Beyond IID: data-driven decision-making in heterogeneous environments

Anonymous Author(s)
Affiliation
Address
email

Abstract

In this work, we study data-driven decision-making and depart from the classical identically and independently distributed (i.i.d.) assumption. We present a new framework in which historical samples are generated from unknown and different distributions, which we dub *heterogeneous environments*. These distributions are assumed to lie in a heterogeneity ball with known radius and centered around the (also) unknown future (out-of-sample) distribution on which the performance of a decision will be evaluated. We quantify the asymptotic worst-case regret that is achievable by central data-driven policies such as Sample Average Approximation, but also by rate-optimal ones, as a function of the radius of the heterogeneity ball. Our work shows that the type of achievable performance varies considerably across different combinations of problem classes and notions of heterogeneity. We demonstrate the versatility of our framework by comparing achievable guarantees for the heterogeneous version of widely studied data-driven problems such as pricing, ski-rental, and newsvendor. En route, we establish a new connection between data-driven decision-making and distributionally robust optimization.

1 Introduction

In optimization under uncertainty, the desirability of a decision (e.g., inventory) depends on an unknown future outcome (e.g., demand). Typically, past data is collected to be indicative of the future, and hence inform our decision. However, ideal data-driven decision-making requires postulating beliefs about the reliability of the past data, and importantly, whether the future may *deviate* from it. In practice, past data may depend on contexts, some of which can be controlled for, and some of which cannot (unobserved confounders). This may introduce data heterogeneity that is not "correctable."

In this paper, we consider a framework for modeling how the future may deviate from past data. At a high level, it accomplishes three goals: $i.)$ capture different forms of future deviation, including no deviation, under the same umbrella; $ii.)$ understand the performance of a central policy, which makes the decision that optimizes the *average* objective value over past data, under different forms of unexpected deviation; $iii.)$ if the central policy performs poorly, then derive modifications that are robust to different forms of anticipated deviation, and quantify the achievable performance.

This framework is general, capturing different problems, and our analysis leads to insightful problem-specific conclusions. The only assumption is that the policy defined in $ii.)$ above can be computed, which requires perfect counterfactual evaluation of any decision on all data points, as opposed to settings where past data are affected by previous actions [5, 16, 33, 41, 42, 43]. The policy described in $ii.)$ is widely studied across different fields, known as[1] *Sample Average Approximation (SAA)* [37, 36, 39], although we emphasize that we also derive *new policies beyond SAA* when it falters.

[1] It is also known under different names, e.g., Empirical Optimization or Empirical Risk Minimization.

Submitted to 36th Conference on Neural Information Processing Systems (NeurIPS 2022). Do not distribute.

1.1 Framework description

Framework (Section 2.1). Our first contribution lies in studying a framework which models the future outcome as being drawn from an unknown distribution, and past data as being drawn independently from (also unknown) "nearby" distributions. In this way, past data is indicative of the future, and we emphasize that the samples of data can be drawn from different nearby distributions. In accord, we call this setting a *heterogeneous environment*. Using varying definitions of "nearby distribution" (e.g. based on Kolmogorov/Wasserstein distances; see Section 1.3), we analyze different forms of deviation possible between the past and future, and let a radius ϵ bound the allowed deviation. When $\epsilon = 0$, our framework captures the classical independent and identically distributed (i.i.d.) setting, in which past data are drawn from the same distribution as the future outcome [37, 45, 4]. This framework presents another possible parametrization to interpolate between the i.i.d. setting and the adversarial one and shares conceptually similar goals to previous approaches considering algorithm analysis beyond worst-case [51, 7, 53, 25, 10, 24].

Performance measure (Section 2.2). We consider the performance of different data-driven policies, which map past data into a decision for a given problem; SAA defines a feasible data-driven policy for any problem. Regret is measured as the difference in objective between the policy's decision and the optimal decision knowing the future distribution, taking an expectation over the draws of past data (which affect the policy's decision), any intrinsic randomness in the policy, and the outcome realization (which affects the objective of both the policy's decision and the optimal decision). We then take a worst case regret over all possible distributions that could have been chosen for the future outcome and the data points, to evaluate the performance of a fixed data-driven policy. We note that this performance depends on the problem (see Section 1.3 for examples), the number of data points n, the definition of "nearby distribution," and the radius ϵ.

1.2 General results

Our second contribution lies in developing a set of general results for this framework. A central result we present in Section 3 is a general reduction we develop in Theorem 1 in which we establish the following: asymptotically, as the number of samples goes to ∞, the worst-case regret of a policy (in a broad subclass we call sample-size-agnostic) is upper-bounded by that of an alternative problem with two major simplifications: $i.$) the worst-case is now taken over only a single distribution for historical environments as opposed to a sequence of heterogeneous distributions; and $ii.$) the decision-maker (DM) has access to the exact historical distribution. This upper bound can be seen as a uniform distributionally robust optimization problem under a regret performance metric, and offers a systematic way to derive upper bounds for policies of interest. This result is closely related to equation (7) in [46] and we discuss in detail the connections to their result after stating Theorem 1.

We develop problem-independent results that hold whenever the objective function is Lipschitz as a function of the future outcome that materializes (Section 4). For such environments, we establish in Theorem 2 that SAA scales linearly in ϵ for both the Kolmogorov and Wasserstein forms of deviation. We also show that for this class of problems, the Lipschitz constant characterizes the hardness of the instance.

1.3 Problem-specific, deviation-specific results

Our third contribution lies in deriving problem-specific conclusions in our general heterogeneous environments. In particular, we consider three classical problems in the context of our framework—newsvendor, pricing, and ski-rental—that have received significant attention in the literature across Operations Research, Economics, and Computer Science in the known-prior, data-driven, distributionally robust, adversarial, and advice-augmented settings. Problem-specific results are obtained by using: $i.$) our general reduction and results to upper-bound their asymptotic worst-case regret, and $ii.$) a combination of algorithm design and the development of impossibility results (lower bounds) on the worst-case regret. We analyze the regret for both the SAA policy as well as our robustified policies.

Problem descriptions. First, we consider the newsvendor problem (Section 4.1), in which a DM must decide how much capacity or inventory to plan in the face of unknown demand. The objective is to minimize the sum of *underage cost*, paid for each unit by which capacity falls short of realized demand, and *overage cost*, paid for each unit of capacity in excess of demand that needs to sit idle.

This is a foundational problem in Operations Research, with early focus on the setting where the demand distribution is known [see 47], and since then studied in various distributionally robust [54, 22, 49] settings but also in data-driven ones [40, 15, 6]. Newsvendor captures the fundamental tradeoff in guessing an unknown value when there are asymmetric consequences for going under/over.

Second, we consider the single-item pricing problem (Section 5.1), in which a DM must find the best price to offer to a customer with unknown willingness to pay (wtp). The objective is to maximize revenue. This is a foundational problem in Economics, with early focus on the setting with known prior [see 48, 52], and more recently in data-driven [21, 29, 2, 1, 17] and distributionally robust [12, 23] settings. Pricing captures the fundamental tradeoff between margin and volume.

Finally, we consider the ski-rental problem (Section 5.2), in which a DM must decide each day whether to buy skis or keep renting, without knowing the length of their trip. The objective is to minimize total cost, where buying eliminates the rental cost going forward. Although typically presented as an online problem that requires a decision to be made every day, any (randomized) buy-or-rent policy can be described by a (randomized) duration for which to rent skis, before committing to buying (if the trip has not ended by then). This is a foundational problem in Computer Science, critical to the development of competitive analysis [see 11] which considers the *adversarial* setting where nothing is known, and recently also important to the development of *advice-augmented algorithms* which receive a prediction on the length of the ski trip [50, 19]. Ski rental captures the fundamental tradeoff between proceeding with a suboptimal process vs. paying the cost of refactoring.

Definitions of "nearby distribution." Our framework generally models past data to be drawn from distributions that lie within an ϵ-distance of the future distribution, for *any* notion of distance between probability distributions. For our problems of interest, we focus on two specific distance metrics: Kolmogorov and Wasserstein (see Section 2.1 for formal definitions). This dichotomy suffices to demonstrate the importance of distance metric in determining performance in our problems of interest, and moreover, represents contrasting inductive philosophies about how the future might deviate from the past: Wasserstein allows *every* data point to be erroneous by ϵ, whereas Kolmogorov allows an ϵ fraction of data points to be *arbitrarily* erroneous. We note that similar notions of "nearby distributions" were considered in [51] in the online learning setting and in [7, Example 5] for the expert selection problem. Our work differs from these two papers because we consider an offline setting and our benchmark is stronger since we focus on the oracle knowing the ground-truth out-of-sample distribution, as opposed to the "best action in hindsight" benchmark.

SAA vs. robustified policies. Finally, performance is demarcated by the policy used. We first provide tight asymptotic performance bounds for the SAA policy, which we note is agnostic to the heterogeneity that might arise. We then analyze the best-possible asymptotic performance bounds that can be achieved by any policy when the notion of distance and the radius ϵ are known.

In Table 1, we present a high level summary of our results across problem classes. We assume that the support of the unknown environment is in $[0, M]$ for some positive real number M which M can be interpreted as parametrizing the "extent" of uncertainty one faces (e.g., maximal values of demand in newsvendor) We track the dependence on both the heterogeneity radius ϵ, as well as M. These results are for the asymptotic setting in which $n \to \infty$.

			Kolmogorov	Wasserstein
Lipschitz	Newsvendor (§4.1)	SAA	$\Theta(M\epsilon)$	$\Theta(\epsilon)$
		best policy	$\Theta(M\epsilon)$	$\Theta(\epsilon)$
Beyond Lipschitz	Pricing (§5.1)	SAA	$\Theta(M\epsilon)$	$\Omega(M)$
		best policy	$\Theta(M\epsilon)$	$\Theta(\sqrt{M\epsilon})$
	Ski-rental (§5.2)	SAA	$\Theta(M\epsilon)$	$\Theta(1)$
		best policy	$\tilde{\Theta}(\epsilon)^*$	$\Theta(\sqrt{\epsilon})$

Table 1: **SAA and achievable performance for different problem classes and heterogeneity balls.** *(Here $\tilde{\Theta}$ provides rate order up to logarathimic factors.)

For newsvendor problems, through a combination of the general result for Lipschitz-continuous problems and lower bounds, we show that SAA actually achieves the optimal rate for asymptotic worst-case regret both as a function of ϵ and M and that this rate is linear.

We then turn to problems whose objective is not Lipschitz. For the pricing problem, we establish that while under Kolmogorov distance, SAA is near-optimal and the optimal performance is again of order $M\epsilon$, the picture is starkly different under the Wasserstein distance. In this case, SAA has arbitrarily poor performance. We then show that a robustification procedure that deflates SAA by an appropriate factor can yield the best achievable asymptotic worst-case regret that shrinks to zero at rate $\sqrt{M\epsilon}$. To our understanding, analyzing these non-SAA policies (which are required for good performance on pricing) deviates from standard analyses used in learning theory, as we elaborate on in Remark 2.

A second prototypical problem that is not Lipschitz is the ski-rental problem. While SAA was near-optimal for the Kolmogorov distance for pricing and newsvendor, we establish the surprising fact that SAA, even under the Kolmogorov distance, has arbitrarily poor performance. Indeed, its asymptotic worst-case regret scales as $\Theta\left(M\epsilon\right)$, while the optimal policy achieves a performance of $\Theta\left(\epsilon\right)$. As a consequence, the performance of SAA can be arbitrarily worse than the optimal one as the support grows to infinity. Meanwhile, for the Wasserstein distance, we show that the performance of SAA is also suboptimal. Indeed, SAA incurs a regret of $\Theta(1)$; but by inflating SAA appropriately, one can construct a policy that achieves order $\sqrt{\epsilon}$ performance, which is the best dependence on ϵ. These examples demonstrate that SAA fails for both heterogeneity types for ski-rental, but also more broadly the intricate interplay between problem class and type of heterogeneity.

Connection to offline data corruption models. Different notions of offline data corruptions and robustness have been previously studied in statistics [30, 26], learning theory [55, 28, 35, 34, 55, 38, 18, 8, 20] and, more recently in revenue management [13, 12, 14, 23]. Most approaches consider ϵ-contamination settings [30] in which the adversary can shift a small proportion of data arbitrarily far from the true distribution. They consider either oblivious adversaries who fix a *single* corrupted distribution from which samples are generated or adaptive one who observe samples before corrupting them. [8] establishes that adaptive adversaries yield equivalent performance to oblivious ones in many settings. Closest to us are [12, 23] who study optimal auctions. They assume the distribution is common across all past observations; the former focuses on regret and the the latter on a ratio metric when data is generated from MHR distributions close in Kolmogorov distance. The framework we develop is anchored around a setting that allows for *heterogeneous* "nearby" distributions for each past observation, and is general in that it allows us to unify a variety of problems/metrics and highlight how these affect the levels of achievable performance.

Connection to online and batch learning in non-stationary environments. Our work also relate broadly to papers studying learnability in changing environment. [27, 9] adopt a universal learning approach and analyze algorithms which are able to achieve vanishing long-term average loss for general data-generation processes. [44, 3, 46] derive generalization guarantees in a setting where the training set distribution may differ from the test set distribution.

2 Problem formulation: data-driven decisions in heterogeneous environments

We consider a general stochastic optimization problem in which a DM wants to optimize an objective in a stochastic environment. Formally, we denote by $\mathcal{X}$ the decision space and by Ξ the space of environment realizations. The objective of the DM is represented by a function $g : \mathcal{X} \times \Xi \to \mathbb{R}$ which maps each pair of decision $x \in \mathcal{X}$ and environment realization $\xi \in \Xi$ to an objective value.

The DM faces a subset of possible probability measures $\mathcal{P} \subset \Delta\left(\Xi\right)$, where $\Delta\left(\Xi\right)$ denotes the space of probability measures supported on Ξ. For every probability measure $\mu \in \mathcal{P}$, the goal of the DM is to optimize the objective defined as

$$\mathcal{G}\left(\cdot, \mu\right) : \begin{cases} \mathcal{X} \to \mathbb{R} \\ x \mapsto \mathbb{E}_{\xi \sim \mu}\left[g(x, \xi)\right]. \end{cases} \tag{1}$$

We extend the definition of $\mathcal{G}\left(\cdot, \mu\right)$ to allow for randomized decisions $\rho \in \Delta\left(\mathcal{X}\right)$, in which case we define $\mathcal{G}\left(\rho, \mu\right) = \mathbb{E}_{x \sim \rho}\left[\mathcal{G}\left(x, \mu\right)\right]$ to also take an expectation over the randomized decision.

Remark that depending on the nature of the problem, the objective could represent for example a function to maximize (profit) or to minimize (cost). In turn, we formally define for every $\mu \in \mathcal{P}$, the optimal objective $\text{opt}(\mu)$ as the objective value achieved by an *optimal action*, given by an *oracle* operator `ORACLE` that is a mapping from $\Delta\left(\Xi\right)$ to $\mathcal{X}$.

2.1 Heterogeneous environments

In the absence of distributional knowledge on the underlying environment, a common approach consists in assuming that historical samples are independent and identically drawn from a fixed probability measure μ. We relax the assumption that the samples are identically distributed and introduce a generalization of this framework that we refer to as *ϵ-heterogeneity*.

Let $d : \mathcal{P} \times \mathcal{P} \to \mathbb{R}$ be a metric on $\mathcal{P}$. Given a target probability measure $\mu \in \mathcal{P}$ and a level of heterogeneity ϵ, we define the heterogeneity ball anchored around μ and with radius ϵ as

$$\mathcal{U}_\epsilon(\mu) := \{\nu \in \mathcal{P},\ \text{s.t}\ d(\mu, \nu) \leq \epsilon\}.$$

The heterogeneity ball $\mathcal{U}_\epsilon(\mu)$ includes all distributions that are within a distance of ϵ from μ. Next, we formally define a data-driven decision-making problem in a heterogeneous environment.

Definition 1 (Data-driven decision-making problem in heterogeneous environment)**.** *A data-driven decision-making problem in a heterogeneous environment is defined by the tuple: $(\mathcal{X}, \Xi, \mathcal{P}, d, g, \texttt{ORACLE})$, where $\mathcal{X}$ is the set of possible actions, Ξ is the environment space, $\mathcal{P}$ is the space of measures nature can select from, d is a metric on the space of measures, g is an objective function, and* ORACLE *is a mapping from environment measures to target actions.*

Focal metrics. The definition above is general, and the framework can be applied to a variety of metrics. In this work, we aim at developing an understanding of the effect of the extent and the type of heterogeneity in various problem classes. We focus on two prototypical and central distances: *Kolmogorov* and *Wasserstein*. We formally define them in one dimension as follows.

Definition 2 (Kolmogorov and Wasserstein distances)**.** *Given an environment space $\Xi \subset \mathbb{R}$ and a subset of probability measures $\mathcal{P}$ supported on Ξ, the Kolmogorov (resp. Wasserstein) distance d_K (resp. d_W) is defined for every $\mu_1, \mu_2 \in \mathcal{P}$ as*

$$d_K(\mu_1, \mu_2) = \sup_{\xi \in \Xi} |F_1(\xi) - F_2(\xi)| \qquad \textit{and} \qquad d_W(\mu_1, \mu_2) = \int_{\xi \in \Xi} |F_1(\xi) - F_2(\xi)|\, d\xi,$$

where F_1 (resp. F_2) is the cumulative distribution associated with μ_1 (resp. μ_2).

Note that when Ξ is a compact interval of $\mathbb{R}$, both distances are related as for every μ_1, $\mu_2 \in \mathcal{P}$,

$$d_W(\mu_1, \mu_2) \leq \operatorname{diam} \Xi \cdot d_K(\mu_1, \mu_2). \tag{2}$$

2.2 Data-driven policies and performance

Data-driven policies. In most settings, the probability measure μ is unknown to the DM and the optimal objective $\text{opt}(\mu)$ is not achievable. In turn, the DM only observes a sequence of historical samples $\xi_1, \xi_2, \ldots, \xi_n$ representing the previous environment realizations they observed. A data-driven policy is then a mapping from the past samples to a (randomized) decision. Define the empirical probability measure $\hat{\mu}_{\boldsymbol{\xi}^n}$ as

$$\hat{\mu}_{\boldsymbol{\xi}^n}(\xi) := \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}\{\xi = \xi_i\}, \qquad \text{for every } \xi \in \Xi.$$

In a setting in which the order of the samples does not matter (as the one we will consider), to represent a policy taking as input $\boldsymbol{\xi}^n := (\xi_1, \ldots, \xi_n)$, one only needs to consider mappings that take as an input the number of samples and the empirical measure. We define a data-driven policy π as a mapping from $\mathbb{N}^* \times \mathcal{P}$ to $\Delta(\mathcal{X})$, the space of measures on the action space $\mathcal{X}$, that associates $(n, \hat{\mu})$ to a distribution of actions $\pi(n, \hat{\mu})$. We let Π denote the set of all such mappings. We assume that such a mapping can be made with knowledge of the metric d and radius ϵ. Moreover, we say that π is *sample-size-agnostic* if its decision does not depend on n, in which case we write $\pi(\hat{\mu})$ for brevity.

A widely-studied policy that achieves excellent performance across a vast range of data-driven decision-making problems is Sample Average Approximation (SAA). Given an empirical distribution $\hat{\mu}$, SAA (denoted by π^{SAA}) selects the target action $\texttt{ORACLE}(\hat{\mu})$, the best action associated with $\hat{\mu}$.

Performance through worst-case regret. Given an out-of-sample distribution $\mu \in \mathcal{P}$ the best achievable objective is $\text{opt}(\mu)$. For a sequence of historical distributions $\nu_1, \ldots, \nu_n$ in $\mathcal{U}_\epsilon(\mu)$, the expected objective of a policy π is

$$\mathbb{E}_{\xi_i \sim \nu_i}\left[\mathcal{G}\left(\pi(n, \hat{\mu}_{\boldsymbol{\xi}^n}), \mu\right)\right].$$

In turn, we define the expected (absolute) regret of a policy $\pi \in \Pi$ against an out-of-sample distribution $\mu \in \mathcal{P}$ and a sequence of n historical distributions $\nu_1, \ldots, \nu_n \in \mathcal{U}_\epsilon(\mu)$ as

$$\mathcal{R}_n\left(\pi, \mu, \nu_1, \ldots, \nu_n\right) := \left|\mathbb{E}_{\xi_i \sim \nu_i}\left[\mathcal{G}\left(\pi(n, \hat{\mu}_{\boldsymbol{\xi}^n}), \mu\right)\right] - \text{opt}(\mu)\right|.$$

We note that we use the absolute value in this definition, as our framework applies to minimization and maximization problem. Once the type of problem is specified (i.e., the operator associated with ORACLE), this definition of regret coincides with the usual regret for minimization or maximization problems, as we will see in Sections 4 and 5, when describing specific applications.

