &= \Delta_{S_2}(\widehat{\mu}_1) - \Delta_{S_2}(\mu^\star) + \frac{2}{|S_2|}\sum_{i\in S_2}\zeta_i\left(\widehat{\mu}_1(Z_i)-\mu^\star(Z_i)\right) \\ &= \left[\Delta_{S_2}(\widehat{\mu}_1) - \Delta_{S_2}(\bar{\mu})\right] + \left[\Delta_{S_2}(\bar{\mu}) - \Delta_{S_2}(\mu^\star)\right] + \frac{2}{|S_2|}\sum_{i\in S_2}\zeta_i\left(\widehat{\mu}_1(Z_i)-\mu^\star(Z_i)\right)\end{aligned} \tag{B.1.6}$$

First notice by definition of $\mu^\star$ that ζ has conditional mean $\mathbb{E}[\zeta \mid Z] = 0$. Hence, conditional on S_1, $\mathbb{E}\left[\zeta\left(\widehat{\mu}_1(Z) - \mu^\star(Z)\right) \mid S_1\right] = 0$. Since $\zeta_i\left(\widehat{\mu}_1(Z_i)-\mu^\star(Z_i)\right)$ are bounded in $[-B^2, B^2]$ and i.i.d. conditional on S_1, Hoeffding inequality [183, Ch. 2] yields

$$\frac{1}{|S_2|}\sum_{i\in S_2}\zeta_i\left(\widehat{\mu}_1(Z_i)-\mu^\star(Z_i)\right) \leq B^2\sqrt{\frac{2\log(1/\delta)}{|S_2|}} \text{ with probability at least } 1-\delta. \tag{B.1.7}$$

Similarly, Hoeffding inequality implies with probability at least $1 - \delta$,

$$\Delta_{S_2}(\bar{\mu}) - \Delta_{S_2}(\mu^\star) \le \mathbb{E}(\ell(\theta(X);Y) - \bar{\mu}(Z))^2 - \mathbb{E}(\ell(\theta(X);Y) - \mu^\star(Z))^2 + 2B^2\sqrt{\frac{2\log(1/\delta)}{|S_2|}}. \quad \text{(B.1.8)}$$

Note the definition of the conditional risk $\mu^\star(Z) = \mathbb{E}[\ell(\theta(X);Y) \mid Z]$ implies

$$\mathbb{E}(\ell(\theta(X);Y) - \bar{\mu}(Z))^2 - \mathbb{E}(\ell(\theta(X);Y) - \mu^\star(Z))^2 = \mathbb{E}(\bar{\mu}(Z) - \mu^\star(Z))^2 = \left\|\bar{\mu} - \mu^\star\right\|_{L^2}^2 .$$

Hence, on the event where inequalities (B.1.7) and (B.1.8) hold (with probability at least $1 - 2\delta$),

$$\frac{1}{|S_2|}\sum_{i \in S_2}\left(\widehat{\mu}_1(Z_i) - \mu^\star(Z_i)\right)^2 \le \Delta_{S_2}(\widehat{\mu}_1) - \Delta_{S_2}(\bar{\mu}) + \left\|\bar{\mu} - \mu^\star\right\|_{L^2}^2 + 4B^2\sqrt{\frac{2\log(1/\delta)}{|S_2|}}.$$

Noticing $\sqrt{a+b+c} \le \sqrt{(a)_+} + \sqrt{b} + \sqrt{c}$ for $b, c \ge 0$, we obtain the first result.

Proof of Theorem 5 with a convex model class

Now assume further that $\mathcal{H}$ is convex. We adopt an alternative three-element decomposition:

$$\mathsf{W}_\alpha(\mu^\star) - \widehat{\mathsf{W}}_{\alpha,k}(\widehat{\mu}_1) = \underbrace{\widehat{\mathsf{W}}_{\alpha,k}(\widehat{\mu}_1) - \widehat{\mathsf{W}}_{\alpha,k}(\bar{\mu})}_{(a):\ \text{first stage}} + \underbrace{\widehat{\mathsf{W}}_{\alpha,k}(\bar{\mu}) - \mathsf{W}_\alpha(\bar{\mu})}_{(b):\ \text{second stage}} + \underbrace{\mathsf{W}_\alpha(\bar{\mu}) - \mathsf{W}_\alpha(\mu^\star)}_{(c):\ \text{approximation error}}. \quad \text{(B.1.9)}$$

The approximation error term (c) can be bounded by

$$\begin{aligned}\left|\mathsf{W}_\alpha(\bar{\mu}) - \mathsf{W}_\alpha(\mu^\star)\right| &\le \frac{1}{\alpha}\sup_\eta\left|\mathbb{E}\left(\bar{\mu}(Z) - \eta\right)_+ - \mathbb{E}\left(\mu^\star(Z) - \eta\right)_+\right| \\ &\le \frac{1}{\alpha}\mathbb{E}\left|\bar{\mu}(Z) - \mu^\star(Z)\right| = \frac{1}{\alpha}\left\|\bar{\mu} - \mu^\star\right\|_{L^1} .\end{aligned}$$

