PRETRAIN-THEN-CALIBRATE: ENHANCING OPTIMALITY-FEASIBILITY TRADEOFF IN SAFE CONTEXTUAL OPTIMIZATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Contextual optimization, or decision-focused learning, aims to learn context-based decision rules for downstream decision-making tasks. In addition to maintaining statistical efficiency like in conventional supervised learning, this problem entails delicate interplays with optimization tractability and the effectiveness of data-optimization integration. In this paper, we consider *safe* contextual optimization where, in addition to an optimality criterion, the decision rule also satisfies a high-probability-enforced risk criterion. Natural attempts to integrate chance-constraint methods such as robust optimization, with established contextual optimization pipelines such as predict-then-optimize, suffers from a high price of optimality in exchange for an adequate feasibility guarantee. This motivates us to propose a *pretrain-then-calibrate* framework to significantly enhance the optimality-feasibility tradeoff. Specifically, our framework translates any heuristic decision rules into rigorous statistically feasible rules via what we call a *feasibility calibrator*, constructed from a careful interpolation against a safe reference rule to possess certifiable joint feasibility-optimality guarantees. At a high level, our approach exploits the low-dimensional interpolation from readily obtainable good heuristics, and is advantageously model-agnostic, modular and consequently implementation-friendly. In synthetic and real-data experiments on operations management and portfolio optimization, pretrain-and-calibrate results in solutions significantly more optimal than a range of benchmarks including conformal and (distributionally) robust optimization methods without sacrificing feasibility, thus achieving demonstrably improved optimality-feasibility tradeoffs.

1 INTRODUCTION

Contextual optimization builds data-to-decision pipelines that optimize decisions based on pre-observed covariate or contextual information (Sadana et al., 2025). This problem, also known as decision-focused learning, can be viewed as a generalization of classical supervised learning, with downstream optimization performance as the solution quality criterion. At a high level, the goal is to obtain a data-driven *decision rule* that maps context to a good decision; for example, an inventory rule maps demand signals to an order, and a portfolio rule maps market information to an asset allocation. Because of the focus on downstream decisions, contextual optimization entails training and error-tolerance considerations different from those of conventional prediction, and faces new challenges arising from computational tractability in optimization, statistical learning efficiency, and their interplay (Elmachtoub & Grigas, 2022; El Balghiti et al., 2023).

In this paper, we especially focus on *safe* contextual optimization. Namely, we seek a decision rule that not only performs well according to an optimality criterion, but also satisfies a constraint that typically represents a risk criterion and is itself subject to uncertainty. A natural way to formulate such problems is through a so-called probabilistic or chance constraint, i.e., the decision needs to satisfy an uncertain "safety" condition with high probability. Chance-constrained models have been applied to inventory management (Charnes et al., 1958), portfolio selection (Agnew et al., 1969), supply-chain network design (Li et al., 2017), transportation and logistics (Sluijk et al., 2022), facility location and emergency service planning (Peng et al., 2020), energy-system operation

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

and maintenance (Porras et al., 2023), production and capacity planning (Li et al., 2013), natural-resource management (Lejeune & Kettunen, 2017), and healthcare policy (Mildebrath et al., 2023).

While there has been an array of approaches, chance-constrained optimization is, first of all, generally intractable: even for simple (such as linear) safety conditions, the associated chance constraint can easily be nonconvex unless very restrictive distributional assumptions hold (Ben-Tal et al., 2009). To this end, the majority of the literature aims to find solutions that are feasible with statistical confidence while avoiding excessive conservatism, which can worsen objective performance. A common method is robust optimization (Bertsimas et al., 2011), which considers the worst-case parameter value over an uncertainty or ambiguity set that captures a high-likelihood region of the uncertainty. While this uncertainty set can be learned and calibrated from data (Hong et al., 2021), the paradigm of robust optimization separates the objective function from uncertainty-set construction, which often makes the "protection" against uncertainty too strong and subsequently leads to overly conservative solutions. Another popular approach is scenario optimization, which uses samples to approximate the constraint, for example through an adequate number of sampled constraints (Campi & Garatti, 2018), empirical violation probability (Luedtke & Ahmed, 2008) or more generally sampling-and-discarding (Campi & Garatti, 2011; Garatti & Campi, 2022). Both robust optimization and scenario optimization have a range of variants designed to reduce conservatism, primarily by tuning hyperparameters such as the set size and the number of scenarios.