In this work, we will assess the performance of a policy based on its *worst-case expected regret*. Given an instance $\mathcal{I} = (\mathcal{X}, \Xi, \mathcal{P}, d, g, \texttt{ORACLE})$, we define the worst-case expected regret of a data-driven policy $\pi \in \Pi$ for a degree of heterogeneity $\epsilon \geq 0$ as

$$\mathfrak{R}^{\pi}_{\mathcal{I},n}(\epsilon) := \sup_{\mu \in \mathcal{P}} \sup_{\nu_1, \ldots, \nu_n \in \mathcal{U}_\epsilon(\mu)} \mathcal{R}_n\left(\pi, \mu, \nu_1, \ldots, \nu_n\right). \tag{3}$$

We finally define the *worst-case asymptotic regret* of a data-driven policy $\pi \in \Pi$ in a ϵ-heterogeneous environment as the number of samples n grows large as

$$\mathfrak{R}^{\pi}_{\mathcal{I},\infty}(\epsilon) := \limsup_{n \to \infty} \mathfrak{R}^{\pi}_{\mathcal{I},n}(\epsilon). \tag{4}$$

We sometimes write $\mathfrak{R}^{\pi}_{K,\infty}(\epsilon)$ or $\mathfrak{R}^{\pi}_{W,\infty}(\epsilon)$ instead of $\mathfrak{R}^{\pi}_{\mathcal{I},\infty}(\epsilon)$ when the problem class is specified to highlight the dependence on the Kolmogorov or Wasserstein metrics.

3 Reduction to distributionally robust optimization in the asymptotic regime

A significant challenge in the analysis of problem (4) stems from the difficulty to characterize a potentially complex worst-case sequence of historical probability measures $\nu_1, \ldots, \nu_n$ and out-of-sample measure μ. We illustrate in the appendix that analyzing every element in the sequence of historical probability measures as a function of n is a priori challenging as nature will attempt to use different distributions for each of the past environments. We will show that, asymptotically, it suffices to consider cases in which nature selects a common distribution for all of the past environments. Formally, we define for a sample-size-agnostic policy $\pi \in \Pi$ the following optimization problem:

$$\sup_{\mu \in \mathcal{P}} \sup_{\nu \in \mathcal{U}_\epsilon(\mu)} \left|\text{opt}(\mu) - \mathcal{G}\left(\pi(\nu), \mu\right)\right|. \tag{5}$$

While Problem (5) resembles Problem (3), it differs on two crucial dimensions: $i.)$ in Problem (5), nature selects the *same* distribution ν in the heterogeneity ball for all past environments; $ii.)$ in Problem (5), the policy π is assumed to know the actual distribution of the past environments ν, whereas in Problem (3), the policy only observed the empirical distribution of past realizations.

Next, we will establish that Problems (5) and (3) are tightly connected under some mild assumptions. We introduce the following definition on the distance d defining the heterogeneity.

Definition 3 (Empirical triangular convergence)**.** *We say that a distance d on the space of probability measures satisfies the empirical triangular convergence (ETC) property if and only if for every triangular array sequence of probability measures $(\mu_{i,n})_{1 \leq i \leq n, n \in \mathbb{N}^*}$ all belonging to $\mathcal{P}$, we have*

$$\lim_{n \to \infty} d(\hat{\mu}_n, \bar{\mu}_n) = 0 \qquad a.s.,$$

where $\hat{\mu}_n(\xi) := \frac{1}{n}\sum_{i=1}^n \mathbb{1}\left\{\xi = \xi_{i,n}\right\}$ for samples $\xi_{i,n} \sim \mu_{i,n}$ and $\bar{\mu}_n = \frac{1}{n}\sum_{i=1}^n \mu_{i,n}$.

The ETC property requires that the sequence of empirical distributions converges to the average of ground truth distributions for arbitrary triangular arrays. In what follows we also impose convexity on the metric d. We note that both properties are satisfied by many common distances on probability spaces, including the Kolmogorov and Wasserstein distances on compact sets, as we discuss in the appendix. We now state our first main result which establishes a fundamental reduction.

Theorem 1 (Upper bound reduction)**.** *Let $\mathcal{I} = (\mathcal{X}, \Xi, \mathcal{P}, d, g, \texttt{ORACLE})$ be a data-driven decision problem in a heterogeneous environment and let $\pi \in \Pi$ be a sample-size-agnostic policy. Assume that g is bounded on $\mathcal{X} \times \Xi$ and that d is a convex metric which satisfies the empirical triangular convergence property. Then*

$$\mathfrak{R}^{\pi}_{\mathcal{I},\infty}(\epsilon) \leq \lim_{\eta \to \epsilon^+} \sup_{\mu \in \mathcal{P}} \sup_{\nu \in \mathcal{U}_\eta(\mu)} \left|opt(\mu) - \mathcal{G}\left(\pi(\nu), \mu\right)\right|. \tag{6}$$

Theorem 1 establishes, that under some mild conditions, in the asymptotic regime, one may restrict attention to problem (5). This reduction significantly simplifies the analysis of sample-size-agnostic policies[2] as problem (5) does not involve any form of randomness in the input received by the policy, and only allows the adversary to select a single distribution.

Remark 1 (Relation to [46])**.** *We note that [46, eq. (7)] presents a finite-time bound on the regret of SAA in a similar setting to ours. Therefore by taking the limit in their bound, one may derive an asymptotic performance guarantee for SAA. In our notation, their result imply that,*

$$\mathfrak{R}^{\pi^{SAA}}_{\mathcal{I},\infty} \leq 2 \cdot \sup_{\mu \in \mathcal{P}} \sup_{\nu \in \mathcal{U}_\epsilon(\mu)} \sup_{x \in \mathcal{X}} |\mathcal{G}(x,\mu) - \mathcal{G}(x,\nu)|. \tag{7}$$

Theorem 1 differs from (7) *because our reduction applies to any policy in the wide range of sample-size-agnostic policies. In particular, this enables us to characterize the minimax optimal asymptotic regret even when SAA is not (rate) optimal (see Sections 5.1 and 5.2).*

We also show in the appendix that Theorem 1 holds whenever Ξ is finite, for any distance metric, even without a sample-size-agnostic policy.

4 Lipschitz-continuous problems

In this section, we study a broad class of problems for which the underlying objective function $g(x,\cdot)$ is Lipschitz-continuous as a function of the environment.

Theorem 2 (Upper bounds for Lipschitz-continuous objectives)**.** *Suppose that $\Xi = [0, M]$ for some $M > 0$. Let $\mathcal{I}_K = (\mathcal{X}, \Xi, \mathcal{P}, d_K, g, \texttt{ORACLE})$ and $\mathcal{I}_W = (\mathcal{X}, \Xi, \mathcal{P}, d_W, g, \texttt{ORACLE})$ be instances of the data-driven decision problem under the Kolmogorov and Wasserstein distances, respectively. Assume that for every $x \in \mathcal{X}$, the function $g(x,\cdot)$ is L-Lipschitz-continuous and that* ORACLE *is either a minimization or maximization oracle. Then for every $\epsilon \geq 0$,*

$$\mathfrak{R}^{\pi^{SAA}}_{\mathcal{I}_K,\infty}(\epsilon) \leq 2L \cdot M \cdot \epsilon \qquad \textit{and} \qquad \mathfrak{R}^{\pi^{SAA}}_{\mathcal{I}_W,\infty}(\epsilon) \leq 2L \cdot \epsilon.$$

Theorem 2 implies that for this class of problems, the asymptotic regret of SAA vanishes as the radius of the heterogeneity ball goes to 0. Furthermore, this theorem shows that the dependence of the asymptotic worst-case regret in the heterogeneity parameter ϵ is at most linear. We show next that this dependence is the best possible for this class of problems and this notion of heterogeneity.

4.1 Newsvendor problem

Recall the newsvendor problem described in Section 1.3.Using $M > 0$ to denote an upper bound on demand, we let $\Xi = \mathcal{X} = [0, M]$, and our regret bounds may (or may not) depend on M. The objective function is the newsvendor cost which is parameterized by two quantities: c_u the underage cost and c_o the overage cost. The cost of carrying an inventory level $x \in \mathcal{X}$ and then observing demand $\xi \in \Xi$ is,

$$g(x,\xi) = c_u (\xi - x)^+ + c_o (x - \xi)^+ ,$$

where $(\cdot)^+ := \max\{\cdot, 0\}$ is the positive part operator. The measure space $\mathcal{P}$ is the set of probability measures on Ξ without restrictions and for every $\mu \in \mathcal{P}$, we define the cost of the oracle as $\text{opt}(\mu) = \min_{x \in \mathcal{X}} \mathcal{G}(x,\mu)$. Therefore the oracle mapping is defined as $\texttt{ORACLE} : \mu \mapsto \arg\min_{x \in \mathcal{X}} \mathcal{G}(x,\mu)$.

We note that the objective function g satisfies the L-Lipschitz-continuity property on the environment realization variable ξ with $L = \max(c_u, c_o)$. Theorem 2 thus directly yields an upper bound on the achievable regret for this class of problems, for both Kolmogorov and Wasserstein, through the regret of SAA. We next show that these dependencies are the best possible, under any policy.

Proposition 1 (Heterogeneous newsvendor)**.** *For the data-driven newsvendor problem in a heterogenous environment under the Kolmogorov and Wasserstein distances, given any $\epsilon \geq 0$,*

$$\begin{aligned} \frac{c_u + c_o}{2} \cdot \epsilon \cdot M \quad &\leq \inf_{\pi \in \Pi} \mathfrak{R}^{\pi}_{K,\infty}(\epsilon) \leq \mathfrak{R}^{\pi^{SAA}}_{K,\infty}(\epsilon) \leq \quad 2\max(c_u, c_o) \cdot \epsilon \cdot M, \\ \frac{c_u + c_o}{2} \cdot \epsilon \quad &\leq \inf_{\pi \in \Pi} \mathfrak{R}^{\pi}_{W,\infty}(\epsilon) \leq \mathfrak{R}^{\pi^{SAA}}_{W,\infty}(\epsilon) \leq \quad 2\max(c_u, c_o) \cdot \epsilon. \end{aligned}$$

[2]Note that considering sample-size-agnostic policies is not very restrictive. As we will see in Section 4 and 5 this class is sufficient to achieve optimal rates for the asymptotic worst-case regret as a function of ϵ.

Proposition 1 first shows that SAA is robust to deviation under multiple forms of heterogeneity and that the regret scales linearly with the radius of the heterogeneity ball for both distances. We show furthermore that the linear dependence is, for the newsvendor problem, the best rate achievable for any data-driven policy. We also note that the lower bound obtained for the newsvendor depends linearly on the Lipschitz constant L through the factor $(c_u + c_o)$.

Beyond newsvendor, the asymptotic worst-case regret of SAA over the general class of L-Lipschitz-continuous functions is of the order $\Theta(L \cdot \epsilon)$. This suggests that the performance of SAA may deteriorate for problems in which the objective function g is not smooth.

5 Beyond Lipschitz-continuous problems

5.1 Pricing problem

Recall the pricing problem presented in Section 1.3. The environment space $\Xi = [0, M]$ represents the set of possible wtp. Similarly, the set of possible prices $\mathcal{X}$ is $[0, M]$. The set of considered probability measures $\mathcal{P}$ is the whole set of probability measures on $[0, M]$ without restriction. The revenue generated by the DM with a price decision $x \in \mathcal{X}$ when facing a customer with wtp $\xi \in \Xi$ is,

$$g(x, \xi) = x \cdot \mathbb{1}\{x \leq \xi\}.$$

Therefore for any probability measure $\mu \in \mathcal{P}$ associated with the cumulative distribution F, we have that $\mathcal{G}(x, \mu) = x \cdot (1 - F(x))$. The goal of the DM is to set a price x in order to maximize its revenue. Equivalently, the goal is to minimize the regret against the oracle which posts the optimal price given a wtp distribution. Formally, the oracle operator is defined as, $\texttt{ORACLE} : \mu \mapsto \arg\max_{x \in \mathcal{X}} \mathcal{G}(x, \mu)$.

Kolmogorov pricing. Remark that the pricing objective is not Lipschitz-continuous and that one cannot use the argument presented in Section 4.1. We show in the appendix that despite the lack of smoothness of the objective, SAA is still near-optimal in the asymptotic regime under the Kolmogorov distance. We derive an upper bound on the asymptotic worst-case regret of SAA and a matching universal lower bound showing that the best achievable performance is order $\Theta(M\epsilon)$.

Failure of SAA in pricing with Wasserstein heterogeneity. In opposition to the strong performance achieved by SAA for pricing against Kolmogorov heterogeneity, our next proposition shows that this result does not hold if one considers instead the Wasserstein heterogeneity ball.

Proposition 2 (SAA and Wasserstein pricing)**.** *For the data-driven pricing problem in a heterogenous environment under the Wasserstein distance, we have for $\epsilon > 0$ that*

$$\mathfrak{R}^{\pi^{SAA}}_{W,\infty}(\epsilon) = M.$$

Proposition 2 shows that, in stark contrast with Lipschitz problems, for the pricing problem, the asymptotic regret of SAA under Wasserstein heterogeneity does not shrink to zero as ϵ goes to zero. As a matter of fact, the regret of M is the worst possible, as the revenue of the oracle is bounded from above by M and the revenue of any data-driven policy is bounded from below by 0. Quite notably, this result holds for any level of heterogeneity ϵ, and an arbitrarily small deviation from the i.i.d. case leads to extremely poor performance, even with infinite data. We next propose a policy which "robustifies" SAA and we analyze it under the Wasserstein distance.

Robustification of SAA for Wasserstein pricing. The key challenge in pricing is that for two distributions μ and ν such that $d_W(\mu, \nu)$ is arbitrarily small, a price x could have a high revenue for μ but a low revenue for ν, and the difference in revenue between distributions may be arbitrarily large. This enabled us to construct an instance for which the optimal price for the in-sample distribution ν generates a low revenue for the out-of-sample distribution μ.

Remark 2 (Proof Sketch, Comparison with Standard Techniques)**.** *To the best of our understanding, our analysis of policies beyond SAA requires a new technique. Assume $M = 1$ for brevity, and consider the problem of pricing. We need to upper-bound the loss $\mathcal{G}(\texttt{ORACLE}(\mu), \mu) - \mathcal{G}(\pi(\nu), \mu)$. A standard analysis of SAA, which satisfies $\pi^{SAA}(\nu) = \texttt{ORACLE}(\nu)$, would write,*

$$\mathcal{G}(\texttt{ORACLE}(\mu), \mu) - \mathcal{G}(\texttt{ORACLE}(\nu), \mu) \leq 2 \sup_{x \in \mathcal{X}} |\mathcal{G}(x, \mu) - \mathcal{G}(x, \nu)|.$$

However, the bug with this approach is that as described above, for certain decisions x, the error $|\mathcal{G}(x, \mu) - \mathcal{G}(x, \nu)|$ can be non-vanishing even if the distance $d_W(\mu, \nu)$ is vanishing. Our fix is the

following. For a generic policy π, we write

$$\mathcal{G}(\texttt{ORACLE}\,(\mu)\,,\mu) - \mathcal{G}(\pi(\nu),\mu) \leq (\mathcal{G}(\texttt{ORACLE}\,(\mu)\,,\mu) - \mathcal{G}(\pi(\mu),\nu)) + (\mathcal{G}(\texttt{ORACLE}\,(\nu)\,,\nu) - \mathcal{G}(\pi(\nu),\mu)) \tag{8}$$

The key observation is that if we set $\pi(\nu)$ not to $\texttt{ORACLE}\,(\nu)$, but to $\texttt{ORACLE}\,(\nu) - \delta$, then both large parentheses in (8) can be (identically) upper-bounded using the relation we introduce in Proposition 3.

Proposition 3. *Let μ and $\nu \in \mathcal{P}$ be two different distributions and let x_1 and x_2 be two different prices such that $x_1 < x_2$. We have that,*

$$\mathcal{G}\,(x_2,\nu) - \mathcal{G}\,(x_1,\mu) \leq (x_2 - x_1) + M \cdot \frac{d_W\,(\mu,\nu)}{x_2 - x_1}.$$

Proposition 3 unveils an interesting tradeoff faced by a DM who prices in a heterogeneous environment. By selecting appropriately the difference between two prices x_1 and x_2, one can ensure that the gap in revenues for two different distributions can be controlled by the Wasserstein distance.

We next restrict attention to the class of policies which additively *deviate* from the action selected by SAA and leverage Proposition 3 to define the correct deviation parameter.

Definition 4 (δ-SAA policies)**.** *For any one-dimensional problem we say that a policy is a δ-SAA policy if for every empirical distribution $\hat{\mu}$, the policy selects the closest action in $\mathcal{X}$ to $\texttt{ORACLE}\,(\hat{\mu})+\delta$, for some $\delta \in \mathbb{R}$ that could be positive or negative. We denote this policy by $\pi^{SAA(\delta)}$.*

We note that when $\delta = 0$, we recover the usual SAA policy. Our next result establishes the best possible regret dependence on ϵ and M for the class of pricing problems, using a δ-SAA policy.

Theorem 3 (Deviation for Wasserstein pricing)**.** *For the data-driven pricing problem in a heterogenous environment under the Wasserstein distance, for any $\epsilon > 0$, let $\tilde{\delta} = -\sqrt{M \cdot \epsilon}$. Then,*

$$\frac{1}{4} \cdot \sqrt{M \cdot \epsilon} \leq \inf_{\pi \in \Pi} \mathfrak{R}^{\pi}_{W,\infty}(\epsilon) \leq \mathfrak{R}^{\pi^{SAA(\tilde{\delta})}}_{W,\infty}(\epsilon) \leq 4 \cdot \sqrt{M \cdot \epsilon}.$$

Theorem 3 shows that by appropriately deflating the price posted by SAA, one is able to be robust against Wasserstein heterogeneity and achieve a worst-case asymptotic regret that vanishes to 0 as the degree of heterogeneity ϵ goes to 0. This proves that one may operate efficiently for pricing under Wasserstein heterogeneity as ϵ goes to 0. Our result establishes that in contrast with Lipschitz problems, it is now impossible to achieve a linear dependence in ϵ (since $\sqrt{\epsilon} > \epsilon$), with the regret growing at a rate of $\Omega(\sqrt{M \cdot \epsilon})$ under any policy. A small radius ϵ has a significantly higher impact in pricing problems than in, e.g., newsvendor problems.

We note that the deflated policy proposed in Theorem 3 improves over SAA when one can make use of the knowledge of ϵ. It is worth noting that SAA does not need such knowledge. Furthermore, one can show that this deflated policy incurs a $\Omega\left(\sqrt{\epsilon}\right)$ regret for pricing under Kolmogorov distance and thus performs poorly compared to SAA in that setting. A natural question would be to understand the best achievable performance without knowledge of ϵ and/or d.

5.2 Ski-rental problem

We now consider the ski-rental problem in which renting skis costs \$1 per unit of time while buying them costs \$$b$ up-front, for some real value b. The environment space Ξ represents the set of possible lengths of the ski trip and a decision x represents the duration after which skis should be bought (if the ski trip has not ended by that time). We call x the *rental duration*. Let $\mathcal{X} = \Xi = [0, M]$, where we note that setting $x = M$ indicates that skis should never be bought. The set of probability measures $\mathcal{P}$ is the whole space of probability measures supported on $[0, M]$. For every rental duration $x \in \mathcal{X}$ and any trip length $\xi \in \Xi$, the cost incurred is

$$g(x,\xi) = \xi \cdot \mathbb{1}\{\xi \leq x\} + (b + x) \cdot \mathbb{1}\{\xi > x\}\,.$$

Finally, as the goal is to minimize cost we have that, $\texttt{ORACLE} : \mu \mapsto \arg\min_{x \in \mathcal{X}} \mathcal{G}\,(x,\mu)\,.$

Wasserstein ski-rental. We proved in Section 5.1 that SAA performs arbitrarily poorly in pricing under Wasserstein heterogeneity. We now show that SAA falters similarly for the Wasserstein ski-rental problem. Furthermore, we design and analyze a δ-SAA policy which inflates the action selected by SAA (i.e., $\delta > 0$) and achieves the best rate possible as a function of the heterogeneity radius. Formally, we show the following.