The second-stage error term (b) can be bounded using Proposition 5 by

$$\left|\widehat{\mathsf{W}}_{\alpha,k}(\bar{\mu}) - \mathsf{W}_\alpha(\bar{\mu})\right| \le C\frac{B}{\alpha}\sqrt{\frac{\log(2/\delta)}{|S_2|}} \text{ with probability at least } 1 - \delta.$$

The first-stage error term (a) can be bounded, similarly to Equation (B.1.5) by

$$\left|\widehat{\mathsf{W}}_{\alpha,k}(\widehat{\mu}_1) - \widehat{\mathsf{W}}_{\alpha,k}(\bar{\mu})\right| \leq \frac{1}{\alpha}\sup_{\eta}\left|\frac{1}{|S_2|}\sum_{i\in S_2}((\widehat{\mu}_1(Z_i)-\eta)_+ - (\bar{\mu}(Z_i)-\eta)_+)\right|$$
$$\leq \frac{1}{\alpha|S_2|}\sum_{i\in S_2}|\widehat{\mu}_1(Z_i)-\bar{\mu}(Z_i)| \leq \frac{1}{\alpha}\left(\frac{1}{|S_2|}\sum_{i\in S_2}(\widehat{\mu}_1(Z_i)-\bar{\mu}(Z_i))^2\right)^{1/2}. \tag{B.1.10}$$

Again, we have the identity

$$\frac{1}{|S_2|}\sum_{i\in S_2}(\widehat{\mu}_1(Z_i)-\bar{\mu}(Z_i))^2 = \Delta_{S_2}(\widehat{\mu}_1) - \Delta_{S_2}(\bar{\mu}) + \frac{2}{|S_2|}\sum_{i\in S_2}(\ell(\theta(X_i);Y_i)-\bar{\mu}(Z_i))(\widehat{\mu}_1(Z_i)-\bar{\mu}(Z_i)).$$

Since we assume the model class $\mathcal{H}$ is convex and $\widehat{\mu}_1 \in \mathcal{H}$, the first-order condition of $\bar{\mu} \in \arg\min_{\mu\in\mathcal{H}} \mathbb{E}(\ell(\theta(X);Y)-\mu(Z))^2$ gives

$$\mathbb{E}[(\ell(\theta(X);Y)-\bar{\mu}(Z))(\widehat{\mu}_1(Z)-\bar{\mu}(Z)) \mid Z, S_1] \leq 0,$$

so Hoeffding inequality implies with probability at least $1-\delta$,

$$\frac{1}{|S_2|}\sum_{i\in S_2}(\ell(\theta(X_i);Y_i)-\bar{\mu}(Z_i))(\widehat{\mu}_1(Z_i)-\bar{\mu}(Z_i)) \leq B^2\sqrt{\frac{2\log(1/\delta)}{|S_2|}}. \tag{B.1.11}$$

Hence,

$$\frac{1}{|S_2|}\sum_{i\in S_2}(\widehat{\mu}_1(Z_i)-\bar{\mu}(Z_i))^2 \leq \Delta_{S_2}(\widehat{\mu}_1) - \Delta_{S_2}(\bar{\mu}) + 2B^2\sqrt{\frac{2\log(1/\delta)}{|S_2|}} \text{ w.p. } \geq 1-\delta.$$

Using $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$ for all $a, b \geq 0$, we obtain the desired result.

B.1.4 Proof of Theorem 6

For ease of notation, we suppress any dependence on the prediction model $\theta(X)$ under evaluation. Consider any $\alpha_1, \alpha_2 \in (0,1]$ with $\alpha_1 < \alpha_2$. For convenience denote $\xi_1 := \widehat{P}^{-1}_{1-\alpha_1}(\widehat{h}(Z))$ and $\xi_2 := \widehat{P}^{-1}_{1-\alpha_2}(\widehat{h}(Z))$, so $\xi_1 \geq \xi_2$ and $\widehat{\mathsf{W}}_{\alpha_1}(\widehat{h}) \geq \widehat{\mathsf{W}}_{\alpha_2}(\widehat{h})$. Denote by $\widehat{\mathcal{P}}$ the empirical probability measure induced by $(Z_i : i \in S_2)$. Notice that

$$\begin{aligned}
\widehat{\mathbb{E}}[\widehat{h}(Z)-\xi_2]_+ - \widehat{\mathbb{E}}[\widehat{h}(Z)-\xi_1]_+ &= \widehat{\mathbb{E}}[\widehat{h}(Z)-\xi_2; \widehat{h}(Z) > \xi_2] - \widehat{\mathbb{E}}[\widehat{h}(Z)-\xi_1; \widehat{h}(Z) \geq \xi_1] \\
&= \underbrace{\widehat{\mathbb{E}}[\widehat{h}(Z)-\xi_1; \xi_2 < \widehat{h}(Z) < \xi_1]}_{\leq 0} + (\xi_1-\xi_2)\underbrace{\widehat{\mathbb{P}}(\widehat{h}(Z) > \xi_2)}_{\leq \alpha_2} \\
&\leq (\xi_1-\xi_2)\alpha_2.
\end{aligned}$$