Despite the above progress, extending these methods to contextual settings presents substantial challenges in attaining a desirable *optimality–feasibility tradeoff*. The major approaches in contextual optimization, including predict-then-optimize (i.e., train a prediction model to feed directly into the downstream objective), integrated estimation-and-optimization (i.e., train a prediction model using the downstream objective as the loss function) (Qi et al., 2026), and end-to-end learning (i.e., train the decision rule as a parameterized map) (Zhang et al., 2026; Qin et al., 2025), all focus on prediction–optimization integration at the objective level. When deterministic constraints are present, we can use a Lagrangian reformulation (Fioretto et al., 2020), smooth the discontinuous indicators (Peña-Ordieres et al., 2020; Chamon & Ribeiro, 2020), or, for the first two approaches above, run constrained optimization using the trained predictor. When uncertainty is present in the constraint, however, these techniques do not automatically ensure feasibility and, to this end, the guarantees associated with established chance-constraint methods and those associated with contextual optimization methods are generally incompatible. For instance, suppose we use robust optimization as the engine to obtain solutions with chance-constraint guarantees. A naive adoption of the resulting uncertainty set, by incorporating the context in the set with a high-probability conditional prediction coverage (if such a set is at all learnable) (Patel et al., 2024; Wang et al., 2023; Im et al., 2025), will tend to be conservative. That is, achieving statistical feasibility might come with a high cost of optimality.

Main Idea and Contributions. Motivated by the above challenges, we propose a *pretrain-then-calibrate* framework that, at a high level, converts a pretrained decision rule under any feasibility heuristic into a rigorously feasibility-guaranteed rule. To explain this, note that, while it is difficult to obtain contextual feasibility guarantees due to the methodological incompatibility described above, obtaining a good heuristic decision is relatively much easier to do. These heuristics include, importantly, end-to-end decision rules that directly leverage modern neural network architectures and other AI models. The challenge mostly lies in the high optimality price to pay if one needs to refine the training algorithm directly to attain rigorous feasibility. In this regard, we take a novel path of *not* refining the training algorithm directly, but instead conducting a light-weighted "boosting" of any pretrained feasibility-heuristic decision rule to equip it with concrete confidence guarantees. In this way, the price on optimality is much smaller than that in the existing direct algorithmic approaches.

More specifically, our framework calibrates a pretrained decision rule via two ingredients. First is the interpolation with a reference rule that is always safe. In many common problems, we will see that such a reference rule naturally exists. Second is a *feasibility calibrator* to obtain the value of the interpolating parameter. Depending on the type of target statistical guarantee, we provide four corresponding statistically efficient calibrators: a Dvoretzky–Kiefer–Wolfowitz (DKW) calibrator has a distribution-free high confidence guarantee independent of both the complexity of the pretrained decision-rule class and the number of interpolated parameters; two (normalized and unnormalized) Gaussian calibrators use covariance to obtain high confidence data-dependent

feasibility guarantees; and a conformal calibrator with an in-expectation feasibility guarantee. For these calibrators, we establish joint optimality-feasibility guarantees along the interpolation path.

Our approach has several significant benefits that we will illustrate in the sequel. First, it is *statistically efficient* compared to directly modified training algorithms because we exploit the low-dimensionality of the interpolation path. Importantly, this dimension is very carefully chosen so that a monotonicity pattern is ensured: The decision rule is always "more" feasible as the parameter varies towards the reference rule, which then guarantees and facilitates the search of an ultimate statistically feasible decision. Second, it is *agnostic* to the feasibility-heuristic, in the sense that the interpolation and the calibrator works whatever heuristic decision we begin with. Third, it is easy to use in practice, since our methodology is entirely *modular* and can be applied simply as an add-on to the heuristic. In numerical experiments, our feasibility calibrator significantly outperforms existing benchmarks on real-world datasets drawn from newsvendor and portfolio optimization problems.

We give a more detailed literature review in Appendix A, including contextual optimization, chance-constrained optimization, conformal and (distributionally) robust approaches, and constrained machine learning.