Theorem 4. *For the data-driven ski-rental problem in a heterogenous environment under the Wasserstein distance, we have for $\epsilon > 0$ that,*

$$\mathfrak{R}^{\pi^{SAA}}_{W,\infty}(\epsilon) = \Theta\left(1\right) \qquad \textit{and} \qquad \inf_{\pi\in\Pi} \mathfrak{R}^{\pi}_{W,\infty}(\epsilon) = \Theta\left(\sqrt{\epsilon}\right).$$

Theorem 4 formalizes the failure of SAA for ski-rental under Wasserstein distance. A notable fact about the ski-rental problem under Wasserstein heterogeneity is that both the regret of SAA and of the optimal data-driven decision do not scale with the size of the support. For pricing, SAA scaled linearly with M and the optimal policy scales in $\sqrt{M}$. The fact that $\mathfrak{R}^{\pi^{\text{SAA}}}_{W,\infty}(\epsilon) = O(1)$ in ski-rental requires a separate non-trivial proof in Theorem 4.[3]

Kolmogorov ski-rental. Under the Kolmogorov distance, we had seen that SAA was near-optimal for both the newsvendor and pricing problems. We establish next that the performance of SAA for the ski-rental problem under Kolmogorov distance is, surprisingly, highly suboptimal.

Theorem 5. *For the data-driven ski-rental problem in a heterogenous environment under the Kolmogorov distance, we have for $\epsilon > 0$ that,*

$$\mathfrak{R}^{\pi^{SAA}}_{K,\infty}(\epsilon) = \Theta\left(M\cdot\epsilon\right) \quad \textit{and} \quad \inf_{\pi\in\Pi} \mathfrak{R}^{\pi}_{K,\infty}(\epsilon) = \tilde{\Theta}\left(\epsilon\right),$$

where $\tilde{\Theta}$ provides rates order up to logarithmic factors.

Theorem 5 shows that the asymptotic worst-case regret of SAA scales linearly with the radius of heterogeneity ϵ and with the size of the support M. We also characterize the rate of the optimal asymptotic worst-case regret as a function of ϵ. The upper bound on the asymptotic regret is obtained in [19] through a variant of the SAA policy, which caps the maximum number of days to rent. We complement this result by providing a matching lower bound. Note that, in contrast to the pricing and newsvendor problems, the scaling of SAA is not optimal in M. As the scale of the support grows, the asymptotic worst-case regret of SAA is considerably worse than the optimal achievable rate.

6 Conclusion

All in all, the present results offer a systematic way of understanding and quantifying the implications of operating in heterogeneous environments. Our framework enables us to develop a common language to compare achievable performance across central classes of problems and to unveil novel insights about the performance of a central policy, Sample Average Approximation, when slightly deviating from the widely studied i.i.d. regime. In settings where SAA fails, we also design robustification procedures achieving rate-optimal asymptotic guarantees. A key takeaway of this analysis across a broad class of problems and for different heterogeneity structures is that it is necessary to understand the structure of the problem we are facing but also the nature of the heterogeneity when designing data-driven policies that are robust to these environments.

To derive our performance guarantees, we established a crucial connection between data-driven decision making in heterogeneous environments and distributionally robust optimization. While this work leverages this connection essentially to derive bounds on achievable performances, we believe that this connection may be of interest to develop new policies from first principles.

This work also opens many additional questions. First, our analysis characterizes the performance in the asymptotic regime where the sample size grows large, and could be complemented by quantifying the performance of policies with finite samples in heterogeneous environments. Furthermore, our results for Lipschitz problems pin down a key driver of performance, whereas isolating the properties of the policies and the framework elements that drive different levels of achievable performance beyond the Lipschitz case remains open. Another key question would be to understand whether there exists a "best of both worlds" policy that does not use the knowledge of the type of heterogeneity and performs well across heterogeneity types (as opposed to the robustified policy designed for pricing and ski-rental) . Finally, we believe that additional exciting and practical complement to this work include incorporating contexts and deriving statistical tests to characterize the type of heterogeneity along with its radius and provide an empirical validation of the procedures developed here.

[3] The fact that $\mathfrak{R}^{\pi^{\text{SAA}}}_{W,\infty}(\epsilon) \leq M$ in pricing was vacuous.

References

[1] Allouah, A., Bahamou, A. and Besbes, O. [2022], 'Pricing with samples', *Operations Research* **70**(2), 1088–1104.

[2] Babaioff, M., Gonczarowski, Y. A., Mansour, Y. and Moran, S. [2018], Are two (samples) really better than one?, *in* 'Proceedings of the 2018 ACM Conference on Economics and Computation', pp. 175–175.

[3] Ben-David, S., Blitzer, J., Crammer, K., Kulesza, A., Pereira, F. and Vaughan, J. W. [2010], 'A theory of learning from different domains', *Machine learning* **79**(1), 151–175.

[4] Bertsimas, D., Gupta, V. and Kallus, N. [2018], 'Robust sample average approximation', *Mathematical Programming* **171**(1), 217–282.

[5] Besbes, O., Gur, Y. and Zeevi, A. [2014], 'Stochastic multi-armed-bandit problem with non-stationary rewards', *Advances in neural information processing systems* **27**.

[6] Besbes, O. and Mouchtaki, O. [2021], 'How big should your data really be? data-driven newsvendor: learning one sample at a time', *working paper, Columbia University* .

[7] Bilodeau, B., Negrea, J. and Roy, D. M. [2020], 'Relaxing the iid assumption: Adaptively minimax optimal regret via root-entropic regularization', *arXiv preprint arXiv:2007.06552* .

[8] Blanc, G., Lange, J., Malik, A. and Tan, L.-Y. [2021], 'On the power of adaptivity in statistical adversaries', *arXiv preprint arXiv:2111.10352* .

[9] Blanchard, M. [2022], 'Universal online learning: An optimistically universal learning rule', *arXiv preprint arXiv:2201.05947* .

[10] Block, A., Dagan, Y., Golowich, N. and Rakhlin, A. [2022], 'Smoothed online learning is as easy as statistical learning', *arXiv preprint arXiv:2202.04690* .

[11] Borodin, A. and El-Yaniv, R. [2005], *Online computation and competitive analysis*, cambridge university press.

[12] Brustle, J., Cai, Y. and Daskalakis, C. [2020], Multi-item mechanisms without item-independence: Learnability via robustness, *in* 'Proceedings of the 21st ACM Conference on Economics and Computation', pp. 715–761.

[13] Cai, Y. and Daskalakis, C. [2017], Learning multi-item auctions with (or without) samples, *in* '2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)', IEEE, pp. 516–527.

[14] Chen, X. and Wang, Y. [2020], 'Dynamic pricing with demand learning in the presence of outlier customers', *https://ssrn.com/abstract=3650656* .

[15] Cheung, W. C. and Simchi-Levi, D. [2019], 'Sampling-based approximation schemes for capacitated stochastic inventory control models', *Mathematics of Operations Research* **44**(2), 668–692.

[16] Cheung, W. C., Simchi-Levi, D. and Zhu, R. [2022], 'Hedging the drift: Learning to optimize under nonstationarity', *Management Science* **68**(3), 1696–1713.

[17] Daskalakis, C. and Zampetakis, M. [2020], More revenue from two samples via factor revealing sdps, *in* 'Proceedings of the 21st ACM Conference on Economics and Computation', pp. 257–272.

[18] Diakonikolas, I., Kamath, G., Kane, D., Li, J., Moitra, A. and Stewart, A. [2019], 'Robust estimators in high-dimensions without the computational intractability', *SIAM Journal on Computing* **48**(2), 742–864.

[19] Diakonikolas, I., Kontonis, V., Tzamos, C., Vakilian, A. and Zarifis, N. [2021], Learning online algorithms with distributional advice, *in* 'International Conference on Machine Learning', PMLR, pp. 2687–2696.

[20] Duchi, J. C. and Namkoong, H. [2021], 'Learning models with uniform performance via distributionally robust optimization', *The Annals of Statistics* **49**(3), 1378 – 1406.

[21] Fu, H., Immorlica, N., Lucier, B. and Strack, P. [2015], Randomization beats second price as a prior-independent auction, *in* 'Proceedings of the Sixteenth ACM Conference on Economics and Computation', pp. 323–323.

[22] Gallego, G. and Moon, I. [1993], 'The distribution free newsboy problem: review and extensions', *Journal of the Operational Research Society* **44**(8), 825–834.

[23] Guo, W., Jordan, M. and Zampetakis, E. [2021], 'Robust learning of optimal auctions', *Advances in Neural Information Processing Systems* **34**.

[24] Haghtalab, N., Han, Y., Shetty, A. and Yang, K. [2022], 'Oracle-efficient online learning for beyond worst-case adversaries', *arXiv preprint arXiv:2202.08549* .

[25] Haghtalab, N., Roughgarden, T. and Shetty, A. [2022], Smoothed analysis with adaptive adversaries, *in* '2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)', IEEE, pp. 942–953.

[26] Hampel, F. R. [1971], 'A general qualitative definition of robustness', *The annals of mathematical statistics* **42**(6), 1887–1896.

[27] Hanneke, S. [2021], 'Learning whenever learning is possible: Universal learning under general stochastic processes.', *J. Mach. Learn. Res.* **22**, 130–1.

[28] Haussler, D. [1992], 'Decision theoretic generalizations of the pac model for neural net and other learning applications', *Information and computation* **100**(1), 78–150.

[29] Huang, Z., Mansour, Y. and Roughgarden, T. [2018], 'Making the most of your samples', *SIAM Journal on Computing* **47**(3), 651–674.

[30] Huber, P. J. [1992], Robust estimation of a location parameter, *in* 'Breakthroughs in statistics', Springer, pp. 492–518.

[31] Kantorovich, L. V. and Akilov, G. P. [2016], *Functional analysis*, Elsevier.

[32] Kantorovich, L. V. and Rubinshtein, S. [1958], 'On a space of totally additive functions', *Vestnik of the St. Petersburg University: Mathematics* **13**(7), 52–59.

[33] Karnin, Z. S. and Anava, O. [2016], 'Multi-armed bandits: Competing with optimal sequences', *Advances in Neural Information Processing Systems* **29**.

[34] Kearns, M. J., Schapire, R. E. and Sellie, L. M. [1994], 'Toward efficient agnostic learning', *Machine Learning* **17**(2), 115–141.

[35] Kearns, M. and Li, M. [1993], 'Learning in the presence of malicious errors', *SIAM Journal on Computing* **22**(4), 807–837.

[36] Kim, S., Pasupathy, R. and Henderson, S. G. [2015], 'A guide to sample average approximation', *Handbook of simulation optimization* pp. 207–243.

[37] Kleywegt, A. J., Shapiro, A. and Homem-de Mello, T. [2002], 'The sample average approximation method for stochastic discrete optimization', *SIAM Journal on Optimization* **12**(2), 479–502.

[38] Klivans, A. R., Long, P. M. and Servedio, R. A. [2009], 'Learning halfspaces with malicious noise.', *Journal of Machine Learning Research* **10**(12).

[39] Lam, H. [2021], 'On the impossibility of statistically improving empirical optimization: A second-order stochastic dominance perspective', *arXiv preprint arXiv:2105.13419* .

[40] Levi, R., Perakis, G. and Uichanco, J. [2015], 'The data-driven newsvendor problem: new bounds and insights', *Operations Research* **63**(6), 1294–1306.

[41] Luo, H., Wei, C.-Y., Agarwal, A. and Langford, J. [2018], Efficient contextual bandits in non-stationary worlds, *in* 'Conference On Learning Theory', PMLR, pp. 1739–1776.

[42] Lykouris, T., Mirrokni, V. and Paes Leme, R. [2018], Stochastic bandits robust to adversarial corruptions, *in* 'Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing', pp. 114–122.

[43] Lykouris, T., Simchowitz, M., Slivkins, A. and Sun, W. [2021], Corruption-robust exploration in episodic reinforcement learning, *in* 'Conference on Learning Theory', PMLR, pp. 3242–3245.

[44] Mansour, Y., Mohri, M. and Rostamizadeh, A. [2009], 'Domain adaptation: Learning bounds and algorithms', *arXiv preprint arXiv:0902.3430* .

[45] Mohajerin Esfahani, P. and Kuhn, D. [2018], 'Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations', *Mathematical Programming* **171**(1), 115–166.

[46] Mohri, M. and Muñoz Medina, A. [2012], New analysis and algorithm for learning with drifting distributions, *in* 'International Conference on Algorithmic Learning Theory', Springer, pp. 124–138.

[47] Morse, P. M. and Kimball, G. E. [1946], Methods of operations research, Technical report, Center for Naval Analyses, Alexandria VA, Operations Evaluation Group.

[48] Myerson, R. B. [1981], 'Optimal auction design', *Mathematics of operations research* **6**(1), 58–73.

[49] Perakis, G. and Roels, G. [2008], 'Regret in the newsvendor model with partial information', *Operations Research* **56**(1), 188–203.

[50] Purohit, M., Svitkina, Z. and Kumar, R. [2018], 'Improving online algorithms via ml predictions', *Advances in Neural Information Processing Systems* **31**.

[51] Rakhlin, A., Sridharan, K. and Tewari, A. [2011], 'Online learning: Stochastic and constrained adversaries', *arXiv preprint arXiv:1104.5070* .

[52] Riley, J. and Zeckhauser, R. [1983], 'Optimal selling strategies: When to haggle, when to hold firm', *The Quarterly Journal of Economics* **98**(2), 267–289.

[53] Roughgarden, T. [2021], *Beyond the worst-case analysis of algorithms*, Cambridge University Press.

[54] Scarf, H. [1958], 'A min-max solution of an inventory problem', *Studies in the mathematical theory of inventory and production* .

[55] Servedio, R. A. [2003], 'Smooth boosting and learning with malicious noise', *The Journal of Machine Learning Research* **4**, 633–648.

[56] Shorack, G. R. [1979], 'The weighted empirical process of row independent random variables with arbitrary distribution functions', *Statistica Neerlandica* **33**(4), 169–189.

[57] Williams, D. [1991], *Probability with martingales*, Cambridge university press.

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes] See Section 1 for pointers to each contribution.
   (b) Did you describe the limitations of your work? [Yes] See Section 6
   (c) Did you discuss any potential negative societal impacts of your work? [N/A]
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [Yes]
   (b) Did you include complete proofs of all theoretical results? [Yes] All proofs are in the appendix.
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [N/A]
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [N/A]
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [N/A]
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [N/A]
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [N/A]
   (b) Did you mention the license of the assets? [N/A]
   (c) Did you include any new assets either in the supplemental material or as a URL? [N/A]
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]

This appendix is organized as follows. We present in Appendix A some preliminary results that we use across sections.

Appendix B complements the discussion in Section 3. We first present in Appendix B.1 a discussion about the empirical triangular convergence property. Appendix B.2 provides a counter example that shows that Theorem 1 does not hold for finite n and that nature may benefit from heterogeneity there. Appendix B.3 provides a complementary result to Theorem 1: we show that when the environment realization space Ξ is finite, one does not need to restrict attention to sample agnostic policies. Appendix B.4 provides a proof of Theorem 1.

Appendix C presents the proofs of all results stated in Section 4.

Appendix D complements the discussion in Section 5. In particular, we present in Appendix D.1 a proposition showing that SAA is achieving the optimal rate of the asymptotic worst-case regret for the pricing problem under Kolmogorov distance. We also present proofs of the results stated in Section 5.1. Finally, we present in Appendix D.2 proofs of the results stated in Section 5.2.

Note that every lemma or proposition stated in a section of the appendix is proved in the same section.

A Preliminary results

Proposition 4 (Connection between Wasserstein and Kolmogorov)**.** *Suppose that that Ξ is a compact subset of $\mathbb{R}$. Let $\mathcal{I}_K = (\mathcal{X}, \Xi, \mathcal{P}, d_K, g, \texttt{ORACLE})$ and $\mathcal{I}_W = (\mathcal{X}, \Xi, \mathcal{P}, d_W, g, \texttt{ORACLE})$ be two data-driven decision problems in heterogeneous environment that differ only along the distance dimension. Then, for every data-driven policy $\pi \in \Pi$ and any sample size $n \geq 1$ we have*

$$\mathfrak{R}^{\pi}_{\mathcal{I}_K,n}(\epsilon) \leq \mathfrak{R}^{\pi}_{\mathcal{I}_W,n}(\textit{diam}\,(\Xi) \cdot \epsilon),$$

where diam $(\Xi) := \sup_{\xi,\xi' \in \Xi} \xi - \xi'$

Proof of Proposition 4. For every $\mu \in \mathcal{P}$, let $\mathcal{U}^K_\epsilon(\mu)$ (resp. $\mathcal{U}^W_\epsilon(\mu)$) denote the heterogeneity ball centered around μ and with radius ϵ with respect to the Kolmogorov (resp. Wasserstein) distance.

We first argue that for any probability measure $\mu \in \mathcal{P}$, and any $\epsilon \geq 0$, we have that, $\mathcal{U}^K_\epsilon(\mu) \subset \mathcal{U}^W_{\text{diam}(\Xi)\cdot\epsilon}(\mu)$. Indeed, for every $\nu \in \mathcal{U}^K_\epsilon(\mu)$ we have that

$$d_W(\mu,\nu) \overset{(a)}{\leq} \text{diam}\,(\Xi) \cdot d_K(\mu,\nu) \leq \text{diam}\,(\Xi) \cdot \epsilon,$$

where (a) follows from (2). Therefore, for every $n \geq 1$, $\epsilon \geq 0$ and $\mu \in \mathcal{P}$ we have

$$\sup_{\nu_1,\ldots,\nu_n \in \mathcal{U}^K_\epsilon(\mu)} \mathcal{R}_n(\pi,\mu,\nu_1,\ldots,\nu_n) \leq \sup_{\nu_1,\ldots,\nu_n \in \mathcal{U}^W_{\text{diam}(\Xi)\cdot\epsilon}(\mu)} \mathcal{R}_n(\pi,\mu,\nu_1,\ldots,\nu_n).$$

By taking the supremum over $\mu \in \mathcal{P}$ we obtain the desired result. □

Lemma 1. *Let μ and ν be two probability measures on $\mathbb{R}$ with respective cumulative distributions F and H. Then, for every $\xi_1 \leq \xi_2$, we have that,*

$$F(\xi_1) \leq H(\xi_2) + \frac{d_W(\mu,\nu)}{\xi_2 - \xi_1}.$$

Proof of Lemma 1.

$$\begin{aligned} d_W(\mu,\nu) &= \int_{\xi\in\Xi} |F(\xi) - H(\xi)| d\xi \\ &\geq \int_{\xi_1}^{\xi_2} F(\xi) - H(\xi) d\xi \geq \int_{\xi_1}^{\xi_2} F(\xi_1) - H(\xi_2) d\xi = (\xi_2 - \xi_1) \cdot (F(\xi_1) - H(\xi_2)). \end{aligned}$$

□

B Additional results and proofs of section Section 3

B.1 Empirical triangular convergence property

We presented in Definition 3 a property on distances that is the key driver of Theorem 1. We now argue that this assumption is satisfied by many common distances on probability spaces. Indeed, it is known in that case that the Kolmogorov distance satisfies the empirical triangular convergence property. Formally, the following theorem in [56] applies.

Proposition 5 ([56], Theorem 2.1)**.** *If $\Xi = [0, M]$, the Kolmogorov distance satisfies the empirical triangular convergence property.*

A direct corollary of Proposition 5 and equation (2) is that the empirical triangular convergence property can be transferred to the Wasserstein distance.

Corollary 1. *If $\Xi = [0, M]$, the Wasserstein distance satisfies the empirical triangular convergence property.*

B.2 Finite sample heterogeneity

Theorem 1 shows that in the asymptotic regime, one may restrict attention to a problem in which nature selects a single historical distribution ν. We next show that this result does not necessarily hold for finite n, even for a well-behaved policy such as SAA.