Hence, by Lemma 36,

$$\widehat{\mathsf{W}}_{\alpha_1}(\widehat{h}) - \widehat{\mathsf{W}}_{\alpha_2}(\widehat{h}) = \left(\frac{\widehat{\mathbb{E}}[\widehat{h}(Z)-\xi_1]_+}{\alpha_1} + \xi_1\right) - \left(\frac{\widehat{\mathbb{E}}[\widehat{h}(Z)-\xi_2]_+}{\alpha_2} + \xi_2\right) \geq \frac{\widehat{\mathbb{E}}[\widehat{h}(Z)-\xi_1]_+}{\alpha_1\alpha_2}(\alpha_2-\alpha_1),$$

meaning

$$|\alpha_1-\alpha_2| \leq \frac{\alpha_1\alpha_2|\widehat{\mathsf{W}}_{\alpha_1}(\widehat{h}) - \widehat{\mathsf{W}}_{\alpha_2}(\widehat{h})|}{\widehat{\mathbb{E}}[\widehat{h}(Z) - \widehat{P}^{-1}_{1-\alpha_1}(\widehat{h}(Z))]_+}.$$

Now suppose $\alpha^\star \geq \underline{\alpha}$. Notice the boundedness of $h(Z)$ implies W_α and $\widehat{\mathsf{W}}_{\alpha,k}$ are continuous and nonincreasing in α, so the definitions (2.2.1) and (2.2.4) imply $\mathsf{W}_{\alpha^\star}(\theta) = \bar{\ell} = \widehat{\mathsf{W}}_{\widehat{\alpha}}(\widehat{h})$. Plugging $\widehat{\alpha}$ and $\alpha^\star$ into the inequality above, we know with probability at least $1-\delta$,

$$|\alpha^\star - \widehat{\alpha}| \leq \frac{\widehat{\alpha}\alpha^\star\left|\widehat{\mathsf{W}}_{\alpha^\star}(\widehat{h}) - \mathsf{W}_{\alpha^\star}(\theta)\right|}{\widehat{\mathbb{E}}\left[\widehat{h}(Z) - \widehat{P}^{-1}_{1-\alpha^\star\wedge\widehat{\alpha}}\left(\widehat{h}(Z)\right)\right]_+} \leq \frac{\widehat{\alpha}U(\delta)}{\widehat{\mathbb{E}}\left[\widehat{h}(Z) - \widehat{P}^{-1}_{1-\underline{\alpha}\wedge\widehat{\alpha}}\left(\widehat{h}(Z)\right)\right]_+}.$$

B.2 Additional experiment details

In this section, we present additional experiments for the Functional Map of the World (FMoW) dataset. Due to the ever-changing nature of aerial images and the uneven availability of data from different regions, it is imperative that ML models maintain good performance under temporal (learn from the past and generalize to future) and spatial distribution shifts (learn from one region and generalize to another). Without having access to the out-of-distribution samples, our diagnostic raises awareness on brittleness of model performance against subpopulation shifts.

B.2.1 Dataset Description

The original Functional Map of the World (FMoW) dataset by [166] consists of over 1 million images from over 200 countries. We use a variant, FMoW-WILDS, proposed by Koh *et al.* [101], which temporally groups observations to simulate distribution shift across time. Each data point includes an RGB satellite image x, and a corresponding label y on the land / building use of the image (there are 62 different classes). FMoW-WILDS splits data into non-overlapping time periods: we train and validate models $\theta(\cdot)$ on data collected from years 2002-2013, and simulate distribution shift by looking at data collected during 2013-2018. Data collected during 2002-2013 ("in-distribution") is split into training (n =76,863), validation (n =19,915), and test (n =11,327). Data collected during 2013-2018 ("out-of-distribution") is split into two sets: one consisting of observations from years 2013-2016 (n =19,915), and another consisting of observations from years 2016-2018 (n =22,108). All data splits contain images from a diverse array of geographic regions. We evaluate the worst-case subpopulation performance on in-distribution validation data, and study model performance under distribution shift on data after 2016.

B.2.2 Models Evaluated

We consider *DenseNet* models as reported by Koh *et al.* [101], including the vanilla empirical risk minimization (ERM) model and models trained with robustness interventions (IRM [175] method; Koh *et al.* [101] notes that ERM's performance closely match or outperform "robust" counterparts even under distribution shift. We also evaluate ImageNet pre-trained *DPN*-68 model from Miller *et al.* [177]. As separate experiments, we also consider *ResNet*-18 and *VGG*-11 from Miller *et al.* [177], and the results are reported in B.2.6.

CLIP (Contrastive Language-Image Pre-training) is a newly proposed model pre-trained on 400M image-text pairs, and has been shown to exhibit strong zero-shot performance on out-of-distribution samples [121]. Although not specifically designed for classification tasks, CLIP can be used for classification by predicting the class whose encoded text is the closest to the encoded image. We consider the weight-space ensembled *CLIP WiSE* models proposed in [178] as it is observed that these models exhibit robust behavior on FMoW. *CLIP WiSE* models are constructed by linearly combining the model weights of *CLIP ViT-B16 Zeroshot model* and *CLIP ViT-B16 FMoW end-to-end finetuned* model.