2 METHODOLOGY

We formulate the safe contextual optimization problem and describe our pretrain-then-calibrate framework. Consider a decision variable $\boldsymbol{w} \in \Omega \subseteq \mathbb{R}^{d_w}$ where Ω is convex, an uncertain parameter $\boldsymbol{z} \in \mathcal{Z} \subseteq \mathbb{R}^{d_z}$, and an observed context $\boldsymbol{x} \in \mathcal{X} \subseteq \mathbb{R}^{d_x}$. Let $c(\boldsymbol{w}, \boldsymbol{z})$ be the cost and $g(\boldsymbol{w}, \boldsymbol{z}) \leq 0$ the constraint. We aim to find a decision rule $\boldsymbol{w}(\cdot) \in \mathcal{H}$ solving

$$\begin{aligned} \min_{\boldsymbol{w}(\cdot)\in\mathcal{H}} \quad & \mathbb{E}[c(\boldsymbol{w}(\boldsymbol{x}), \boldsymbol{z})] \\ \text{s.t.} \quad & \mathbb{P}\{g(\boldsymbol{w}(\boldsymbol{x}), \boldsymbol{z}) \leq 0\} \geq 1-\alpha, \end{aligned} \tag{1}$$

where $\mathcal{H}$ is the decision-rule class and $\alpha \in (0,1)$ is the allowed violation probability. The expectation and probability are taken over the unknown joint distribution of $(\boldsymbol{x}, \boldsymbol{z})$. A joint requirement on several constraints can be expressed through $g = \max_i g_i$. For example, in portfolio optimization, $g(\boldsymbol{w}, \boldsymbol{z}) \leq 0$ represents that an asset allocation rule triggers a large loss, and $-c(\boldsymbol{w}, \boldsymbol{z})$ denotes the portfolio return.

Pretraining. Given training data $\mathcal{D}_n := \{(\boldsymbol{x}_i, \boldsymbol{z}_i)\}_{i=1}^n$, assume we already get a pretrained rule $\boldsymbol{w}^{\text{pre}} : \mathcal{X} \to \Omega$ from some procedure, such as empirical constrained optimization, smoothing, Lagrangian, and predict-then-optimize. This pretrained rule is supposedly good in terms of optimality performance thanks to the strengths of these existing methods (In Appendix B.1, we present the performance guarantee for empirical constrained optimization as an example). We denote $\mathscr{T}$ as the sigma-field generated by training data, which we need to describe further our framework.

Reference rule and interpolation path. We assume there is a known reference rule $\boldsymbol{w}_0 : \mathcal{X} \to \Omega$ satisfying $g(\boldsymbol{w}_0(\boldsymbol{x}), \boldsymbol{z}) \leq 0$ almost surely. In practice, it can be obtained by combining known support bounds with a feasible decision that satisfies the constraint over that support. Its construction depends on the application and can use context-specific bounds when available. We construct a collection of convex combinations

$$\boldsymbol{w}_\lambda^{\text{tune}}(\boldsymbol{x}) = (1-\lambda)\boldsymbol{w}^{\text{pre}}(\boldsymbol{x}) + \lambda \boldsymbol{w}_0(\boldsymbol{x}), \qquad \lambda \in [0,1]. \tag{2}$$

We call λ the *safety parameter*. We construct a finite grid $0 \leq \lambda_1 < \cdots < \lambda_p = 1$ and set

$$\boldsymbol{w}_j := \boldsymbol{w}_{\lambda_j}^{\text{tune}}, \quad R_j := \mathbb{P}\left(g\left(\boldsymbol{w}_j\left(\boldsymbol{x}\right), \boldsymbol{z}\right) \leq 0 \mid \mathscr{T}\right), \quad C_j := \mathbb{E}\left[c\left(\boldsymbol{w}_j\left(\boldsymbol{x}\right), \boldsymbol{z}\right) \mid \mathscr{T}\right].$$

We call $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_p$ *path candidates*. Here R_j is the population constraint-satisfaction probability and C_j is its expected cost evaluated at $\boldsymbol{w}_j$. The final path candidate $\boldsymbol{w}_p = \boldsymbol{w}_0$ is the reference rule.

Intuitively, the interpolation path provides a smooth transition on the decision space from the pretrained rule, which supplies objective performance, to a reference rule that provides an almost-surely safe endpoint. The safety parameter λ measures the adjustment toward the reference rule. Convexity of Ω keeps the interpolated rule in the decision space, and convexity of $g(\cdot, \boldsymbol{z})$

makes constraint satisfaction *nested* as the safety parameter λ increases (formalized in Section 3.1). Calibration can therefore evaluate a one-dimensional family and select a proper safety adjustment. The interpolation path (1) explicitly exposes the optimality–feasibility tradeoff between the pretrained and reference rules; and (2) is computationally efficient since it directly works on the decision space and does not update pretrained model parameters.