Proposition 6. *There exists a data-driven decision problem in a heterogeneous environment $\mathcal{I} = (\mathcal{X}, \Xi, \mathcal{P}, d, g, \texttt{ORACLE})$ such that,*

$$\mathfrak{R}^{\pi^{SAA}}_{\mathcal{I},2}(\epsilon) > \sup_{\mu \in \mathcal{P}} \sup_{\nu \in \mathcal{U}_\eta(\mu)} \mathcal{R}_2\left(\pi^{SAA}, \mu, \nu, \nu\right).$$

Proof of Proposition 6. Let k be a positive integer.

We consider a variant of the ski-rental problem presented in Section 5.2, in which we set $b = 2k + 1$, $\Xi = [0, 3k + 2]$, $\mathcal{X} = \{0, \ldots, 3l + 2\}$ and $\mathcal{P} = \Delta(\{k, k + 1, 3k + 2\})$. Note that we restrict here the space of distributions to the ones that are supported on $\{k, k + 1, 3k + 2\}$. For the sake of simple notations, we denote by $\pi^{\text{SAA}}(\xi_1, \xi_2)$ the decision of SAA after observing samples ξ_1 and ξ_2.

We next enumerate all possible decisions of SAA as a function of the samples observed. The order of the samples does not affect the decision and we have,

$$\pi^{\text{SAA}}(k, k) = \pi^{\text{SAA}}(k + 1, k + 1) = \pi^{\text{SAA}}(k, k + 1) = 3k + 2, \quad \pi^{\text{SAA}}(3k + 2, 3k + 2) = 0,$$
$$\pi^{\text{SAA}}(k, 3k + 2) \overset{(a)}{=} k \quad \text{and} \quad \pi^{\text{SAA}}(k + 1, 3k + 2) \overset{(b)}{=} 0, \tag{9}$$

where we next give a detailed derivation of (a) and (b).

First, for every action $x \in \mathcal{X}$, denote by $\hat{c}(x) = \frac{1}{2}g(x, k) + \frac{1}{2}g(x, 3k + 2)$ the empirical cost of the action x against the samples k and $3k + 2$. We note that in that case, all actions are dominated by one of the following ones $\{0, k, 3k + 2\}$ as it is never optimal to purchase between skiing days. Furthermore, we have that,

$$\hat{c}(0) = b = 2k{+}1 \qquad \hat{c}(k) = \frac{1}{2}k{+}\frac{1}{2}(k + b) = 2k{+}\frac{1}{2} \qquad \hat{c}(3k{+}2) = \frac{1}{2}k{+}\frac{1}{2}(3k + 2) = 2k{+}1.$$

This implies (a). Similarly, we now let $\hat{c}(x) = \frac{1}{2}g(x, k+1) + \frac{1}{2}g(x, 3k+2)$. In that case, all actions are dominated by one of the following ones $\{0, k + 1, 3k + 2\}$ and,

$$\hat{c}(0) = b = 2k + 1 \qquad \hat{c}(k + 1) = \frac{1}{2}(k + 1) + \frac{1}{2}(k + 1 + b) = 2k + \frac{3}{2}$$
$$\hat{c}(3k + 2) = \frac{1}{2}(k + 1) + \frac{1}{2}(3k + 2) = 2k + \frac{3}{2}.$$

This implies (b).

Let $\epsilon = \infty$. Remark that for every $\mu \in \mathcal{P}$ and for any action $x \in \mathcal{X}$, we have

$$\begin{aligned}\mathcal{G}(x,\mu) - \text{opt}(\mu) &= \sum_{\xi \in \{k,k+1,3k+2\}} \mu(\xi) \cdot (\mathcal{G}(x,\delta_\xi) - \text{opt}(\mu)) \\ &\overset{(a)}{\leq} \sum_{\xi \in \{k,k+1,3k+2\}} \mu(\xi) \cdot (\mathcal{G}(x,\delta_\xi) - \text{opt}(\delta_\xi)) \leq \max_{\xi \in \{k,k+1,3k+2\}} \mathcal{G}(x,\delta_\xi) - \text{opt}(\delta_\xi),\end{aligned}$$

where (a) holds because $\text{opt}(\mu) \geq \sum_{\xi \in \{k,k+1,3k+2\}} \mu(\xi) \cdot \text{opt}(\delta_\xi)$.

Therefore the maximum regret incurred by SAA is at most

$$\mathfrak{R}^{\pi^{\text{SAA}}}_{\mathcal{I},2}(\epsilon) \leq \max_{x \in \{0,k,3k+1\}} \max_{\xi \in \{k,k+1,3k+2\}} \mathcal{G}(x,\delta_\xi) - \text{opt}(\delta_\xi) \overset{(a)}{\leq} 2k,$$

where (a) is obtained by enumerating all possible cases as presented in Table 2.

	$\xi = k$	$\xi = k+1$	$\xi = 3k+2$
$x = 0$	$k+1$	k	0
$x = k$	0	$2k$	k
$x = 3k+2$	0	0	$k+1$

Table 2: Different values of $\mathcal{G}(x,\delta_\xi) - \text{opt}(\delta_\xi)$.

We also note from Table 2 that the worst-case regret is achieved when SAA selects action k and when $\mu = \delta_{k+1}$. We finally, show that this worst-case regret can only be obtained when nature uses different historical distribution $\nu_1 \neq \nu_2$. First note that the out-of-sample distribution must be δ_{k+1} in the worst-case to achieve a regret of $2k$. Furthemore, we have seen in (9) that SAA selects k if and only if it observes samples k and $3k+2$. Therefore, a regret of $2k$ can only be achieved if the probability of selecting action k is 1. By setting $\nu_1 = \delta_k$ and $\nu_2 = \delta_{3k+2}$ we ensure that SAA selects action k with probability 1. In contrast, for any distribution ν such that $\nu(k) > 0$ and $\nu(3k+2) > 0$, the probability that SAA selects action k when observing two samples from ν is strictly smaller than 1.

This concludes the proof and shows that,

$$\mathfrak{R}^{\pi^{\text{SAA}}}_{\mathcal{I},2}(\epsilon) > \sup_{\mu \in \mathcal{P}} \sup_{\nu \in \mathcal{U}_\epsilon(\mu)} \mathcal{R}_2\left(\pi^{\text{SAA}}, \mu, \nu, \nu\right).$$

□

B.3 Finite space of environment realizations

We next show that when Ξ is finite, one does not need to restrict attention to sample agnostic policies.

Theorem 6 (Reduction finite space)**.** *Let $\mathcal{I} = (\mathcal{X}, \Xi, \mathcal{P}, d, g, \texttt{ORACLE})$ be a data-driven decision problem in a heterogeneous environment. Assume that Ξ is finite and that g is bounded on $\mathcal{X} \times \Xi$. Then for every $\pi \in \Pi$ we have,*

$$\mathfrak{R}^{\pi}_{\mathcal{I},\infty}(\epsilon) \leq \limsup_{n \to \infty} \lim_{\eta \to \epsilon^+} \sup_{\mu \in \mathcal{P}} \sup_{\nu \in \mathcal{U}_\eta(\mu)} |\mathcal{G}(\pi(n,\nu),\mu) - \textit{opt}(\mu)|.$$

Proof of Theorem 6. We prove this statement for the Kolmogorov distance. A similar analysis yields the result for Wasserstein distance.

Step 1: Fix $n \geq 1$ and $\epsilon \geq 0$. We first show that for any distribution μ and any $\alpha > 0$ we have

$$\begin{aligned}\sup_{\nu_1,\ldots,\nu_n \in \mathcal{U}_\epsilon(\mu)} |\mathbb{E}_{\xi_i \sim \nu_i}[\mathcal{G}(\pi(n,\hat{\mu}_{\boldsymbol{\xi}^n}),\mu)] - \text{opt}(\mu)| &\leq \sup_{\nu \in \mathcal{U}_{\epsilon+\alpha}(\mu)} |\mathcal{G}(\pi(n,\nu),\mu) - \text{opt}(\mu)| \\ &\quad + 4C \cdot |\Xi| \exp(-2n\alpha^2) \qquad (10)\end{aligned}$$

For every sequence $\nu_1, \ldots, \nu_n \in \mathcal{U}_\epsilon(\mu)$, let $\bar{\nu}_n = \frac{1}{n}\sum_{i=1}^n \nu_i$ and let E_n be the event $\{d_K(\hat{\mu}_{\boldsymbol{\xi^n}}, \bar{\nu}_n) \leq \alpha\}$ and let C be a constant that bounds $|g|$ over $\mathcal{X} \times \Xi$. We have that,

$$\begin{aligned}|\mathbb{E}_{\xi_i \sim \nu_i}\left[\mathcal{G}\left(\pi(n, \hat{\mu}_{\boldsymbol{\xi^n}}), \mu\right)\right] - \text{opt}(\mu)| &\overset{(a)}{\leq} \sup_{\nu \in \mathcal{U}_{\epsilon+\alpha}(\mu)} |\mathcal{G}\left(\pi(n,\nu),\mu\right) - \text{opt}(\mu)|\mathbb{P}\left(E_n\right) + 4C \cdot \left(1 - \mathbb{P}\left(E_n\right)\right) \\ &\overset{(b)}{\leq} \sup_{\nu \in \mathcal{U}_{\epsilon+\alpha}(\mu)} |\mathcal{G}\left(\pi(n,\nu),\mu\right) - \text{opt}(\mu)| + 4C \cdot |\Xi| \exp(-2n\alpha^2),\end{aligned}$$

where (a) holds because under the even E_n, we have that $\hat{\mu}_{\boldsymbol{\xi^n}} \in \mathcal{U}_{\epsilon+\alpha}(\mu)$ since $d_K(\hat{\mu}_{\boldsymbol{\xi^n}}, \mu) \leq d_K(\hat{\mu}_{\boldsymbol{\xi^n}}, \bar{\nu}) + d_K(\bar{\nu}, \mu) \leq \alpha + \epsilon$ and (b) follows from the next lemma.

Lemma 2. *Let Ξ be such that $|\Xi| < \infty$. For any distributions $\nu_1, \ldots, \nu_n$ over Ξ, and any $\alpha > 0$, we have*

$$\Pr_{\hat{\nu}}[d_K(\hat{\nu}, \bar{\nu}) \leq \epsilon] \geq 1 - 2|\Xi| \exp(-2n\epsilon^2),$$

where $\hat{\nu}(\xi) := \frac{1}{n}\sum_{i=1}^n \mathbb{1}\{\xi = \xi_i\}$ for samples $\xi_i \sim \nu_i$ and $\bar{\nu} = \frac{1}{n}\sum_{i=1}^n \nu_i$.

We finally derive (10) by taking the supremum over historical distributions.

Step 2: Let $L = \lim_{\eta \to \epsilon^+} \limsup_{n\to\infty} \sup_{\mu \in \mathcal{P}} \sup_{\nu \in \mathcal{U}_\eta(\mu)} |\mathcal{G}\left(\pi(n,\nu),\mu\right) - \text{opt}(\mu)|$. Fix $\delta > 0$ and let α be such that,

$$\limsup_{n\to\infty} \sup_{\mu \in \mathcal{P}} \sup_{\nu \in \mathcal{U}_{\epsilon+\alpha}(\mu)} |\mathcal{G}\left(\pi(n,\nu),\mu\right) - \text{opt}(\mu)| \leq L + \frac{\delta}{2}.$$

Let $N = \frac{1}{2\alpha^2}\log\left(\frac{8C\cdot|\Xi|}{\delta}\right)$. Then, for every $n \geq N$ we have by equation (10) that

$$\sup_{\nu_1,\ldots,\nu_n \in \mathcal{U}_\epsilon(\mu)} |\mathbb{E}_{\xi_i \sim \nu_i}\left[\mathcal{G}\left(\pi(n, \hat{\mu}_{\boldsymbol{\xi^n}}), \mu\right)\right] - \text{opt}(\mu)| \leq \sup_{\nu \in \mathcal{U}_{\epsilon+\alpha}(\mu)} |\mathcal{G}\left(\pi(n,\nu),\mu\right) - \text{opt}(\mu)| + \frac{\delta}{2}.$$

By taking the supremum over μ and then the limit over n we finally conclude that,

$$\limsup_{n\to\infty} \mathfrak{R}^{\pi}_{\mathcal{I},n}(\epsilon) \leq \limsup_{n\to\infty} \sup_{\nu \in \mathcal{U}_{\epsilon+\alpha}(\mu)} \sup_{\mu \in \mathcal{P}} |\mathcal{G}\left(\pi(n,\nu),\mu\right) - \text{opt}(\mu)| + \frac{\delta}{2} \leq L + \delta.$$

Since this holds for any $\delta > 0$, we obtain the desired result, by letting δ go to 0.

□

Proof of Lemma 2. Denote by $F_{\hat{\nu}}$ and $F_{\bar{\nu}}$ the cdf associated to both probability measures. Then, for any $\xi \in \Xi$, by Hoeffding's inequality, we have

$$\Pr[|F_{\hat{\nu}}(\xi) - F_{\bar{\nu}}(\xi)| > \epsilon] < 2\exp(-2n\epsilon^2).$$

Taking a union bound over $\xi \in \Xi$ completes the proof. □

B.4 Proofs of results

Proof of Theorem 1. Fix $\epsilon > 0$. To show this theorem, it is sufficient to prove that for every sequence of distributions $(\mu_n)_{n\in\mathbb{N}^*}$, $\boldsymbol{\nu} := (\nu_{i,n})_{1\leq i\leq n, n\in\mathbb{N}^*}$, such that $\nu_{i,n} \in \mathcal{U}_\epsilon(\mu_n)$ for every i and n, we have that

$$\limsup_{n\to\infty} \mathcal{R}_n\left(\pi, \mu_n, \nu_{1,n}, \ldots, \nu_{n,n}\right) \leq \lim_{\eta\to\epsilon^+} \sup_{\mu\in\mathcal{P}} \sup_{\nu\in\mathcal{U}_\eta(\mu)} |\text{opt}(\mu) - \mathcal{G}\left(\pi(\nu),\mu\right)|. \tag{11}$$

Indeed, for every $\alpha > 0$, we can construct a sequence of distribution, $(\mu_n)_{n\in\mathbb{N}^*}$, $\boldsymbol{\nu} := (\nu_{i,n})_{1\leq i\leq n, n\in\mathbb{N}^*}$, such that for every i and n, we have $\nu_{i,n} \in \mathcal{U}_\epsilon(\mu)$ and

$$\mathfrak{R}^{\pi}_{\mathcal{I},n}(\epsilon) \leq \mathcal{R}_n\left(\pi, \mu_n, \nu_{1,n}, \ldots, \nu_{n,n}\right) + \alpha.$$

Assuming that (11) holds and by taking the limit we obtain that for every $\alpha > 0$,

$$\mathfrak{R}^{\pi}_{\mathcal{I},\infty}(\epsilon) \leq \lim_{\eta\to\epsilon^+} \sup_{\mu\in\mathcal{P}} \sup_{\nu\in\mathcal{U}_\eta(\mu)} |\text{opt}(\mu) - \mathcal{G}\left(\pi(\nu),\mu\right)| + \alpha.$$

We obtain the desired result by taking α to 0. We now prove (11).

For every sample path $(\xi_{i,n})_{1\leq i\leq n, n\in\mathbb{N}^*}$ s.t. $\xi_{i,n} \sim \nu_{i,n}$ we define the sequence of empirical distributions $(\hat{\nu}_n)_{n\in\mathbb{N}^*}$ such that for every $n \geq 1$ and every measurable set A, we have $\hat{\nu}_n(A) := \frac{1}{n}\sum_{i=1}^n \mathbb{1}\{\xi_{i,n} \in A\}$. We also define the sequence $(\bar{\nu}_n)_{n\in\mathbb{N}^*}$ such that for every $n \geq 1$, we have $\bar{\nu}_n = \frac{1}{n}\sum_{i=1}^n \nu_{i,n}$.

In turn, let $(Z_n)_{n\in\mathbb{N}^*}$ be the sequence of random variables defined for every $n \geq 1$ as $Z_n := |\text{opt}(\mu_n) - \mathcal{G}(\pi(n, \hat{\nu}_n), \mu_n)|$. Remark that there exists a constant K such that for every $n \geq 1$, $|Z_n| \leq K$ almost surely because g is bounded on $\mathcal{X} \times \Xi$.

Therefore we have that,

$$\begin{aligned}
\limsup_{n\to\infty} \mathcal{R}_n(\pi, \mu_n, \nu_{1,n}, \ldots, \nu_{n,n}) &= \limsup_{n\to\infty} |\mathbb{E}_{\boldsymbol{\nu}}[\text{opt}(\mu_n) - \mathcal{G}(\pi(n, \hat{\nu}_n), \mu_n)]| \\
&\overset{(a)}{\leq} \limsup_{n\to\infty} \mathbb{E}_{\boldsymbol{\nu}}[|\text{opt}(\mu_n) - \mathcal{G}(\pi(n, \hat{\nu}_n), \mu_n)|] \\
&= \limsup_{n\to\infty} \mathbb{E}_{\boldsymbol{\nu}}[Z_n] \overset{(b)}{\leq} \mathbb{E}_{\boldsymbol{\nu}}\left[\limsup_{n\to\infty} Z_n\right]
\end{aligned}$$

where (a) is a consequence of Jensen's inequality and (b) holds by the reversed Fatou lemma [57, Section 5.4] which can be applied because $(Z_n)_{n\in\mathbb{N}}$ are almost surely bounded.

We next show that $\mathbb{E}_{\boldsymbol{\nu}}[\limsup_{n\to\infty} Z_n] \leq \lim_{\eta\to\epsilon^+} \sup_{\mu\in\mathcal{P}} \sup_{\nu\in\mathcal{U}_\eta(\mu)} |\text{opt}(\mu) - \mathcal{G}(\pi(\nu), \mu)|$. To do so, we show a stronger statement and show that $\limsup_{n\to\infty} Z_n \leq \lim_{\eta\to\epsilon^+} \sup_{\mu\in\mathcal{P}} \sup_{\nu\in\mathcal{U}_\eta(\mu)} |\text{opt}(\mu) - \mathcal{G}(\pi(\nu), \mu)|$ almost surely.

Fix a realization ω from the set of realizations such that ETC is verified. Remark that, for every $n \geq 1$, we have

$$\begin{aligned}
d(\hat{\nu}_n(\omega), \mu_n) &\leq d(\hat{\nu}_n(\omega), \bar{\nu}_n) + d(\bar{\nu}_n, \mu_n) \\
&= d(\hat{\nu}_n(\omega), \bar{\nu}_n) + d\left(\frac{1}{n}\sum_{i=1}^n \nu_{i,n}, \mu_n\right) \\
&\overset{(a)}{\leq} d(\hat{\nu}_n(\omega), \bar{\nu}_n) + \frac{1}{n}\sum_{i=1}^n d(\nu_{i,n}, \mu_n) \overset{(b)}{\leq} d(\hat{\nu}_n(\omega), \bar{\nu}_n) + \epsilon,
\end{aligned} \tag{12}$$

where (a) follows from the convexity of d and (b) holds because $d(\nu_{i,n}, \mu) \leq \epsilon$ for all $i \leq n$.

We now define the sequence $(\delta_n(\omega))_{n\in\mathbb{N}^*}$ such that for every n, $\delta_n(\omega) := \max(d(\hat{\nu}_n(\omega), \mu_n), \epsilon)$. Remark that since d satisfies the empirical triangular convergence property, we have $d(\hat{\nu}_n(\omega), \bar{\nu}_n) \to 0$ as $n \uparrow \infty$. This, together with (12), implies that $\delta_n(\omega) \to \epsilon$ as $n \uparrow \infty$.

Furthermore, for every $n \geq 1$ we have that,

$$\begin{aligned}
Z_n(\omega) &\overset{(a)}{=} |\text{opt}(\mu_n) - \mathcal{G}(\pi(\hat{\nu}_n(\omega)), \mu_n)| \\
&\leq \sup_{\mu\in\mathcal{P}} \sup_{\nu\in\mathcal{U}_{d(\mu_n, \hat{\nu}_n(\omega))}(\mu)} |\text{opt}(\mu) - \mathcal{G}(\pi(\nu), \mu)| \\
&\leq \sup_{\mu\in\mathcal{P}} \sup_{\nu\in\mathcal{U}_{\delta_n(\omega)}(\mu)} |\text{opt}(\mu) - \mathcal{G}(\pi(\nu), \mu)|.
\end{aligned}$$

where (a) holds because π is sample-size-agnostic.