To illustrate the usage of our method, we choose the *CLIP WiSE* model that has similar ID validation accuracy as the *DenseNet* Models. This turns out to be putting 60% weight on *CLIP ViT-B16 Zeroshot model* and 40% weight on *CLIP ViT-B16 FMoW end-to-end finetuned*. *DenseNet* Models have average ID validation loss $2.4 - 2.8$, but *CLIP WiSE* has average ID validation loss 1.6. To ensure fair comparison, we calibrate the temperature parameter such that the average loss of *CLIP WiSE* matches the worst average loss of the models considered. We deliberately make *CLIP WiSE* no better than any *DenseNet* Models, in the hope that out metric will recover its robustness property.

B.2.3 Flexibility of our metric

We implement Algorithm 5 by partitioning the ID validation data into two; we estimate $\mu^\star(Z)$ using XGBoost on one sample, and estimate $\mathsf{W}_\alpha(\cdot)$ at varying subpopulation size α on the other. By switching the role of each split, our final estimator averages two versions of $\widehat{\mathsf{W}}_{\alpha,k}(\widehat{\mu})$.

A less conservative Z

In Section 2.4, we report results when Z is defined over all metadata consisting of (longitude, latitude, cloud cover, region, year), as well as the label Y. Defining subpopulations over such a wide range of variables may be overly conservative in some scenarios, and to illustrate the flexibility of our approach, we now showcase a more tailored definition of subpopulations. Since FMoW-

#	**Text Prompt**
1	"CLASSNAME"
2	"a picture of a CLASSNAME."
3	"a photo of a CLASSNAME."
4	"an image of an CLASSNAME"
5	"an image of a CLASSNAME in asia."
6	"an image of a CLASSNAME in africa."
7	"an image of a CLASSNAME in the americas."
8	"an image of a CLASSNAME in europe."
9	"an image of a CLASSNAME in oceania."
10	"satellite photo of a CLASSNAME"
11	"satellite photo of an CLASSNAME"
12	"satellite photo of a CLASSNAME in asia."
13	"satellite photo of a CLASSNAME in africa."
14	"satellite photo of a CLASSNAME in the americas."
15	"satellite photo of a CLASSNAME in europe."
16	"satellite photo of a CLASSNAME in oceania."
17	"an image of a CLASSNAME"

Table B.1: Text prompts for CLIP text encoders

WILDS is specifically designed for spatiotemporal shifts, a natural choice of Z is to condition on (region, year). Motivated by our observation that some classes are harder to predict than others (Figure 2.3(b)), we also consider Z= (region, year, label Y). We plot our findings in Figure B.1. If we simply define Z= (year, region), the corresponding worst-case subpopulation performance is less pessimistic. However, when we add labels to Z, we again see a drastic decrease in the worst-case subpopulation performance, and that *CLIP WiSE-FT* outperforms all other models by a significant amount. This is consistent with our motivation in defining subpopulations over labels; our procedure automatically takes into account the interplay between class labels and spatiotemporal information.

Using semantics of the labels

Alternatively, we may wish to define subpopulations over rich natural language descriptions on the input X. To illustrate the flexibility of our procedure in such scenarios, we consider subpopulations defined over the semantic meaning of the class names: CLIP-encoded class names using the 17 prompts reported in Table B.1. For comparison, we report the (estimated) worst-case subpopu-

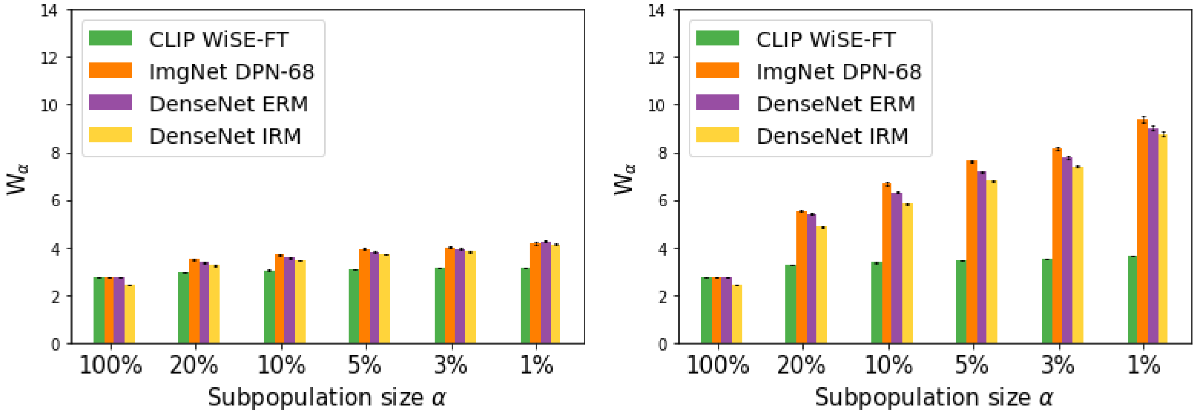


Figure B.1. In the left panel Z = (year, region); in the right panel Z = (year, region, label Y). Here we take Z to contain only spatial and temporal information, a less conservative counterpart to the experiment reported in the main text. We again see that introduction of labels in Z drastically increase our metric, showing varying difficulty in learning different labels.

lation performance (2.1.2) when we take Z = (all metadata, encoded labels) and (all metadata, label Y, encoded labels) in Figure B.2. We observe that in this case, the semantics of the class names do not contribute to further deterioration in robustness, and the relative ordering across models remains unchanged.