Feasibility calibration. We use i.i.d. validation data $\mathcal{D}_{\text{val}} := \left\{\left(\boldsymbol{x}_i^{\text{val}}, \boldsymbol{z}_i^{\text{val}}\right)\right\}_{i=1}^{n_{\text{val}}}$ to select a proper safety parameter $\lambda \in \{\lambda_j\}_{j=1}^{p}$ using our four calibrators introduced in Section 2.1. For validation sample i and the candidate j, let $Y_{ij} := \mathbf{1}\left\{g\left(\boldsymbol{w}_j\left(\boldsymbol{x}_i^{\text{val}}\right), \boldsymbol{z}_i^{\text{val}}\right) \le 0\right\}$ and define

$$\left(\widehat{R}_j, \widehat{C}_j\right) := n_{\text{val}}^{-1} \sum_{i=1}^{n_{\text{val}}} \left(Y_{ij}, c\left(\boldsymbol{w}_j\left(\boldsymbol{x}_i^{\text{val}}\right), \boldsymbol{z}_i^{\text{val}}\right)\right), \quad \widehat{\Sigma}_{jk} := n_{\text{val}}^{-1} \sum_{i=1}^{n_{\text{val}}} Y_{ij} Y_{ik} - \widehat{R}_j \widehat{R}_k, \quad \widehat{\sigma}_j^2 := \widehat{\Sigma}_{jj}. \tag{3}$$

Here $\widehat{R}_j$ and $\widehat{C}_j$ are the empirical validation counterparts of R_j and C_j. Matrix $\widehat{\Sigma}$ is the empirical covariance matrix of constraint-satisfaction indicators, and $\widehat{\sigma}_j^2$ is empirical variance for candidate j.

2.1 Four Calibrators

We design four feasibility calibrators as follows: conformal risk calibrator (C), DKW calibrator (D), normalized Gaussian calibrator (N), and unnormalized Gaussian calibrator (U), targeting on different requirements and guarantees. We denote feasibility confidence as $\beta \in (0, 1/2)$.

Conformal risk calibration (C). We apply conformal risk control (Angelopoulos et al., 2024) to the binary violation indicators $1 - Y_{ij}$. When $\alpha \ge 1/(n_{\text{val}} + 1)$, select $j_{\mathsf{C}}^* := \min\left\{j \in [p] : (1 + \sum_{i=1}^{n_{\text{val}}}(1 - Y_{ij}))/(n_{\text{val}} + 1) \le \alpha\right\}$. The reference rule makes this set nonempty. Selecting the first passing candidate gives the smallest adjustment satisfying the criterion. This choice controls expected violation; its objective value depends on how cost varies along the path. If $\alpha < 1/(n_{\text{val}} + 1)$, set $j_{\mathsf{C}}^* = p$.

Next, the three methods $\mathsf{m} \in \{\mathsf{D}, \mathsf{N}, \mathsf{U}\}$ share the path selection procedure as described in Algorithm 1. That is, they select the candidate with the smallest validated cost that achieves the lower confidence bound. For each method m, the confidence bound depends on a parameter M_j^{m} defined as follows.

Algorithm 1 High-Confidence Path Selection

Input: $\{\widehat{R}_j, \widehat{C}_j, M_j^{\mathsf{m}}\}_{j \in [p]}$, where M_j^{m} is defined in (4) and (5) respectively

Step 1: $\widehat{\mathcal{J}}_{\mathsf{m}} \leftarrow \left\{j \in [p] : \widehat{R}_j - M_j^{\mathsf{m}} \ge 1 - \alpha\right\}$.

Step 2: If $\widehat{\mathcal{J}}_{\mathsf{m}} \neq \varnothing$, choose $j_{\mathsf{m}}^* \in \operatorname{argmin}_{j \in \widehat{\mathcal{J}}_{\mathsf{m}}} \widehat{C}_j$; otherwise, set $j_{\mathsf{m}}^* \leftarrow p$.

Output: $\boldsymbol{w}_{j_{\mathsf{m}}^*}$.