Hence, by taking the limit we obtain that,

$$\begin{aligned}
\limsup_{n\to\infty} Z_n(\omega) &\leq \limsup_{n\to\infty} \sup_{\mu\in\mathcal{P}} \sup_{\nu\in\mathcal{U}_{\delta_n(\omega)}(\mu)} |\text{opt}(\mu) - \mathcal{G}(\pi(\nu), \mu)| \\
&= \lim_{\eta\to\epsilon^+} \sup_{\mu\in\mathcal{P}} \sup_{\nu\in\mathcal{U}_\eta(\mu)} |\text{opt}(\mu) - \mathcal{G}(\pi(\nu), \mu)|.
\end{aligned}$$

This concludes the proof. □

C Additional results and proofs of section Section 4

Proof of Theorem 2. *Step 1:* We first prove the result for the Wasserstein distance.

We show this result by bounding the value of the uniform DRO problem (5) and by applying Theorem 1. In particular we show the following.

$$\sup_{\mu\in\mathcal{P}}\sup_{\nu\in\mathcal{U}_\epsilon(\mu)} \left|\mathrm{opt}(\mu) - \mathcal{G}\left(\pi^{\mathrm{SAA}}(\nu),\mu\right)\right| \leq 2L\cdot\epsilon. \tag{13}$$

We first show that for any fixed action $x\in\mathcal{X}$, the difference in expected objectives achieved by x under two distributions is controlled by the Wasserstein distance between these two distributions. Formally, we show the following result.

Lemma 3. *Let g be an objective function and let $x\in\mathcal{X}$ be such that $g(x,\cdot)$ is L-lipschitz-continuous. Then for every $\mu,\nu\in\mathcal{P}$, we have that,*

$$\left|\mathcal{G}\left(x,\mu\right) - \mathcal{G}\left(x,\nu\right)\right| \leq L\cdot d_W(\mu,\nu).$$

First, assume that ORACLE is a maximization oracle. For every $\mu,\nu\in\mathcal{P}$, we have that,

$$\begin{aligned}
\mathrm{opt}(\mu) - \mathcal{G}\left(\pi^{\mathrm{SAA}}(\nu),\mu\right) &\overset{(a)}{=} \mathrm{opt}(\mu) - \mathcal{G}\left(\mathtt{ORACLE}\left(\nu\right),\mu\right)\\
&= \mathcal{G}\left(\mathtt{ORACLE}\left(\mu\right),\mu\right) - \mathcal{G}\left(\mathtt{ORACLE}\left(\mu\right),\nu\right)\\
&\qquad + \mathcal{G}\left(\mathtt{ORACLE}\left(\mu\right),\nu\right) - \mathcal{G}\left(\mathtt{ORACLE}\left(\nu\right),\nu\right)\\
&\qquad + \mathcal{G}\left(\mathtt{ORACLE}\left(\nu\right),\nu\right) - \mathcal{G}\left(\mathtt{ORACLE}\left(\nu\right),\mu\right)\\
&\overset{(b)}{\leq} L\cdot d_W(\mu,\nu) + 0 + L\cdot d_W(\mu,\nu),
\end{aligned}$$

where (a) follows by noting that $\pi^{\mathrm{SAA}}(\nu) = \mathtt{ORACLE}\left(\nu\right)$ by definition and (b) is obtained by applying Lemma 3 to the first and the last differences and remarking that $\mathcal{G}\left(\mathtt{ORACLE}\left(\mu\right),\nu\right) - \mathcal{G}\left(\mathtt{ORACLE}\left(\nu\right),\nu\right) \leq 0$.

The latter implies that

$$\sup_{\mu\in\mathcal{P}}\sup_{\nu\in\mathcal{U}_\epsilon(\mu)} \mathrm{opt}(\mu) - \mathcal{G}\left(\pi^{\mathrm{SAA}}(\nu),\mu\right) \leq 2L\cdot\sup_{\mu\in\mathcal{P}}\sup_{\nu\in\mathcal{U}_\epsilon(\mu)} d_W(\mu,\nu) \leq 2L\cdot\epsilon.$$

One can show the same result for a minimization oracle.

By Theorem 1, we finally conclude that,

$$\mathfrak{R}^{\pi^{\mathrm{SAA}}}_{\mathcal{I}_W,\infty}(\epsilon) \leq \lim_{\eta\to\epsilon^+}\sup_{\mu\in\mathcal{P}}\sup_{\nu\in\mathcal{U}_\eta(\mu)} \left|\mathrm{opt}(\mu) - \mathcal{G}\left(\pi^{\mathrm{SAA}}(\nu),\mu\right)\right| \leq \lim_{\eta\to\epsilon^+} 2L\cdot\eta = 2L\cdot\epsilon.$$

Step 2: We now prove the result for the Kolmogorov distance. We have that,

$$\mathfrak{R}^{\pi^{\mathrm{SAA}}}_{K,\infty}(\epsilon) \overset{(a)}{\leq} \mathfrak{R}^{\pi^{\mathrm{SAA}}}_{W,\infty}(\mathrm{diam}\left(\Xi\right)\cdot\epsilon) \overset{(b)}{\leq} 2L\cdot M\cdot\epsilon,$$

where (a) follows from Proposition 4 and (b) holds by step 1 and because $\mathrm{diam}\left([0,M]\right) = M$. $\square$

Proof of Proposition 1. *Step 1:* We first show that the upper bounds on the performance of SAA hold for both distances. Remark that for every $x\in\mathcal{X}$, the newsvendor cost $g(x,\cdot)$ is L-Lipschitz, for $L = \max\left(c_u,c_o\right)$. Therefore Theorem 2 implies that,

$$\mathfrak{R}^{\pi^{\mathrm{SAA}}}_{K,\infty}(\epsilon) \leq 2\max\left(c_u,c_o\right)\cdot M\cdot\epsilon \qquad \text{and} \qquad \mathfrak{R}^{\pi^{\mathrm{SAA}}}_{W,\infty}(\epsilon) \leq 2\max\left(c_u,c_o\right)\cdot\epsilon$$

Step 2: We now show the lower bound for the Komlogorov distance.

For every $p\in[0,1]$, let $\mathcal{B}_M(p)$ denote the two point mass distribution which puts mass p at M and mass $1-p$ at 0. Furthermore, let $q := \frac{c_u}{c_u+c_o}$ denote the critical fractile. Fix $\epsilon\leq\min\left(q,1-q\right)$ and let $\tilde{\nu} := \mathcal{B}_M\left(1-q\right)$, $\mu_- := \mathcal{B}_M\left(1-q-\epsilon\right)$ and $\mu_+ := \mathcal{B}_M\left(1-q+\epsilon\right)$.

For a fixed $n \geq 1$, we have that,

$$\begin{aligned}
\inf_{\pi\in\Pi} \mathfrak{R}^{\pi}_{K,n}(\epsilon) &\overset{(a)}{\geq} \inf_{\pi\in\Pi} \sup_{\mu\in\{\mu_-,\mu_+\}} \sup_{\nu\in\mathcal{U}_\epsilon(\mu)} \mathcal{R}_n\left(\pi,\mu,\nu,\ldots,\nu\right) \\
&\overset{(b)}{\geq} \inf_{\pi\in\Pi} \sup_{\mu\in\{\mu_-,\mu_+\}} \mathcal{R}_n\left(\pi,\mu,\tilde{\nu},\ldots,\tilde{\nu}\right) \\
&= \inf_{\pi\in\Pi} \sup_{\mu\in\{\mu_-,\mu_+\}} \mathbb{E}_{\xi_i\sim\tilde{\nu}}\left[\mathcal{G}\left(\pi(n,\hat{\mu}_{\boldsymbol{\xi}^n}),\mu\right)\right] - \text{opt}(\mu) \\
&\overset{(c)}{=} \inf_{\alpha\in\Delta(\mathcal{X})} \sup_{\mu\in\{\mu_-,\mu_+\}} \mathcal{G}\left(\alpha,\mu\right) - \text{opt}(\mu) \overset{(d)}{\geq} \sup_{\beta\in\Delta(\{\mu_-,\mu_+\})} \inf_{x\in\mathcal{X}} \mathbb{E}_{\mu\sim\beta}\left[\mathcal{G}\left(x,\mu\right) - \text{opt}(\mu)\right]
\end{aligned}$$

where (a) follows from the fact that we restrict the space of out-of-sample distributions to be $\{\mu_-,\mu_+\}$ and we restrict historical distributions $\nu_1,\ldots,\nu_n$ to be all equal. (b) holds as $\tilde{\nu}\in\mathcal{U}_\epsilon(\mu)$ for both μ_- and μ_+. We obtain (c) by remarking that when $\tilde{\nu}$ is fixed, the policy does not need to look at the data as it does not give any extra information on the value of μ. Indeed, for every $\pi\in\Pi$, we construct a probability measure $\alpha\in\Delta\left(\mathcal{X}\right)$ which satisfies for every $x\in\mathcal{X}$,

$$\alpha(x) = \mathbb{P}_{\xi_i\sim\tilde{\nu}}\left(\pi(n,\hat{\mu}_{\boldsymbol{\xi}^n}) = x\right).$$

We then have for every $\mu\in\{\mu_-,\mu_+\}$,

$$\mathbb{E}_{\xi_i\sim\tilde{\nu}}\left[\mathcal{G}\left(\pi(n,\hat{\mu}_{\boldsymbol{\xi}^n}),\mu\right)\right] - \text{opt}(\mu) = \mathbb{E}_{x\sim\alpha}\left[\mathcal{G}\left(x,\mu\right)\right] - \text{opt}(\mu) = \mathcal{G}\left(\alpha,\mu\right) - \text{opt}(\mu).$$

By taking the supremum over μ, we obtain,

$$\sup_{\mu\in\{\mu_-,\mu_+\}} \mathbb{E}_{\xi_i\sim\tilde{\nu}}\left[\mathcal{G}\left(\pi(n,\hat{\mu}_{\boldsymbol{\xi}^n}),\mu\right)\right] - \text{opt}(\mu) = \sup_{\mu\in\{\mu_-,\mu_+\}} \mathcal{G}\left(\alpha,\mu\right) - \text{opt}(\mu)$$

which implies the inequality,

$$\inf_{\pi\in\Pi} \sup_{\mu\in\{\mu_-,\mu_+\}} \mathbb{E}_{\xi_i\sim\tilde{\nu}}\left[\mathcal{G}\left(\pi(n,\hat{\mu}_{\boldsymbol{\xi}^n}),\mu\right)\right] - \text{opt}(\mu) \geq \inf_{\alpha\in\Delta(\mathcal{X})} \sup_{\mu\in\{\mu_-,\mu_+\}} \mathcal{G}\left(\alpha,\mu\right) - \text{opt}(\mu).$$

In addition, the inequality

$$\inf_{\pi\in\Pi} \sup_{\mu\in\{\mu_-,\mu_+\}} \mathbb{E}_{\xi_i\sim\tilde{\nu}}\left[\mathcal{G}\left(\pi(n,\hat{\mu}_{\boldsymbol{\xi}^n}),\mu\right)\right] - \text{opt}(\mu) \leq \inf_{\alpha\in\Delta(\mathcal{X})} \sup_{\mu\in\{\mu_-,\mu_+\}} \mathcal{G}\left(\alpha,\mu\right) - \text{opt}(\mu)$$

holds by inclusion.

Finally, (d) follows from Yao's minimax principle.

Furthermore, remark that for every $x\in[0,M]$ we have,

$$\begin{aligned}
\mathcal{G}\left(x,\mu_+\right) - \text{opt}(\mu_+) &= (q-\epsilon)\cdot c_o\cdot x + (1-q+\epsilon)\cdot c_u\cdot(M-x) - c_o\cdot(q-\epsilon)\cdot M = (c_u+c_o)\cdot(M-x)\cdot\epsilon \\
\mathcal{G}\left(x,\mu_-\right) - \text{opt}(\mu_-) &= (q+\epsilon)\cdot c_o\cdot x + (1-q-\epsilon)\cdot c_u\cdot(M-x) - c_u\cdot(1-q-\epsilon)\cdot M = (c_u+c_o)\cdot x\cdot\epsilon
\end{aligned}$$

Finally, let $\tilde{\beta}$ be the randomization which puts mass $\frac{1}{2}$ on μ_+ and mass $\frac{1}{2}$ on μ_-. We have that,

$$\inf_{x\in\mathcal{X}} \mathbb{E}_{\mu\sim\tilde{\beta}}\left[\mathcal{G}\left(x,\mu\right) - \text{opt}(\mu)\right] = \inf_{x\in\mathcal{X}} \frac{c_u+c_o}{2}\cdot(M-x)\cdot\epsilon + \frac{c_u+c_o}{2}\cdot x\cdot\epsilon = \frac{c_u+c_o}{2}\cdot M\cdot\epsilon.$$

Therefore,

$$\inf_{\pi\in\Pi} \mathfrak{R}^{\pi}_{K,n}(\epsilon) \geq \sup_{\beta\in\Delta(\{\mu_-,\mu_+\})} \inf_{x\in\mathcal{X}} \mathbb{E}_{\mu\sim\beta}\left[\mathcal{G}\left(x,\mu\right) - \text{opt}(\mu)\right] \geq \frac{c_u+c_o}{2}\cdot M\cdot\epsilon.$$

And by taking the limit we obtain the desired result,

$$\inf_{\pi\in\Pi} \limsup_{n\to\infty} \mathfrak{R}^{\pi}_{K,n}(\epsilon) \overset{(a)}{=} \limsup_{n\to\infty} \inf_{\pi\in\Pi} \mathfrak{R}^{\pi}_{K,n}(\epsilon) \geq \frac{c_u+c_o}{2}\cdot M\cdot\epsilon,$$

where (a) holds because π observes n.

Step 3: We now derive the lower bound for the Wasserstein distance. Remark that for every $\epsilon \leq M\cdot\max\left(q,1-q\right)$,

$$\inf_{\pi\in\Pi} \mathfrak{R}^{\pi}_{W,\infty}(\epsilon) \overset{(a)}{\geq} \inf_{\pi\in\Pi} \mathfrak{R}^{\pi}_{K,\infty}\left(\frac{\epsilon}{M}\right) \overset{(b)}{\geq} \frac{c_u+c_o}{2}\cdot\epsilon,$$

where (a) follows from Proposition 4 and (b) is a consequence of step 2.

□

Proof of Lemma 3. Recall that $\mathcal{F}_L$ denotes the set of L-Lipschitz mappings from (Ξ, d_Ξ) to $\mathbb{R}$. Then, we have

$$\begin{aligned}|\mathcal{G}(x,\mu) - \mathcal{G}(x,\nu)| &= L \cdot \left| \mathbb{E}_{\xi\sim\mu}\left[\frac{1}{L}\cdot g(x,\xi)\right] - \mathbb{E}_{\xi\sim\nu}\left[\frac{1}{L}\cdot g(x,\xi)\right]\right| \\ &\overset{(a)}{\leq} L \cdot \sup_{f\in\mathcal{F}_1} |\mathbb{E}_{\xi\sim\mu}[f(\xi)] - \mathbb{E}_{\xi\sim\nu}[f(\xi)]| \overset{(b)}{=} L\cdot d_W(\mu,\nu),\end{aligned}$$

where (a) holds because $\frac{1}{L}g(x,\cdot)$ is 1-Lipschitz and (b) uses the dual representation of the Wasserstein distance [32, 31] defined as follows. For any mapping from Ξ to $\mathbb{R}$, recall that $\text{Lip}(f)$ is the Lipschitz constant of f. We have that for every probability measures $\mu_1, \mu_2 \in \mathcal{P}$,

$$d_W(\mu_1,\mu_2) = \sup_{f \text{ s.t. } \text{Lip}(f)\,\leq\, 1} \left| \int_\xi f(\xi) d\mu_1(\xi) - \int_\xi f(\xi) d\mu_2(\xi)\right|.$$

□

D Additional results and proofs of Section 5

D.1 Additional results and proofs of Section 5.1

Proposition 7 (Kolmogorov Pricing)**.** *For the data-driven pricing problem in a heterogenous environment under the Kolmogorov distance, for any $\epsilon > 0$, we have,*

$$\frac{M\cdot\epsilon}{2} \leq \inf_{\pi\in\Pi} \mathfrak{R}^{\pi}_{K,\infty}(\epsilon) \leq \mathfrak{R}^{\pi^{SAA}}_{K,\infty}(\epsilon) \leq 2M\cdot\epsilon.$$

Proof of Proposition 7. *Step 1-a:* We derive an upper bound on the worst-case asymptotic regret of SAA by using Theorem 1 and by showing the following upper bound

$$\sup_{\mu\in\mathcal{P}} \sup_{\nu\in\mathcal{U}_\epsilon(\mu)} |\text{opt}(\mu) - \mathcal{G}\left(\pi^{\text{SAA}}(\nu),\mu\right)| \leq 2M\cdot\epsilon. \tag{14}$$

For every $\mu,\nu\in\mathcal{P}$, we have that,

$$\begin{aligned}\text{opt}(\mu) - \mathcal{G}\left(\pi^{\text{SAA}}(\nu),\mu\right) &\overset{(a)}{=} \text{opt}(\mu) - \mathcal{G}\left(\texttt{ORACLE}(\nu),\mu\right) \\ &= \left(\text{opt}(\mu) - \mathcal{G}\left(\texttt{ORACLE}(\mu),\nu\right)\right) + \left(\mathcal{G}\left(\texttt{ORACLE}(\mu),\nu\right) - \mathcal{G}\left(\texttt{ORACLE}(\nu),\nu\right)\right) \\ &\quad + \left(\mathcal{G}\left(\texttt{ORACLE}(\nu),\nu\right) - \mathcal{G}\left(\texttt{ORACLE}(\nu),\mu\right)\right) \\ &\overset{(b)}{\leq} M\cdot d_K(\mu,\nu) + 0 + M\cdot d_K(\mu,\nu) = 2M\cdot d_K(\mu,\nu),\end{aligned}$$

where (a) follows by noting that $\pi^{\text{SAA}}(\nu) = \texttt{ORACLE}(\nu)$ by definition and (b) follows from bounding the first and third term by using the fact that for every $x\in\mathcal{X}$, we have for the pricing problem that

$$\mathcal{G}(x,\mu) - \mathcal{G}(x,\nu) = x\cdot(1-F(x)) - x\cdot(1-H(x)) = x\,(H(x)-F(x)) \leq M\cdot d_K(\mu,\nu),$$

where F (resp. H) is the cdf of μ (resp. ν).

By taking supremums over μ and ν we obtain (14). Furthermore, Theorem 1 and equation (14) imply that,

$$\mathfrak{R}^{\pi^{\text{SAA}}}_{K,\infty}(\epsilon) \leq 2M\cdot\epsilon.$$

Step 2: We now show the lower bound for the Komlogorov distance. Let $\tilde{\nu}$, μ_+ and μ_- be probability measures supported on $\{\frac{M}{2}, M\}$ such that, $\tilde{\nu}(M) = \frac{1}{2}$, $\mu_+(M) = \frac{1}{2}+\epsilon$ and $\mu_-(M) = \frac{1}{2}-\epsilon$.