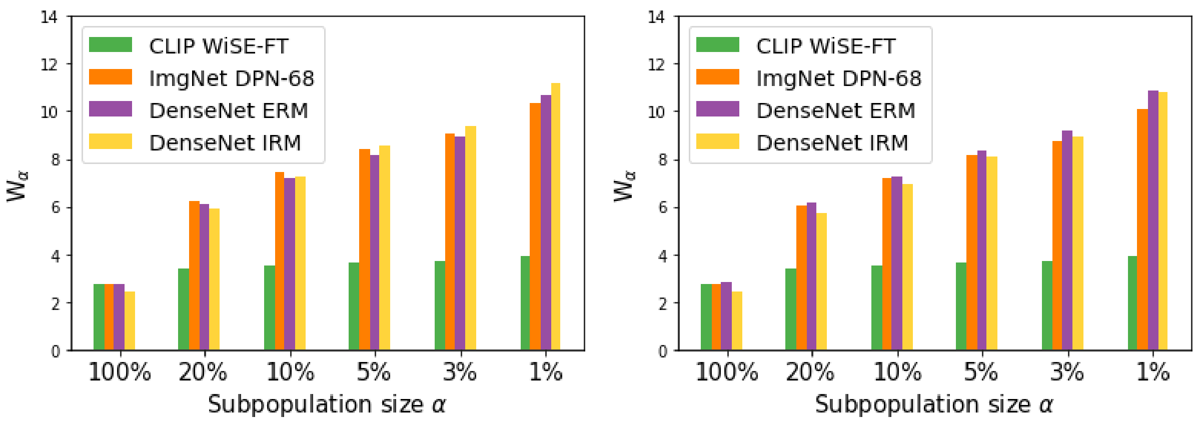


Figure B.2. In the left panel, Z = (all meta, encoded labels); in the right panel, Z = (all meta, label Y, encoded labels). We see that in this case no significant difference is introduced when semantics of the class names are included.

B.2.4 Analysis of spatiotemporal distribution shift

The significant performance drop in the Africa region on data collected from 2016-2018 was also observed in [101, 178]. In Figures B.3-B.4, we plot the number of samples collected from Africa over data splits. In particular, we observe a large number of single-unit and multi-unit residential instances emerge in the OOD data. Data collection systems are often biased against the African continent—often as a result of remnants of colonialism—and addressing such bias is an important topic of future research.

B.2.5 Estimator of model loss

One potential limitation of our approach is $\hat{h}$ does not always estimate the tail losses accurately, and this is important because our approach precisely is designed to counter ML models that perform poorly on tail subpopulations. Figure B.5 plots a histogram of model losses and the estimated conditional risk $\hat{h}$ for *DenseNet ERM* and *CLIP WiSE*, where the y-axis is plotted on a log-scale. It is clear that *DenseNet ERM* has more extreme losses compared to the *CLIP WiSE* model, suggesting that at least part of the reason why *DenseNet ERM* suffers poor loss on subpopulations: it is overly confident when it's incorrect. While a direct comparison is not appropriate since the conditional risk $\mu(Z)$ represent *smoothed* losses, we observe that naive estimators of $\mu(\cdot)$ may consistently underestimate. In this particular instance, since the extent of underestimation is more severe for ImageNet pre-trained models, our experiments are fortuitously providing an even more conservative comparison between the two model classes, instilling confidence in the relative robustness of the *CLIP WiSE* model.

Alternatively, we can directly define the worst-case subpopulation performance (2.1.2) using the 0-1 loss. The discrete nature of the 0-1 loss pose some challenges in estimating $\mu(\cdot)$. While we chose to focus on the cross entropy loss that aligns with model training, we leave a thorough study of 0-1 loss to future work.

B.2.6 Additional comparisons

We use the ensembled *CLIP WiSE* model constructed by averaging the network weights of *CLIP zero-shot* and *CLIP finetuned* models. So far, we used proportion $\lambda = 0.4$ to match the ID validation accuracy of *CLIP WiSE* to that of *DenseNet* models and *DPN*-68. In this subsection, we provide alternative choices:

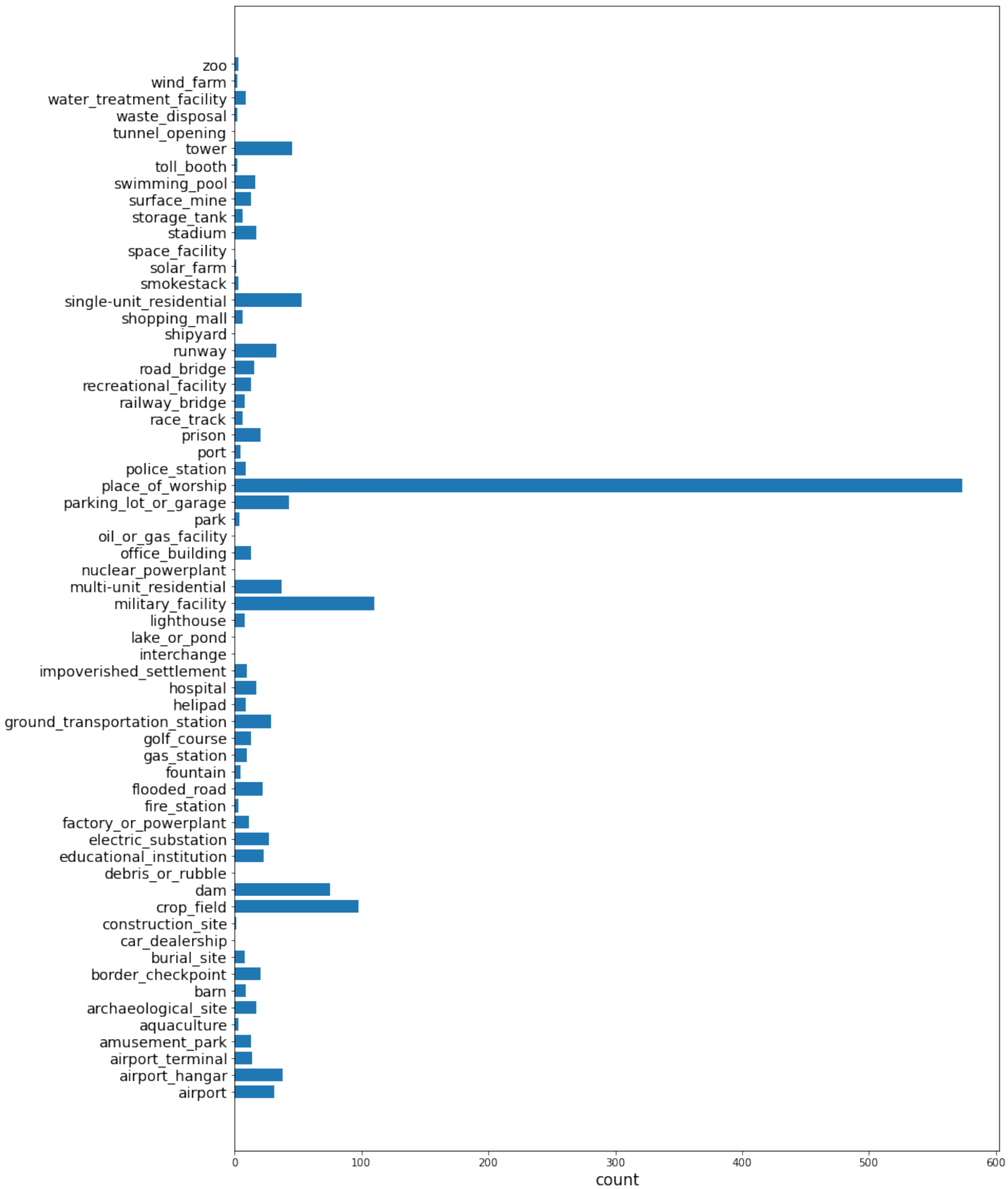


Figure B.3: Instances by class, ID 2002-2013, Africa