For a fixed $n\geq 1$, we have that,

$$\begin{aligned}\inf_{\pi\in\Pi}\mathfrak{R}^{\pi}_{K,n}(\epsilon) &\overset{(a)}{\geq} \inf_{\pi\in\Pi} \sup_{\mu\in\{\mu_-,\mu_+\}} \sup_{\nu\in\mathcal{U}_\epsilon(\mu)} \mathcal{R}_n(\pi,\mu,\nu,\ldots,\nu) \\ &\overset{(b)}{\geq} \inf_{\pi\in\Pi} \sup_{\mu\in\{\mu_-,\mu_+\}} \mathcal{R}_n(\pi,\mu,\tilde{\nu},\ldots,\tilde{\nu}) \\ &= \inf_{\pi\in\Pi} \sup_{\mu\in\{\mu_-,\mu_+\}} \text{opt}(\mu) - \mathbb{E}_{\xi_i\sim\tilde{\nu}}\left[\mathcal{G}\left(\pi(n,\hat{\mu}_{\boldsymbol{\xi^n}}),\mu\right)\right] \\ &\overset{(c)}{=} \inf_{\alpha\in\Delta(\mathcal{X})} \sup_{\mu\in\{\mu_-,\mu_+\}} \text{opt}(\mu) - \mathcal{G}(\alpha,\mu) \overset{(d)}{\geq} \sup_{\beta\in\Delta(\{\mu_-,\mu_+\})} \inf_{x\in\mathcal{X}} \mathbb{E}_{\mu\sim\beta}\left[\text{opt}(\mu) - \mathcal{G}(x,\mu)\right]\end{aligned}$$

where (a) follows from the fact that we restrict the space of out-of-sample distributions to be $\{\mu_-, \mu_+\}$ and we restrict historical distributions $\nu_1, \ldots, \nu_n$ to be all equal. (b) holds as $\tilde{\nu} \in \mathcal{U}_\epsilon(\mu)$ for both μ_- and μ_+. We obtain (c) by remarking that when $\tilde{\nu}$ is fixed, the policy does not need to look at the data as it does not give any extra information on the value of μ. Indeed, for every $\pi \in \Pi$, we construct a probability measure $\alpha \in \Delta(\mathcal{X})$ which satisfies for every $x \in \mathcal{X}$,

$$\alpha(x) = \mathbb{P}_{\xi_i \sim \tilde{\nu}} \left(\pi(n, \hat{\mu}_{\boldsymbol{\xi}^n}) = x\right).$$

We then have for every $\mu \in \{\mu_-, \mu_+\}$,

$$\mathbb{E}_{\xi_i \sim \tilde{\nu}} \left[\mathcal{G}\left(\pi(n, \hat{\mu}_{\boldsymbol{\xi}^n}), \mu\right)\right] - \text{opt}(\mu) = \mathbb{E}_{x \sim \alpha} \left[\mathcal{G}\left(x, \mu\right)\right] - \text{opt}(\mu) = \mathcal{G}\left(\alpha, \mu\right) - \text{opt}(\mu).$$

By taking the supremum over μ, we obtain,

$$\sup_{\mu \in \{\mu_-, \mu_+\}} \mathbb{E}_{\xi_i \sim \tilde{\nu}} \left[\mathcal{G}\left(\pi(n, \hat{\mu}_{\boldsymbol{\xi}^n}), \mu\right)\right] - \text{opt}(\mu) = \sup_{\mu \in \{\mu_-, \mu_+\}} \mathcal{G}\left(\alpha, \mu\right) - \text{opt}(\mu)$$

which implies the inequality,

$$\inf_{\pi \in \Pi} \sup_{\mu \in \{\mu_-, \mu_+\}} \mathbb{E}_{\xi_i \sim \tilde{\nu}} \left[\mathcal{G}\left(\pi(n, \hat{\mu}_{\boldsymbol{\xi}^n}), \mu\right)\right] - \text{opt}(\mu) \geq \inf_{\alpha \in \Delta(\mathcal{X})} \sup_{\mu \in \{\mu_-, \mu_+\}} \mathcal{G}\left(\alpha, \mu\right) - \text{opt}(\mu).$$

In turn, the inequality,

$$\inf_{\pi \in \Pi} \sup_{\mu \in \{\mu_-, \mu_+\}} \mathbb{E}_{\xi_i \sim \tilde{\nu}} \left[\mathcal{G}\left(\pi(n, \hat{\mu}_{\boldsymbol{\xi}^n}), \mu\right)\right] - \text{opt}(\mu) \leq \inf_{\alpha \in \Delta(\mathcal{X})} \sup_{\mu \in \{\mu_-, \mu_+\}} \mathcal{G}\left(\alpha, \mu\right) - \text{opt}(\mu)$$

holds by inclusion.

Finally, (d) follows from Yao's minimax principle.

Let $\tilde{\beta}$ be the prior which puts mass $\frac{1}{2}$ on μ_+ and mass $\frac{1}{2}$ on μ_-.

Since the optimal price must be in the support of the willingness to pay distribution, we only need to consider the revenue generated by posting prices $\frac{M}{2}$ or M. In that case, we have

$$\text{opt}(\mu_+) - \mathcal{G}\left(\frac{M}{2}, \mu_+\right) = M\left(\frac{1}{2} + \epsilon\right) - \frac{M}{2} = M \cdot \epsilon, \quad \text{opt}(\mu_-) - \mathcal{G}\left(\frac{M}{2}, \mu_-\right) = 0,$$

$$\text{opt}(\mu_+) - \mathcal{G}\left(M, \mu_+\right) = 0 \quad \text{and} \quad \text{opt}(\mu_-) - \mathcal{G}\left(M, \mu_-\right) = \frac{M}{2} - M\left(\frac{1}{2} - \epsilon\right) = M \cdot \epsilon.$$

Therefore,

$$\begin{aligned}\inf_{\pi \in \Pi} \mathfrak{R}^{\pi}_{K,n}(\epsilon) &\geq \inf_{x \in \mathcal{X}} \mathbb{E}_{\mu \sim \tilde{\beta}} \left[\text{opt}(\mu) - \mathcal{G}\left(x, \mu\right)\right] \\ &= \min\left(\mathbb{E}_{\mu \sim \tilde{\beta}} \left[\text{opt}(\mu) - \mathcal{G}\left(\frac{M}{2}, \mu\right)\right], \mathbb{E}_{\mu \sim \tilde{\beta}} \left[\text{opt}(\mu) - \mathcal{G}\left(M, \mu\right)\right]\right) = \frac{M \cdot \epsilon}{2}.\end{aligned}$$

We conclude the proof of the lower bound by taking the limsup as n goes to ∞.

□

Proof of Proposition 2. Fix $\eta > 0$. We show that there exists an instance of the data-driven pricing problem under Wasserstein heterogeneity such that the regret of SAA for that instance is larger than $M - \eta$.

We consider the out-of-sample distribution $\mu := \delta_{M-\eta}$, that is to say the probability measure which puts all mass at $M - \eta$. Furthermore, we let $\nu := \delta_{\min(M, M-\eta+\epsilon)}$. Remark that by definition of the Wasserstein distance, we have that,

$$d_W(\mu, \nu) = \min\left(M, M - \eta + \epsilon\right) - (M - \eta) \leq \epsilon.$$

It suffices to show that the asymptotic regret of SAA under our candidate instance is larger than $M - \eta$. We note that for any $\xi_0 \in \Xi$ the oracle of the point mass distribution δ_{ξ_0} is given by $\texttt{ORACLE}\left(\delta_{\xi_0}\right) = \xi_0$, which generates a revenue of ξ_0. Furthermore, we have that,

$$\begin{aligned}\mathbb{E}_{\xi_i \sim \nu} \left[\mathcal{G}\left(\pi^{\text{SAA}}(n, \hat{\mu}_{\boldsymbol{\xi}^n}), \mu\right)\right] &\overset{(a)}{=} \mathcal{G}\left(\pi^{\text{SAA}}(n, \nu), \mu\right) \\ &= \mathcal{G}\left(\min\left(M, M - \eta + \epsilon\right), \mu\right) \\ &= \min\left(M, M - \eta + \epsilon\right) \cdot \mathbb{1}\left\{\min\left(M, M - \eta + \epsilon\right) \leq M - \eta\right\} = 0,\end{aligned}$$

where (a) holds because for every sample size n, the historical samples $\boldsymbol{\xi^n}$ sampled from ν are almost surely all equal to $\min(M, M-\eta+\epsilon)$.

This shows that SAA generates a revenue equal to 0 because it posts a price that is higher than the actual willingness to pay of the customers. We finally conclude that,

$$\mathcal{R}_n\left(\pi^{\text{SAA}}, \mu, \nu, \ldots, \nu\right) = \text{opt}(\mu) - \mathbb{E}_{\xi_i \sim \nu}\left[\mathcal{G}\left(\pi^{\text{SAA}}(n, \hat{\mu}_{\boldsymbol{\xi^n}}), \mu\right)\right] = M - \eta.$$

By taking the supremum over $\eta > 0$ we obtain that,

$$\mathfrak{R}^{\pi^{\text{SAA}}}_{W,n}(\epsilon) \geq M.$$

And by taking the limsup as n goes to ∞, we conclude that $\mathfrak{R}^{\pi^{\text{SAA}}}_{W,\infty}(\epsilon) \geq M$. We complete the proof by noting that $\mathfrak{R}^{\pi^{\text{SAA}}}_{W,\infty}(\epsilon) \leq M$ as the revenue of the oracle is bounded above by M and the one of SAA is bounded below by 0.

□

Proof of Proposition 3. In what follows, we associate to the measure μ (resp. ν) its cumulative distribution F (resp. H).

We have that,

$$\begin{aligned}
\mathcal{G}(x_2, \nu) - \mathcal{G}(x_1, \mu) &= x_2 \cdot (1 - H(x_2)) - x_1 \cdot (1 - F(x_1)) \\
&= x_2 \cdot (F(x_1) - H(x_2)) + (x_2 - x_1) \cdot (1 - F(x_1)) \\
&\overset{(a)}{\leq} (x_2 - x_1) \cdot (1 - F(x_1)) + x_2 \cdot \frac{d_W(\mu, \nu)}{x_2 - x_1} \leq (x_2 - x_1) + M \cdot \frac{d_W(\mu, \nu)}{x_2 - x_1},
\end{aligned}$$

where (a) is a consequence of Lemma 1. □

Proof of Theorem 3. *Step 1:* Fix $\delta < 0$. We first show the following upper bound on the asymptotic worst-case regret of $\pi^{\text{SAA}(\delta)}$.

$$\mathfrak{R}^{\pi^{\text{SAA}(\delta)}}_{W,\infty}(\epsilon) \leq 2\sqrt{M \cdot \epsilon} + |\delta| + M \cdot \frac{\epsilon}{|\delta|} \tag{15}$$

To prove this result, we first bound the quantity $\sup_{\mu \in \mathcal{P}} \sup_{\nu \in \mathcal{U}_\epsilon(\mu)} |\text{opt}(\mu) - \mathcal{G}\left(\pi^{\text{SAA}(\delta)}(\nu), \mu\right)|$ and then apply Theorem 1.

Consider $\mu, \nu \in \mathcal{P}$. We have that,

$$\begin{aligned}
\text{opt}(\mu) - \mathcal{G}\left(\pi^{\text{SAA}(\delta)}(\nu), \mu\right) &= (\text{opt}(\mu) - \text{opt}(\nu)) + \left(\text{opt}(\nu) - \mathcal{G}\left(\pi^{\text{SAA}(\delta)}(\nu), \mu\right)\right) \\
&\overset{(a)}{\leq} 2\sqrt{M \cdot d_W(\mu, \nu)} + |\delta| + M \cdot \frac{d_W(\mu, \nu)}{|\delta|}.
\end{aligned} \tag{16}$$

where (a) follows from the following lemma.

Lemma 4. *Consider the pricing problem. Fix $\delta < 0$. For every μ and $\nu \in \mathcal{P}$, we have that,*

$$\begin{aligned}
\textit{opt}(\mu) - \textit{opt}(\nu) &\leq 2\sqrt{M \cdot d_W(\mu, \nu)}, \\
\textit{opt}(\nu) - \mathcal{G}\left(\pi^{\textit{SAA}(\delta)}(\nu), \mu\right) &\leq |\delta| + M \cdot \frac{d_W(\mu, \nu)}{|\delta|}.
\end{aligned}$$

By taking the supremum in (16), we obtain that

$$\sup_{\mu \in \mathcal{P}} \sup_{\nu \in \mathcal{U}_\epsilon(\mu)} |\text{opt}(\mu) - \mathcal{G}\left(\pi^{\text{SAA}(\delta)}(\nu), \mu\right)| \leq 2\sqrt{M \cdot \epsilon} + |\delta| + M \cdot \frac{\epsilon}{|\delta|}. \tag{17}$$

We conclude that (15) holds by applying Theorem 1 and by using the upper bound derived in (17) on the uniform DRO problem.

Furthermore, we remark that $\delta = -\sqrt{M \cdot \epsilon}$ minimizes the RHS of (15) and yields the bound,

$$\mathfrak{R}^{\pi^{\text{SAA}(\delta)}}_{W,\infty}(\epsilon) \leq 4\sqrt{M \cdot \epsilon}.$$

This shows the first part of the theorem.

Step 2: We next provide a universal lower bound across all possible data-driven policies. Fix $\epsilon > 0$ and let, $\tilde{\nu} = \delta_{\frac{M}{2}}$ the probability measure which puts all the mass at $\frac{M}{2}$. Let $\tilde{\mu}$ be the probability measure defined as,

$$\tilde{\mu}\left(\xi\right) := \begin{cases} 2\sqrt{\frac{\epsilon}{M}} & \text{if } \xi = \frac{M}{2} - \frac{\sqrt{M\cdot\epsilon}}{2} \\ 1 - 2\sqrt{\frac{\epsilon}{M}} & \text{if } \xi = \frac{M}{2}. \end{cases}$$

For a fixed $n \geq 1$, we have that,

$$\begin{aligned}
\inf_{\pi\in\Pi} \mathfrak{R}^{\pi}_{W,n}(\epsilon) &\overset{(a)}{\geq} \inf_{\pi\in\Pi} \sup_{\mu\in\{\tilde{\mu},\tilde{\nu}\}} \sup_{\nu\in\mathcal{U}_{\epsilon}(\mu)} \mathcal{R}_n\left(\pi,\mu,\nu,\ldots,\nu\right) \\
&\overset{(b)}{\geq} \inf_{\pi\in\Pi} \sup_{\mu\in\{\tilde{\mu},\tilde{\nu}\}} \mathcal{R}_n\left(\pi,\mu,\tilde{\nu},\ldots,\tilde{\nu}\right) \\
&= \inf_{\pi\in\Pi} \sup_{\mu\in\{\tilde{\mu},\tilde{\nu}\}} \mathbb{E}_{\xi_i\sim\tilde{\nu}}\left[\text{opt}(\mu) - \mathcal{G}\left(\pi(n,\hat{\mu}_{\boldsymbol{\xi^n}}),\mu\right)\right] \\
&\overset{(c)}{=} \inf_{\alpha\in\Delta(\mathcal{X})} \sup_{\mu\in\{\tilde{\mu},\tilde{\nu}\}} \text{opt}(\mu) - \mathcal{G}\left(\alpha,\mu\right) \overset{(d)}{\geq} \sup_{\beta\in\Delta(\{\tilde{\mu},\tilde{\nu}\})} \inf_{x\in\mathcal{X}} \mathbb{E}_{\mu\sim\beta}\left[\text{opt}(\mu) - \mathcal{G}\left(x,\mu\right)\right]
\end{aligned}$$

where (a) follows from the fact that we restrict the space of out-of-sample distributions to be $\{\tilde{\mu},\tilde{\nu}\}$ and we restrict historical distributions $\nu_1,\ldots,\nu_n$ to be all equal. (b) holds as $\tilde{\nu}\in\mathcal{U}_{\epsilon}(\mu)$ for both $\tilde{\mu}$ and $\tilde{\nu}$. We obtain (c) by remarking that when $\tilde{\nu}$ is fixed, the policy does not need to look at the data as it does not give any extra information on the value of μ. (d) follows from Yao's minimax principle.

Let $\tilde{\beta}$ be the prior which puts mass $\frac{1}{2}$ on $\tilde{\nu}$ and mass $\frac{1}{2}$ on $\tilde{\mu}$. Then, for every $x\in[0,\frac{M}{2}-\frac{\sqrt{M\cdot\epsilon}}{2}]$, we have that,

$$\mathbb{E}_{\mu\sim\tilde{\beta}}\left[\text{opt}(\mu) - \mathcal{G}\left(x,\mu\right)\right] = \mathbb{E}_{\mu\sim\tilde{\beta}}\left[\text{opt}(\mu)\right] - x = \frac{M}{2} - \frac{\sqrt{M\cdot\epsilon}}{4} - x \overset{(a)}{\geq} \frac{\sqrt{M\cdot\epsilon}}{4},$$

where (a) holds because $x \leq \frac{M}{2} - \frac{\sqrt{M\cdot\epsilon}}{2}$. Furthermore, for every $x\in\left(\frac{M}{2}-\frac{\sqrt{M\cdot\epsilon}}{2},\frac{M}{2}\right]$, we have that,

$$\begin{aligned}
\mathbb{E}_{\mu\sim\tilde{\beta}}\left[\text{opt}(\mu) - \mathcal{G}\left(x,\mu\right)\right] &= \frac{M}{2} - \frac{\sqrt{M\cdot\epsilon}}{4} - \mathbb{E}_{\mu\sim\tilde{\beta}}\left[\mathcal{G}\left(x,\mu\right)\right] \\
&= \frac{M}{2} - \frac{\sqrt{M\cdot\epsilon}}{4} - \left(1-\sqrt{\frac{\epsilon}{M}}\right)\cdot x \geq \frac{\sqrt{M\cdot\epsilon}}{4}.
\end{aligned}$$

We note that, for $x > \frac{M}{2}$, we have that $\mathbb{E}_{\mu\sim\tilde{\beta}}\left[\text{opt}(\mu) - \mathcal{G}\left(x,\mu\right)\right] = \frac{M}{2} - \frac{\sqrt{M\cdot\epsilon}}{4}$ because, $\mathcal{G}\left(x,\mu\right) = 0$ when $\mu = \tilde{\mu}$ and $\mu = \tilde{\nu}$.

Therefore,

$$\inf_{\pi\in\Pi} \mathfrak{R}^{\pi}_{W,n}(\epsilon) \geq \sup_{\beta\in\Delta(\{\tilde{\mu},\tilde{\nu}\})} \inf_{x\in\mathcal{X}} \mathbb{E}_{\mu\sim\beta}\left[\text{opt}(\mu) - \mathcal{G}\left(x,\mu\right)\right] \geq \frac{\sqrt{M\cdot\epsilon}}{4}.$$

And by taking the limit we obtain the desired result,

$$\inf_{\pi\in\Pi} \limsup_{n\to\infty} \mathfrak{R}^{\pi}_{W,n}(\epsilon) \overset{(a)}{=} \limsup_{n\to\infty} \inf_{\pi\in\Pi} \mathfrak{R}^{\pi}_{W,n}(\epsilon) \geq \frac{\sqrt{M\cdot\epsilon}}{4},$$

where (a) holds because π observes n.

□

Proof of Lemma 4. *Step 1:* We first prove that, $\text{opt}\left(\mu\right) - \text{opt}\left(\nu\right) \leq 2\sqrt{M\cdot d_W\left(\mu,\nu\right)}$.

If $\texttt{ORACLE}\left(\mu\right) \leq \sqrt{M\cdot d_W\left(\mu,\nu\right)}$, we have that

$$\text{opt}\left(\mu\right) - \text{opt}\left(\nu\right) \leq \mathcal{G}\left(\texttt{ORACLE}\left(\mu\right),\mu\right) \overset{(a)}{\leq} \texttt{ORACLE}\left(\mu\right) \leq \sqrt{M\cdot d_W\left(\mu,\nu\right)},$$

where (a) holds for every $x \in \mathcal{X}$ and any distribution $\tilde{\mu}$, we have that $\mathcal{G}(x, \tilde{\mu}) \leq x$ for the pricing problem.

Assume now that, $\texttt{ORACLE}(\mu) > \sqrt{M \cdot d_W(\mu, \nu)}$. Let $\tilde{x} := \texttt{ORACLE}(\mu) - \sqrt{M \cdot d_W(\mu, \nu)}$. We have,

$$\begin{aligned}
\text{opt}(\mu) - \text{opt}(\nu) &\leq \mathcal{G}(\texttt{ORACLE}(\mu), \mu) - \mathcal{G}(\tilde{x}, \nu) \\
&\overset{(a)}{\leq} \texttt{ORACLE}(\mu) - \tilde{x} + M \cdot \frac{d_W(\mu, \nu)}{\texttt{ORACLE}(\mu) - \tilde{x}} \\
&= 2\sqrt{M \cdot d_W(\mu, \nu)},
\end{aligned}$$

where (a) follows from Proposition 3.

Step 2: We now show that, $\text{opt}(\nu) - \mathcal{G}\left(\pi^{\text{SAA}(\delta)}(\nu), \mu\right) \leq |\delta| + M \cdot \frac{d_W(\mu,\nu)}{|\delta|}$.