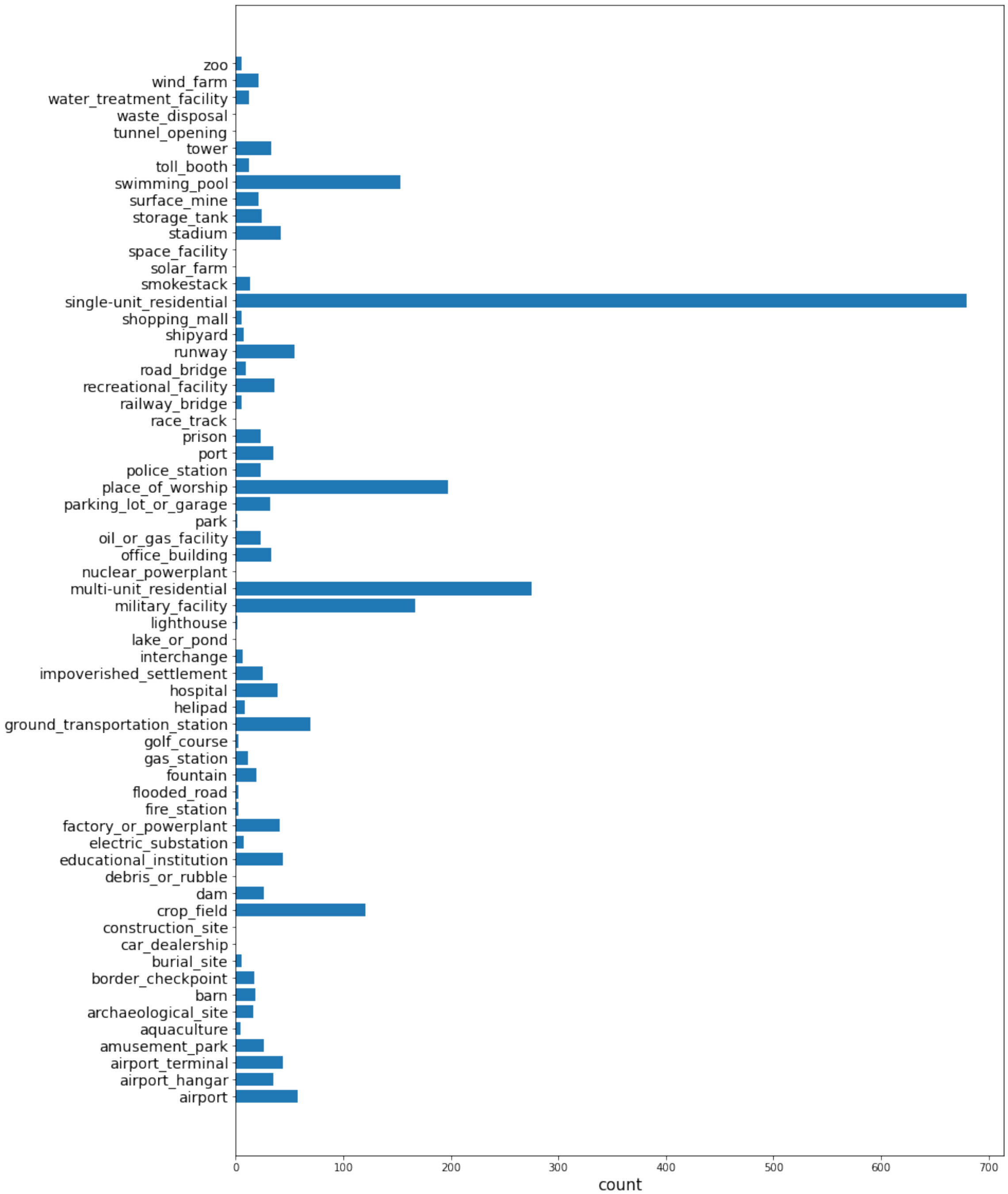


Figure B.4: Instances by class, test 2016-2018, Africa

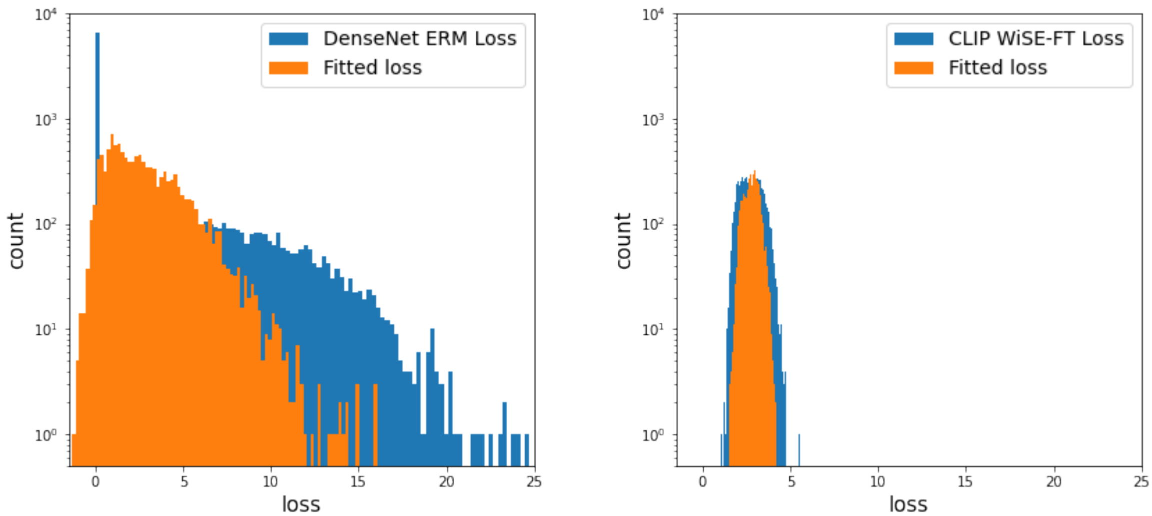


Figure B.5. Histograms of model losses and fitted losses $\hat{h}$. Y-axis count is plotted in **log-scale**. For *DenseNet ERM* model, fitted $\hat{h}$ underestimates the extreme losses (right-tail).

1. $\lambda = 0.24$ to match ID accuracy of *ResNet*-18 of 47%

2. $\lambda = 0.27$ to match ID accuracy of *VGG*-11 of 51%.