Similarly to step 1, we note that if, $\texttt{ORACLE}(\nu) \leq |\delta|$, we have that

$$\text{opt}(\nu) - \mathcal{G}\left(\pi^{\text{SAA}(\delta)}(\nu), \mu\right) \leq \text{opt}(\nu) \leq \texttt{ORACLE}(\nu) \leq |\delta|.$$

Furthermore, if $\texttt{ORACLE}(\nu) > |\delta|$, then

$$\text{opt}(\nu) - \mathcal{G}\left(\pi^{\text{SAA}(\delta)}(\nu), \mu\right) = \mathcal{G}(\texttt{ORACLE}(\nu), \nu) - \mathcal{G}(\texttt{ORACLE}(\nu) - |\delta|, \mu) \overset{(a)}{\leq} |\delta| + M \cdot \frac{d_W(\mu, \nu)}{|\delta|},$$

where (a) follows from Proposition 3.

□

D.2 Additional results and proofs of Section 5.2

Lemma 5. *Consider the ski-rental cost. For every $\mu \in \mathcal{P}$ associated with the cumulative distribution F, we have for every $x \in \mathcal{X}$ that,*

$$\mathcal{G}(x, \mu) = b \cdot (1 - F(x)) + x - \int_0^x F(\xi) d\xi.$$

Proof of Lemma 5. For every $\mu \in \mathcal{P}$ and $x \in \mathcal{X}$, we have that,

$$\begin{aligned}
\mathcal{G}(x, \mu) &= \mathbb{E}_{\xi \sim \mu}[\xi \cdot \mathbb{1}\{\xi \leq x\}] + (b + x) \cdot (1 - F(x)) \\
&\overset{(a)}{=} \left(x \cdot F(x) - \int_0^x F(\xi) d\xi\right) + (b + x) \cdot (1 - F(x)) \\
&= b \cdot (1 - F(x)) + x - \int_0^x F(\xi) d\xi,
\end{aligned}$$

where (a) follows from the Riemman-Stieljes integration by part. □

Lemma 6. *Consider the ski-rental cost. For every $\mu \in \mathcal{P}$ and every $x_1, x_2 \in \mathcal{X}$ such that $x_1 < x_2$ we have that,*

$$\mathcal{G}(x_2, \mu) - \mathcal{G}(x_1, \mu) \leq x_2 - x_1.$$

Proof of Lemma 6. Let $\mu \in \mathcal{P}$ and let $x_1, x_2 \in \mathcal{X}$ such that $x_1 \leq x_2$. We have,

$$\mathcal{G}(x_2, \mu) - \mathcal{G}(x_1, \mu) \overset{(a)}{=} b \cdot (F(x_1) - F(x_2)) + x_2 - x_1 - \int_{x_1}^{x_2} F(\xi) d\xi \overset{(b)}{\leq} x_2 - x_1,$$

where (a) follows from Lemma 5 and (b) holds because F is non-negative and non-decreasing. □

Lemma 7. *For the ski-rental problem, we have that for every $\mu, \nu \in \mathcal{P}$ and any $x \in \mathcal{X}$,*

$$|\mathcal{G}(x, \mu) - \mathcal{G}(x, \nu)| \leq (M + b) \cdot d_K(\mu, \nu) \qquad \textit{and} \qquad |\mathcal{G}(x, \mu) - \mathcal{G}(x, \nu)| \leq b + d_W(\mu, \nu)$$

Proof of Lemma 7. For every $\mu, \nu \in \mathcal{P}$, such that μ (resp. ν) is associated to the cumulative distribution F (resp. H) we have that,

$$\begin{aligned}
|\mathcal{G}(x,\mu) - \mathcal{G}(x,\nu)| &\overset{(a)}{=} \left| b \cdot (H(x) - F(x)) + \int_0^x (H(\xi) - F(\xi))\, d\xi \right| \\
&\leq b \cdot |H(x) - F(x)| + \int_0^x |H(\xi) - F(\xi)| d\xi \\
&\leq b \cdot d_K(\mu,\nu) + \int_{\xi \in \Xi} d_K(\mu,\nu)\, d\xi = (M+b) \cdot d_K(\mu,\nu)
\end{aligned}$$

and

$$\begin{aligned}
|\mathcal{G}(x,\mu) - \mathcal{G}(x,\nu)| &\overset{(a)}{=} \left| b \cdot (H(x) - F(x)) + \int_0^x (H(\xi) - F(\xi))\, d\xi \right| \\
&\leq b \cdot |H(x) - F(x)| + \int_0^x |H(\xi) - F(\xi)| d\xi \overset{(b)}{\leq} b + d_W(\mu,\nu),
\end{aligned}$$

where (a) follows from Lemma 5 and (b) holds because $d_W(\mu,\nu) = \int_{\xi \in \Xi} |H(\xi) - F(\xi)| d\xi$. □

Proposition 8. *Consider the ski-rental problem. Let μ and $\nu \in \mathcal{P}$ be two different distributions and let x_1 and x_2 be two different actions such that $x_1 < x_2$. We have that,*

$$\mathcal{G}(x_2,\mu) - \mathcal{G}(x_1,\nu) \leq b \cdot \frac{d_W(\mu,\nu)}{x_2 - x_1} + d_W(\mu,\nu) + x_2 - x_1.$$

Proof of Proposition 8. Let $x_1 < x_2$. Let F (resp. H) be the cdf associated to μ (resp. ν). We have,

$$\begin{aligned}
\mathcal{G}(x_2,\mu) - \mathcal{G}(x_1,\nu) &\overset{(a)}{=} b \cdot (H(x_1) - F(x_2)) + x_2 - x_1 + \int_0^{x_1} (F(\xi) - H(\xi))\, d\xi - \int_{x_1}^{x_2} F(\xi) d\xi \\
&\leq b \cdot (H(x_1) - F(x_2)) + x_2 - x_1 + \int_0^M |F(\xi) - H(\xi)| d\xi \\
&\overset{(b)}{\leq} b \cdot \frac{d_W(\mu,\nu)}{x_2 - x_1} + x_2 - x_1 + \int_0^M |F(\xi) - H(\xi)| d\xi \\
&\overset{(c)}{=} b \cdot \frac{d_W(\mu,\nu)}{x_2 - x_1} + x_2 - x_1 + d_W(\mu,\nu),
\end{aligned}$$

where (a) follows from Lemma 5 and (b) holds by Lemma 1 and (c) follows from the definition of the Wasserstein distance. □

Proof of Theorem 4. We first characterize the performance of SAA.

Step 1-a: We derive an upper bound on the worst-case asymptotic regret of SAA by considering the uniform DRO problem. According to Theorem 1, it is sufficient to show the following upper bound

$$\sup_{\mu \in \mathcal{P}} \sup_{\nu \in \mathcal{U}_\epsilon(\mu)} |\mathrm{opt}(\mu) - \mathcal{G}\left(\pi^{\mathrm{SAA}}(\nu), \mu\right)| \leq 2b + 2\epsilon. \tag{18}$$

For every $\mu, \nu \in \mathcal{P}$, we have that,

$$\begin{aligned}
\mathcal{G}\left(\pi^{\mathrm{SAA}}(\nu), \mu\right) - \mathrm{opt}(\mu) &\overset{(a)}{=} \mathcal{G}(\texttt{ORACLE}(\nu), \mu) - \mathrm{opt}(\mu) \\
&= (\mathcal{G}(\texttt{ORACLE}(\nu), \mu) - \mathrm{opt}(\nu)) \\
&\quad + (\mathrm{opt}(\nu) - \mathcal{G}(\texttt{ORACLE}(\mu), \nu)) + (\mathcal{G}(\texttt{ORACLE}(\mu), \nu) - \mathrm{opt}(\mu)) \\
&\overset{(b)}{\leq} (b + d_W(\mu,\nu)) + 0 + (b + d_W(\mu,\nu)) \\
&= 2b + 2d_W(\mu,\nu),
\end{aligned}$$

where (a) follows by noting that $\pi^{\mathrm{SAA}}(\nu) = \texttt{ORACLE}(\nu)$ by definition and (b) holds by applying Lemma 7 to the first and third difference. By taking supremums over μ and ν we obtain the inequality (18). Furthermore, Theorem 1 and equation (18) imply that,

$$\mathfrak{R}^{\pi^{\mathrm{SAA}}}_{W,\infty}(\epsilon) \leq 2b + 2\epsilon.$$

Step 1-b: Next, we derive a lower bound on the worst-case asymptotic regret of SAA. Let μ and ν be the probability measures defined,

$$\nu(\xi) = \begin{cases} \frac{3}{4} & \text{if } \xi = \frac{b}{2} \\ \frac{1}{4} & \text{if } \xi = M \end{cases} \qquad \text{and} \qquad \mu(\xi) = \begin{cases} \frac{3}{4} & \text{if } \xi = \frac{b}{2} + \epsilon \\ \frac{1}{4} & \text{if } \xi = M. \end{cases}$$

By construction, we have that $d_W(\mu, \nu) = \frac{3}{4}\epsilon \leq \epsilon$.

Fix $n \geq 1$ and let $\boldsymbol{\xi^n}$ be samples from ν. Let $\hat{p}_n$ be the fraction of samples equal to M. Formally,

$$\hat{p}_n = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}\{\xi_i = M\}.$$

Remark that $\hat{p}_n$ is a sufficient statistic to describe the empirical cdf $\hat{\mu}_{\boldsymbol{\xi^n}}$. We next claim that $\pi^{\text{SAA}}(\hat{\mu}_{\boldsymbol{\xi^n}}) = \frac{b}{2}$ if and only if, $\hat{p}_n < \frac{1}{2}$. Without loss of optimality one may restrict attentions to one of the following actions: purchasing at day 0, never purchasing (i.e. purchasing at the end of day M) or purchasing after a skiing day. Furthermore, by using the expression of the ski-rental cost, we remark that, $\mathcal{G}\left(\frac{b}{2}, \hat{\mu}_{\boldsymbol{\xi^n}}\right) < \mathcal{G}(0, \hat{\mu}_{\boldsymbol{\xi^n}})$ if and only if $\frac{b}{2} + \hat{p}_n \cdot b < b$ and we have $\mathcal{G}\left(\frac{b}{2}, \hat{\mu}_{\boldsymbol{\xi^n}}\right) < \mathcal{G}(M, \hat{\mu}_{\boldsymbol{\xi^n}})$ whenever $M > 2 \cdot b$ (independently of $\hat{p}_n$). Therefore, the action $\frac{b}{2}$ yields the smallest cost if and only if $\hat{p}_n < \frac{1}{2}$.

Let E_n be the event, $\{\hat{p}_n < \frac{1}{2}\}$. We note that,

$$\begin{aligned} \mathbb{P}(E_n) &\geq \mathbb{P}\left(\left|\hat{p}_n - \frac{1}{4}\right| < \frac{1}{4}\right) \overset{(a)}{=} 1 - \mathbb{P}\left(|\hat{p}_n - \mathbb{E}[\hat{p}_n]| \geq \frac{1}{4}\right) \\ &\overset{(b)}{\geq} 1 - 16\text{Var}[\hat{p}_n] = 1 - \frac{16}{n} \cdot \text{Var}[\mathbb{1}\{\xi = M\}] = 1 - \frac{3}{n}, \end{aligned} \tag{19}$$

where (a) follows from the fact that $\mathbb{E}[\hat{p}_n] = \frac{1}{4}$ and (b) is a consequence of Chevyshev's inequality.

We finally conclude that,

$$\begin{aligned} \mathfrak{R}^{\pi^{\text{SAA}}}_{W,n}(\epsilon) &\overset{(a)}{\geq} \mathcal{R}_n\left(\pi^{\text{SAA}}, \mu, \nu, \ldots, \nu\right) \\ &= \mathbb{E}_{\xi_i \sim \nu}\left[\mathcal{G}\left(\pi^{\text{SAA}}(\hat{\mu}_{\boldsymbol{\xi^n}}), \mu\right)\right] - \text{opt}(\mu) \\ &\overset{(b)}{\geq} \left[\mathcal{G}\left(\pi^{\text{SAA}}(\hat{\mu}_{\boldsymbol{\xi^n}}), \mu\right) \,\middle|\, E_n\right] \cdot \mathbb{P}(E_n) - \mathcal{G}(0, \mu) \\ &\overset{(c)}{=} \mathbb{P}(E_n)\, \mathcal{G}\left(\frac{b}{2}, \mu\right) - \mathcal{G}(0, \mu) = \frac{3}{2}\mathbb{P}(E_n) \cdot b - b, \end{aligned}$$

where (a) follows from the fact that $d_W(\mu, \nu) \leq \epsilon$ and (b) holds because the function g is non-negative and because $\text{opt}(\mu) \leq \mathcal{G}(0, \mu)$. (c) holds because under the event E_n, $\pi^{\text{SAA}}(\hat{\mu}_{\boldsymbol{\xi^n}}) = \frac{b}{2}$.

FInally, by taking the limsup we obtain that,

$$\limsup_{n \to \infty} \mathfrak{R}^{\pi^{\text{SAA}}}_{W,n}(\epsilon) \geq \limsup_{n \to \infty} \frac{3}{2}\mathbb{P}(E_n) \cdot b - b \overset{(a)}{=} \frac{b}{2},$$

where (a) follows from (19).

We now characterize the rate of the optimal achievable performance.

Step 2-a: Fix $\delta > 0$. We first show the following upper bound on the asymptotic worst-case regret of $\pi^{\text{SAA}(\delta)}$.

$$\mathfrak{R}^{\pi^{\text{SAA}(\delta)}}_{W,\infty}(\epsilon) \leq b \cdot \frac{\epsilon}{\delta} + 2\epsilon + \delta + 2\sqrt{b \cdot \epsilon}. \tag{20}$$

To prove this result, we first bound the quantity $\sup_{\mu \in \mathcal{P}} \sup_{\nu \in \mathcal{U}_\epsilon(\mu)} |\text{opt}(\mu) - \mathcal{G}\left(\pi^{\text{SAA}(\delta)}(\nu), \mu\right)|$ and then apply Theorem 1.

Consider $\mu, \nu \in \mathcal{P}$. We have that,

$$\begin{aligned} \mathcal{G}\left(\pi^{\text{SAA}(\delta)}(\nu), \mu\right) - \text{opt}(\mu) &= \mathcal{G}\left(\pi^{\text{SAA}(\delta)}(\nu), \mu\right) - \text{opt}(\nu) + \text{opt}(\nu) - \text{opt}(\mu) \\ &\overset{(a)}{\leq} b \cdot \frac{d_W(\mu, \nu)}{\delta} + 2d_W(\mu, \nu) + \delta + 2\sqrt{b \cdot d_W(\mu, \nu)}, \end{aligned} \tag{21}$$

where (a) follows from the following lemma.

Lemma 8. *Consider the ski-rental problem. Fix $\delta > 0$. For every μ and $\nu \in \mathcal{P}$, we have that,*

$$opt\,(\mu) - opt\,(\nu) \leq 2\sqrt{b \cdot d_W\,(\mu,\nu)} + d_W\,(\mu,\nu)\,,$$

$$\mathcal{G}\left(\pi^{SAA(\delta)}(\nu), \mu\right) - opt(\nu) \leq b \cdot \frac{d_W\,(\mu,\nu)}{\delta} + d_W\,(\mu,\nu) + \delta.$$

By taking the supremum in equation (21), we obtain,

$$\sup_{\mu\in\mathcal{P}} \sup_{\nu\in\mathcal{U}_\epsilon(\mu)} |\text{opt}(\mu) - \mathcal{G}\left(\pi^{\text{SAA}(\delta)}(\nu), \mu\right)| \leq b \cdot \frac{\epsilon}{\delta} + 2\epsilon + \delta + 2\sqrt{b\cdot\epsilon}. \tag{22}$$

We conclude that (20) holds by applying Theorem 1 and by using the upper bound derived in (22) on the uniform DRO problem.

Furthermore, we remark that $\delta = \sqrt{b\cdot\epsilon}$ minimizes the RHS of (20) and yields the bound,

$$\mathfrak{R}^{\pi^{\text{SAA}(\delta)}}_{W,\infty}(\epsilon) \leq 4\sqrt{b\cdot\epsilon} + 2\epsilon.$$

This concludes the proof of the upper bound on the asymptotic worst-case regret of the optimal policy.

Step 2-b: We next provide a universal lower bound across all possible data-driven policies. Fix $0 < \epsilon \leq \frac{1}{2}$. Assume $b \geq 1$. Let $\tilde{\nu}$ and $\tilde{\mu}$ be the probability measures defined,

$$\tilde{\nu}(\xi) = \begin{cases} \frac{1}{2} & \text{if } \xi = \frac{b}{2} - \frac{\sqrt{b\cdot\epsilon}}{2} \\ \frac{1}{2} & \text{if } \xi = M \end{cases} \qquad \text{and} \qquad \tilde{\mu}(\xi) = \begin{cases} \frac{1}{2} - \sqrt{\frac{\epsilon}{b}} & \text{if } \xi = \frac{b}{2} - \frac{\sqrt{b\cdot\epsilon}}{2} \\ \sqrt{\frac{\epsilon}{b}} & \text{if } \xi = \frac{b}{2} + \frac{\sqrt{b\cdot\epsilon}}{2} \\ \frac{1}{2} & \text{if } \xi = M. \end{cases}$$

By construction, we have that $d_W\,(\tilde{\mu},\tilde{\nu}) = \epsilon$. For a fixed $n \geq 1$, we have that,

$$\begin{aligned}
\inf_{\pi\in\Pi} \mathfrak{R}^{\pi}_{W,n}(\epsilon) &\overset{(a)}{\geq} \inf_{\pi\in\Pi} \sup_{\mu\in\{\tilde{\mu},\tilde{\nu}\}} \sup_{\nu\in\mathcal{U}_\epsilon(\mu)} \mathcal{R}_n\,(\pi,\mu,\nu,\ldots,\nu) \\
&\overset{(b)}{\geq} \inf_{\pi\in\Pi} \sup_{\mu\in\{\tilde{\mu},\tilde{\nu}\}} \mathcal{R}_n\,(\pi,\mu,\tilde{\nu},\ldots,\tilde{\nu}) \\
&= \inf_{\pi\in\Pi} \sup_{\mu\in\{\tilde{\mu},\tilde{\nu}\}} \mathbb{E}_{\xi_i\sim\tilde{\nu}}\left[\mathcal{G}\left(\pi(n,\hat{\mu}_{\boldsymbol{\xi}^n}),\mu\right) - \text{opt}(\mu)\right] \\
&\overset{(c)}{=} \inf_{\alpha\in\Delta(\mathcal{X})} \sup_{\mu\in\{\tilde{\mu},\tilde{\nu}\}} \mathcal{G}\,(\alpha,\mu) - \text{opt}(\mu) \overset{(d)}{\geq} \sup_{\beta\in\Delta(\{\tilde{\mu},\tilde{\nu}\})} \inf_{x\in\mathcal{X}} \mathbb{E}_{\mu\sim\beta}\left[\mathcal{G}\,(x,\mu) - \text{opt}(\mu)\right]
\end{aligned}$$

where (a) follows from the fact that we restrict the space of out-of-sample distributions to be $\{\tilde{\mu},\tilde{\nu}\}$ and we restrict historical distributions $\nu_1,\ldots,\nu_n$ to be all equal. (b) holds as $\tilde{\nu} \in \mathcal{U}_\epsilon(\mu)$ for both $\tilde{\mu}$ and $\tilde{\nu}$. We obtain (c) by remarking that when $\tilde{\nu}$ is fixed, the policy does not need to look at the data as it does not give any extra information on the value of μ. (d) follows from Yao's minimax principle.