Similar to *DPN*-68, *ResNet*-18 and *VGG*-11 are ImageNet pretrained models fine-tuned on FMoW as evaluated by Miller *et al.* [177]. We refer to the two *CLIP WiSE* models as *CLIP WiSE 24* and *CLIP WiSE 27* respectively, and report all model performances below. Again, we observe that our approach successfully picks out the more robust *CLIP WiSE* models, in contrast to the non-robust models chosen by ID accuracy or ID loss.

	ID, 2002-2013			OOD, 2016-2018	
Model	Accuracy	Loss	$W_{0.10}$	Accuracy	Loss
CLIP WiSE 24	0.47	2.84	3.47	0.45	2.85
ResNet-18	0.48	2.84	5.05	0.40	3.36
CLIP WiSE 27	0.51	3.07	3.54	0.48	3.08
VGG-11	0.51	3.06	6.07	0.45	3.68

Table B.2. Additional experiments showcasing our approach successfully identifies more robust models.

B.3 Simulation

In this section, we illustrate the asymptotic convergence of our two-stage estimator $\widehat{\mathsf{W}}_{\alpha,k}(\widehat{\mu}_1)$ of the worst-case subpopulation performance $\mathsf{W}_\alpha(\theta)$. We conduct a binary classification experiment similar to Duchi and Namkoong [131].

Formally, we randomly generate and fix two vectors $\theta, \theta_0^\star \in \mathbb{R}^d$ on the unit sphere. The data-generating distribution is given by $X \overset{iid}{\sim} \mathsf{N}(\gamma, \Sigma)$ and

$$Y \mid X = \begin{cases} \mathrm{sgn}(X^\top \theta_0^\star) & X^1 \le z_{0.95} = 1.645 \\ -\mathrm{sgn}(X^\top \theta_0^\star) & \text{otherwise.} \end{cases}$$

In this data-generating distribution, there is a drastic difference between subpopulations generated by $X^1 \le z_{0.95}$ and $X^1 > z_{0.95}$; typical prediction models will perform poorly on the latter rare group. The loss function is taken to be the hinge loss $\ell(\theta; x, y) = [1 - y \cdot \theta^\top x]_+$, where $y \in \{\pm 1\}$. We take the first covariate X^1 as our protected attribute Z. Let $d = 5$, $\Sigma = \mathsf{I}_5$, $\gamma = 0$.

We fix $\alpha = 0.3$. To analyze the asymptotic convergence of our two-stage estimator, for sample size ranging in 1,000 to 256,000 doubling each time, we run 40 repeated experiments of the estimation procedure on simulated data. We split each sample evenly into S_1 and S_2 and using gradient boosted trees in the package XGBoost [167] to estimate the conditional risk. On a log-scale, we report the mean estimate across random runs in Figure B.6 alongside error bars. To compute the true worst-case subpopulation performance $\mathsf{W}_\alpha(\mu^\star)$ of the conditional risk $\mu^\star(X^1)$, we first run a Monte Carlo simulation for 150,000 copies of $X^1 \sim \mathsf{N}(0, 1)$. For each sampled X^1, we generate 100,000 copies of $(X^2, X^3, X^4, X^5) \sim \mathsf{N}(\mathbf{0}, \mathsf{I}_4)$ independent of X^1 and compute the mean loss among them to approximate the conditional risk $\mu^\star(X^1)$. Finally, we approximate $\mathsf{W}_\alpha(\mu^\star)$ using the empirical distribution of $\mu^\star(\cdot)$, obtaining 6.47×10^{-1}. We observe convergence toward the true value as sample size n grows, verifying the consistency of our two-stage estimator $\widehat{\mathsf{W}}_{\alpha,k}(\widehat{\mu}_1)$.

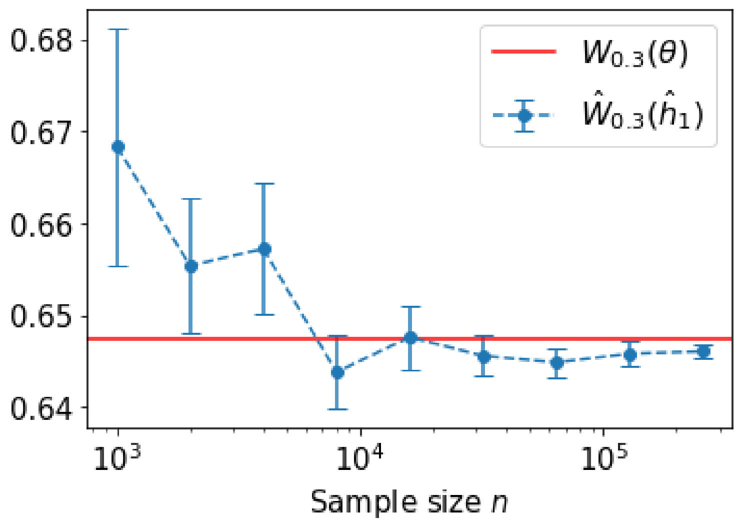


Figure B.6: $\widehat{\mathsf{W}}_{\alpha,k}(\widehat{\mu}_1)$ and $\mathsf{W}_{\alpha}(\theta)$ from simulation experiments with $\alpha = 0.3$