Without loss of optimality one may restrict attentions to one of the following actions: purchasing at day 0 or purchasing after a skiing day. We remark that,

$$\mathcal{G}\,(0,\tilde{\nu}) = b, \quad \mathcal{G}\left(\frac{b}{2} - \frac{\sqrt{b\cdot\epsilon}}{2}, \tilde{\nu}\right) = b - \frac{\sqrt{b\cdot\epsilon}}{2}, \quad \mathcal{G}\left(\frac{b}{2} + \frac{\sqrt{b\cdot\epsilon}}{2}, \tilde{\nu}\right) = b,$$

$$\mathcal{G}\,(0,\tilde{\mu}) = b, \quad \mathcal{G}\left(\frac{b}{2} - \frac{\sqrt{b\cdot\epsilon}}{2}, \tilde{\mu}\right) = b + \frac{\sqrt{b\cdot\epsilon}}{2} \quad \text{and} \quad \mathcal{G}\left(\frac{b}{2} + \frac{\sqrt{b\cdot\epsilon}}{2}, \tilde{\mu}\right) = b + \epsilon.$$

As a consequence, $opt(\tilde{\nu}) = b - \frac{\sqrt{b\cdot\epsilon}}{2}$ and $\text{opt}(\tilde{\mu}) = b$. Let $\tilde{\beta}$ be the prior which puts mass $\frac{1}{2}$ on $\tilde{\nu}$ and mass $\frac{1}{2}$ on $\tilde{\mu}$. We have that,

$$\mathbb{E}_{\mu\sim\tilde{\beta}}\left[\mathcal{G}\,(0,\mu)\right] = b, \quad \mathbb{E}_{\mu\sim\tilde{\beta}}\left[\mathcal{G}\left(\frac{b}{2} - \frac{\sqrt{b\cdot\epsilon}}{2}, \mu\right)\right] = b \text{ and } \mathbb{E}_{\mu\sim\tilde{\beta}}\left[\mathcal{G}\left(\frac{b}{2} + \frac{\sqrt{b\cdot\epsilon}}{2}, \mu\right)\right] = b + \frac{\epsilon}{2}.$$

Therefore,

$$\inf_{\pi\in\Pi} \mathfrak{R}^{\pi}_{W,n}(\epsilon) \geq \inf_{x\in\mathcal{X}} \mathbb{E}_{\mu\sim\tilde{\beta}}\left[\mathcal{G}\,(x,\mu) - \text{opt}(\mu)\right] = \frac{\sqrt{b\cdot\epsilon}}{4}.$$

We conclude the proof of the lower bound by taking the limsup as n goes to ∞. □

Proof of Lemma 8. *Step 1:* We first prove that, $\text{opt}(\mu) - \text{opt}(\nu) \leq 2\sqrt{b \cdot d_W(\mu,\nu)} + d_W(\mu,\nu)$.

First, if $\texttt{ORACLE}(\mu) \geq M - \sqrt{b \cdot d_W(\mu,\nu)}$, we have that

$$\begin{aligned}
\text{opt}(\nu) - \text{opt}(\mu) &= \text{opt}(\nu) - \mathcal{G}(M,\mu) + \mathcal{G}(M,\mu) - \mathcal{G}(\texttt{ORACLE}(\mu),\mu) \\
&\overset{(a)}{\leq} \text{opt}(\nu) - \mathcal{G}(M,\mu) + M - \texttt{ORACLE}(\mu) \\
&\leq \mathcal{G}(M,\nu) - \mathcal{G}(M,\mu) + M - \texttt{ORACLE}(\mu) \\
&\overset{(b)}{=} \int_0^M (F(\xi) - H(\xi))\, d\xi + M - \texttt{ORACLE}(\mu) \leq d_W(\mu,\nu) + \sqrt{b \cdot d_W(\mu,\nu)},
\end{aligned}$$

where (a) follows from Lemma 6 and (b) follows from Lemma 5.

Assume now that, $\texttt{ORACLE}(\mu) < M - \sqrt{b \cdot d_W(\mu,\nu)}$. Let $\tilde{x} := \texttt{ORACLE}(\mu) + \sqrt{b \cdot d_W(\mu,\nu)}$. We have,

$$\begin{aligned}
\text{opt}(\nu) - \text{opt}(\mu) &\leq \mathcal{G}(\tilde{x},\nu) - \mathcal{G}(\texttt{ORACLE}(\mu),\mu) \\
&\overset{(a)}{\leq} \tilde{x} - \texttt{ORACLE}(\mu) + b \cdot \frac{d_W(\mu,\nu)}{\tilde{x} - \texttt{ORACLE}(\mu)} + d_W(\mu,\nu) \\
&= 2\sqrt{b \cdot d_W(\mu,\nu)} + d_W(\mu,\nu),
\end{aligned}$$

where (a) follows from Proposition 8.

Step 2: We now show that, $\mathcal{G}\left(\pi^{\text{SAA}(\delta)}(\nu),\mu\right) - \text{opt}(\nu) \leq b \cdot \frac{d_W(\mu,\nu)}{\delta} + d_W(\mu,\nu) + \delta$.

We note that if, $\texttt{ORACLE}(\nu) \geq M - \delta$, we have that $\pi^{\text{SAA}(\delta)}(\nu)$ plays action M and,

$$\begin{aligned}
\mathcal{G}\left(\pi^{\text{SAA}(\delta)}(\nu),\mu\right) - \text{opt}(\nu) &= \mathcal{G}(M,\mu) - \text{opt}(\nu) \\
&= \mathcal{G}(M,\mu) - \mathcal{G}(M,\nu) + \mathcal{G}(M,\nu) - \mathcal{G}(\texttt{ORACLE}(\nu),\nu) \\
&\overset{(a)}{\leq} d_W(\mu,\nu) + \delta,
\end{aligned}$$

where (a) is obtained by a similar argument to the one developed in step 1. Furthermore, if $\texttt{ORACLE}(\nu) \leq M - \delta$, then

$$\begin{aligned}
\mathcal{G}\left(\pi^{\text{SAA}(\delta)}(\nu),\mu\right) - \text{opt}(\nu) &= \mathcal{G}(\texttt{ORACLE}(\nu) + \delta,\mu) - \mathcal{G}(\texttt{ORACLE}(\nu),\nu) \\
&\overset{(a)}{\leq} b \cdot \frac{d_W(\mu,\nu)}{\delta} + d_W(\mu,\nu) + \delta.
\end{aligned}$$

where (a) follows from Proposition 8.

□

Proof of Theorem 5 . We first characterize the performance of SAA.

Step 1-a: We derive an upper bound on the worst-case asymptotic regret of SAA by using Theorem 1 and by showing the following upper bound

$$\sup_{\mu \in \mathcal{P}} \sup_{\nu \in \mathcal{U}_\epsilon(\mu)} \left|\text{opt}(\mu) - \mathcal{G}\left(\pi^{\text{SAA}}(\nu),\mu\right)\right| \leq 2(M+b) \cdot \epsilon. \tag{23}$$

For every $\mu, \nu \in \mathcal{P}$, we have that,

$$\begin{aligned}
\mathcal{G}\left(\pi^{\text{SAA}}(\nu),\mu\right) - \text{opt}(\mu) &= \mathcal{G}(\texttt{ORACLE}(\nu),\mu) - \text{opt}(\mu) \\
&= \mathcal{G}(\texttt{ORACLE}(\nu),\mu) - \text{opt}(\nu) \\
&\quad + \text{opt}(\nu) - \mathcal{G}(\texttt{ORACLE}(\mu),\nu) + \mathcal{G}(\texttt{ORACLE}(\mu),\nu) - \text{opt}(\mu) \\
&\overset{(a)}{\leq} (M+b) \cdot d_K(\mu,\nu) + 0 + (M+b) \cdot d_K(\mu,\nu) \\
&= 2(M+b) \cdot d_K(\mu,\nu),
\end{aligned}$$

where (a) holds by applying Lemma 7 to the first and third difference. By taking supremums over μ and ν we obtain the inequality (23). Furthermore, Theorem 1 and equation (23) imply that,

$$\mathfrak{R}_{K,\infty}^{\pi^{\text{SAA}}}(\epsilon) \leq 2(M+b) \cdot \epsilon$$

Step 1-b: Next, we derive a lower bound on the worst-case asymptotic regret of SAA. Assume that $M \in \mathbb{N}^*$. For a fixed $p \in (0,1)$, let ν be the probability measure defined as,

$$\nu(k) = \frac{(1-p)^{k-1}p}{\alpha} \qquad \text{for all } k \in \{1, \dots, M\},$$

where $\alpha = p \cdot \sum_{j=1}^{M}(1-p)^{j-1}$.

Fix $\delta > 0$. Given observed samples $\boldsymbol{\xi^n}$ of ν, the decision of SAA is completely characterized by the empirical cdf $\hat{\nu}$. Consider the event,

$$E_n = \{\hat{\nu}(k) \in [(1-\delta)\nu(k), (1+\delta)\nu(k)], \text{ for all } k \in \{0, \dots, M\}\}$$

We next show that if $p > \frac{1}{b}\left(\frac{1+\delta}{1-\delta}\right)$ and the event E_n holds, then SAA never buys the skis, i.e., it buys at the end of day M.

Remark that SAA never buys if for each day k, buying at the end of day k leads to a larger cost than the cost to go buy letting renting for one more day. Formally SAA does not buy if for every $k \in \{1, \dots, M-1\}$,

$$b > 1 + \mathbb{P}_{\hat{\nu}}[\xi \geq k+1 \,|\, \xi \geq k] \cdot b \tag{24}$$

Therefore, it suffices to show that if $p > \frac{1}{b}\left(\frac{1+\delta}{1-\delta}\right)$ and the event E_n holds, equation (24) is satisfied. Remark that,

$$\begin{aligned}
\mathbb{P}_{\hat{\nu}}[\xi \geq k+1 \,|\, \xi \geq k] &= \frac{\mathbb{P}_{\hat{\nu}}[\xi \geq k+1]}{\mathbb{P}_{\hat{\nu}}[\xi \geq k]} \\
&= \frac{\sum_{j=k+1}^{M} \hat{\nu}(j)}{\sum_{j=k}^{M} \hat{\nu}(j)} \\
&= 1 - \frac{\hat{\nu}(k)}{\sum_{j=k}^{M} \hat{\nu}(j)} \\
&\overset{(a)}{\leq} 1 - \frac{\nu(k)}{\sum_{j=k}^{M} \nu(j)} \cdot \frac{1-\delta}{1+\delta} \\
&= 1 - \frac{(1-p)^{k-1}}{\sum_{j=k}^{M}(1-p)^{j-1}} \cdot \frac{1-\delta}{1+\delta} \\
&= 1 - \frac{p}{1-(1-p)^{M-k}} \cdot \frac{1-\delta}{1+\delta} \leq 1 - p \cdot \frac{1-\delta}{1+\delta},
\end{aligned}$$

where (a) holds because the event E_n hold. Therefore,

$$1 + \mathbb{P}_{\hat{\nu}}[\xi \geq k+1 \,|\, \xi \geq k] \cdot b \leq 1 + \left(1 - p \cdot \frac{1-\delta}{1+\delta}\right) \cdot b = b + 1 - p \cdot \frac{1-\delta}{1+\delta} \cdot b \overset{(a)}{<} b,$$

where (a) follows from the fact that $p > \frac{1}{b}\left(\frac{1+\delta}{1-\delta}\right)$. Therefore, (24) holds. In what follows we keep $p > \frac{1}{b}\left(\frac{1+\delta}{1-\delta}\right)$.

Let μ be a probability measure with cdf F and denote by H the cdf of ν. Set $F(\xi) = \max(H(\xi) - \epsilon, 0)$ for all ξ (and $F(M) = H(M) = 1$). By construction, we have $d_K(\mu, \nu) = \epsilon$. For every $n \geq 1$, we have that,

$$\mathfrak{R}_{K,n}^{\pi^{\text{SAA}}}(\epsilon) \geq \mathcal{R}_n(\pi, \mu, \nu, \dots, \nu) \overset{(a)}{\geq} (\mathcal{G}(M, \mu) - \mathcal{G}(0, \mu)) \cdot \mathbb{P}(E_n), \tag{25}$$

where (a) follows from the fact that under the event E_n, SAA never buys and by bounding the cost of the oracle by the one obtained by purchasing at day 0.

We next show that one can choose ν and μ such that $d_K(\mu, \nu) \leq \epsilon$ and,

$$\mathcal{G}(M, \mu) - \mathcal{G}(0, \mu) \geq M \cdot \epsilon + b\frac{1-\delta}{1+\delta+b\cdot(1-\delta)\cdot\delta} - b - \frac{1}{2} \tag{26}$$

Indeed, note that,

$$\mathcal{G}\left(M,\mu\right)-\mathcal{G}\left(0,\mu\right)\overset{(a)}{=}\mathbb{E}_{\mu}[\xi]-b=\mathbb{E}_{\mu}[\xi]-\mathbb{E}_{\nu}[\xi]+\mathbb{E}_{\nu}[\xi]-b$$

First remark that for M large enough ν is close to a geometric distribution with parameter b therefore, its mean is close to $\frac{1}{p}$. In particularly, for M large enough, we have that $\left|\mathbb{E}_{\nu}[\xi]-\frac{1}{p}\right|\leq\frac{1}{2}$. Furthermore if $\epsilon<p$ we have that for every $\xi\in\{1,\ldots M-1\}$, $H(\xi)>\epsilon$ and thus $F(\xi)=H(\xi)-\epsilon$. This implies that, $\mathbb{E}_{\mu}[\xi]=\mathbb{E}_{\nu}[\xi]+\epsilon\cdot M$. As a consequence we have for M large enough that

$$\mathcal{G}\left(M,\mu\right)-\mathcal{G}\left(0,\mu\right)=\mathbb{E}_{\mu}[\xi]-\mathbb{E}_{\nu}[\xi]+\mathbb{E}_{\nu}[\xi]-b\geq M\cdot\epsilon+\frac{1}{p}-b-\frac{1}{2}.$$

Setting $p=\frac{1}{b}\frac{1+\delta}{1-\delta}+\delta$ we obtain the relation (26).

By combining equations (25) and (26), we obtain that,

$$\begin{aligned}\limsup_{n\to\infty}\mathfrak{R}_{K,n}^{\pi^{\text{SAA}}}(\epsilon)&\geq\limsup_{n\to\infty}\left(M\cdot\epsilon+b\frac{1-\delta}{1+\delta+b\cdot(1-\delta)\cdot\delta}-b-\frac{1}{2}\right)\mathbb{P}\left(E_n\right)\\&\overset{(a)}{=}M\cdot\epsilon+b\frac{1-\delta}{1+\delta+b\cdot(1-\delta)\cdot\delta}-b-\frac{1}{2},\end{aligned}$$

where (a) is a consequence of the multiplicative Chernoff bound combined with a union bound. Since this holds for any $\delta>0$ we get by taking δ to 0 that,

$$\limsup_{n\to\infty}\mathfrak{R}_{K,n}^{\pi^{\text{SAA}}}(\epsilon)\geq M\cdot\epsilon-\frac{1}{2}.$$

Step 2-a: We derive the upper bound on the asymptotic worst-case regret of the optimal data-driven algorithm by relying on the bound provided by [19] for their algorithm in the DRO case. They show that their Algorithm 1 is able to achieve a uniform DRO regret $\mathcal{O}\left(b\cdot\epsilon\cdot\log\left(\epsilon^{-1}\right)\right)$. As a consequence by applying Theorem 1 we obtain that,

$$\inf_{\pi\in\Pi}\mathfrak{R}_{K,\infty}^{\pi}(\epsilon)=\mathcal{O}\left(b\cdot\epsilon\cdot\log\left(\epsilon^{-1}\right)\right)$$

Step 2-b: We next provide a universal lower bound across all possible data-driven policies. Fix $0<\epsilon\leq\frac{1}{2}$. Assume $b\geq1$. Let $\tilde{\nu}$ and $\tilde{\mu}$ be the probability measures defined,

$$\tilde{\nu}(\xi)=\begin{cases}\frac{1}{2}+\frac{\epsilon}{2} & \text{if }\xi=\frac{3}{4}b\\ \frac{1}{2}-\frac{\epsilon}{2} & \text{if }\xi=\frac{5}{4}b\end{cases}\qquad\text{and}\qquad\tilde{\mu}(\xi)=\begin{cases}\frac{1}{2}-\frac{\epsilon}{2} & \text{if }\xi=\frac{3}{4}b\\ \frac{1}{2}+\frac{\epsilon}{2} & \text{if }\xi=\frac{5}{4}b\end{cases}$$

By construction, we have that $d_K\left(\mu,\nu\right)=\epsilon$. For a fixed $n\geq1$, we have that,

$$\begin{aligned}\inf_{\pi\in\Pi}\mathfrak{R}_{K,n}^{\pi}(\epsilon)&\overset{(a)}{\geq}\inf_{\pi\in\Pi}\sup_{\mu\in\{\tilde{\mu},\tilde{\nu}\}}\sup_{\nu\in\mathcal{U}_{\epsilon}(\mu)}\mathcal{R}_n\left(\pi,\mu,\nu,\ldots,\nu\right)\\&\overset{(b)}{\geq}\inf_{\pi\in\Pi}\sup_{\mu\in\{\tilde{\mu},\tilde{\nu}\}}\mathcal{R}_n\left(\pi,\mu,\tilde{\nu},\ldots,\tilde{\nu}\right)\\&=\inf_{\pi\in\Pi}\sup_{\mu\in\{\tilde{\mu},\tilde{\nu}\}}\mathbb{E}_{\xi_i\sim\tilde{\nu}}\left[\text{opt}(\mu)-\mathcal{G}\left(\pi(n,\hat{\mu}_{\boldsymbol{\xi}^n}),\mu\right)\right]\\&\overset{(c)}{=}\inf_{\alpha\in\Delta(\mathcal{X})}\sup_{\mu\in\{\tilde{\mu},\tilde{\nu}\}}\text{opt}(\mu)-\mathcal{G}\left(\alpha,\mu\right)\overset{(d)}{\geq}\sup_{\beta\in\Delta(\{\tilde{\mu},\tilde{\nu}\})}\inf_{x\in\mathcal{X}}\mathbb{E}_{\mu\sim\beta}\left[\text{opt}(\mu)-\mathcal{G}\left(x,\mu\right)\right]\end{aligned}$$

where (a) follows from the fact that we restrict the space of out-of-sample distributions to be $\{\tilde{\mu},\tilde{\nu}\}$ and we restrict historical distributions $\nu_1,\ldots,\nu_n$ to be all equal. (b) holds as $\tilde{\nu}\in\mathcal{U}_{\epsilon}(\mu)$ for both $\tilde{\mu}$ and $\tilde{\nu}$. We obtain (c) by remarking that when $\tilde{\nu}$ is fixed, the policy does not need to look at the data as it does not give any extra information on the value of μ. (d) follows from Yao's minimax principle.

Furthermore, remark that,

$$\mathcal{G}\left(0,\tilde{\nu}\right)=b\qquad\text{and}\qquad\mathcal{G}\left(\frac{3}{4}b,\tilde{\nu}\right)=b+b\cdot\left(\frac{1}{4}-\frac{\epsilon}{2}\right)\qquad\text{and}\qquad\mathcal{G}\left(\frac{5}{4}b,\tilde{\nu}\right)=b-\frac{\epsilon\cdot b}{4}$$

$$\mathcal{G}\left(0, \tilde{\mu}\right) = b \qquad \text{and} \qquad \mathcal{G}\left(\frac{3}{4}b, \tilde{\mu}\right) = b + b \cdot \left(\frac{1}{4} + \frac{\epsilon}{2}\right) \qquad \text{and} \qquad \mathcal{G}\left(\frac{5}{4}b, \tilde{\mu}\right) = b + \frac{\epsilon \cdot b}{4}.$$

As a consequence, $opt(\tilde{\nu}) = b - \frac{\epsilon \cdot b}{4}$ and $\text{opt}(\tilde{\mu}) = b$. Let $\tilde{\beta}$ be the prior which puts mass $\frac{1}{2}$ on $\tilde{\nu}$ and mass $\frac{1}{2}$ on $\tilde{\mu}$. We have that,

$$\mathbb{E}_{\mu \sim \beta}\left[\mathcal{G}\left(0, \mu\right)\right] = b \qquad \text{and} \qquad \mathbb{E}_{\mu \sim \beta}\left[\mathcal{G}\left(\frac{3}{4}b, \mu\right)\right] = \frac{9}{8}b \qquad \text{and} \qquad \mathbb{E}_{\mu \sim \beta}\left[\mathcal{G}\left(\frac{5}{4}b, \mu\right)\right] = b.$$

Therefore,

$$\inf_{\pi \in \Pi} \mathfrak{R}^{\pi}_{K,n}(\epsilon) \geq \inf_{x \in \mathcal{X}} \mathbb{E}_{\mu \sim \tilde{\beta}}\left[\mathcal{G}\left(x, \mu\right) - \text{opt}(\mu)\right] = \frac{\epsilon \cdot b}{8}.$$

We conclude the proof of the lower bound by taking the limsup as n to ∞. □