DKW calibration (D). This method applies when $Y_{i1} \le \cdots \le Y_{ip} = 1$ (which holds when g is convex, formalized by Assumption 1). Each validation sample i therefore has a first-satisfaction index, and $\widehat{R}_j$ is the empirical distribution function of these indices evaluated at j. The one-sided DKW inequality (Massart, 1990) controls this empirical distribution simultaneously at all candidates. For every $j \in [p]$, we compute the common margin independently of grid point number:

$$M_j^{\mathsf{D}} := \sqrt{\log(1/\beta)/(2n_{\text{val}})}. \tag{4}$$

Gaussian calibration (N and U). These calibrators apply the Gaussian validation schemes of (Lam & Qian, 2019; Chen et al., 2021) to the constraint indicators of decision rules. Gaussian calibration uses the estimated covariance of these indicators to account for their dependence. The normalized calibrator scales each margin by the candidate's empirical standard deviation, and the unnormalized calibrator uses a common margin. For every $j \in [p]$, we compute the margins:

$$M_j^{\mathsf{N}} := q_{1-\beta}^{\mathsf{N}} \widehat{\sigma}_j / \sqrt{n_{\text{val}}}, \quad M_j^{\mathsf{U}} := q_{1-\beta}^{\mathsf{U}} / \sqrt{n_{\text{val}}}. \tag{5}$$

where $\widehat{\sigma}_j$ is defined in (3), and the quantities $q^{\mathsf{N}}_{1-\beta}, q^{\mathsf{U}}_{1-\beta}$ are computed by the quantiles of the candidate's empirical standard deviation (N) or the probability unit (U). Formally,

$$q^{\mathsf{N}}_{1-\beta} := \inf\left\{q : \mathbb{P}\left(\max_{j\in\widehat{\mathcal{J}}_+} \frac{Z_j}{\widehat{\sigma}_j} \le q \,\middle|\, \mathcal{D}_{\text{val}}, \mathscr{T}\right) \ge 1-\beta\right\}, \qquad q^{\mathsf{U}}_{1-\beta} := \inf\left\{q : \mathbb{P}\left(\max_{j\in[p]} Z_j \le q \,\middle|\, \mathcal{D}_{\text{val}}, \mathscr{T}\right) \ge 1-\beta\right\}.$$

Here, $Z \sim \mathcal{N}\left(0, \widehat{\Sigma}\right)$ with $\widehat{\Sigma}$ defined in (3), and $\widehat{\mathcal{J}}_+ := \{j \in [p] : \widehat{\sigma}_j > 0\}$. In the definition of $q^{\mathsf{N}}_{1-\beta}$, if $\widehat{\mathcal{J}}_+ = \varnothing$, we have $q^{\mathsf{N}}_{1-\beta} = 0$. Since $\beta < 1/2$, the values $q^{\mathsf{N}}_{1-\beta}, q^{\mathsf{U}}_{1-\beta}$ and all three margins M^{D}_j, $M^{\mathsf{N}}_j, M^{\mathsf{U}}_j$ are nonnegative. The quantiles, which are depicted as exact conditional quantiles above, can be well approximated by straightforward Gaussian vector simulation in implementation.

Rationale of Feasibility Calibrators. The four calibrators aim to achieve statistical feasibility without strongly degrading optimality via: (1) the pretrained rule, including end-to-end neural network architectures, often already supplies desirable objective performance; and (2) D, N and U are cost-aware calibrators in the sense that they select a λ with the lowest validated cost among those candidates whose validated feasibility reaches $1-\alpha$. The feasibility check is carefully formulated to account for the validation error of order $O(1/\sqrt{n_{\text{val}}})$ by the margin inflation $M^{\mathsf{D}}_j, M^{\mathsf{N}}_j, M^{\mathsf{U}}_j$. On the other hand, C chooses the smallest safety parameter passing a conformalized feasibility criterion. Although C is not cost-aware in its design, its thinner $O(1/n_{\text{val}})$ margin often empirically leads to better optimality, but with a weakened statistical confidence of feasibility.

Relative Strengths of Feasibility Calibrators. Compared to existing literature, our methodology aims for a more desirable statistical-computational balance. Direct constrained learning methods (Fioretto et al., 2020; Chamon & Ribeiro, 2020) focus on coming up with pretrained rules with desirable optimality, but with heuristic feasibility attainments. Projection, hard constraint repair methods (Liang et al., 2024; Constante-Flores et al., 2025), as well as (distributionally) robust optimization approaches (Wang et al., 2023; Patel et al., 2024), are over-conservative due to their worst-case design and consequently sacrifice optimality substantially. In our two-phase pretrain-then-calibrate framework, we do not first look for rigorous feasibility, so that we can leverage good heuristics with strong optimality performance. Then, we utilize validation mechanism as a bridge to effectively adjust the optimality-superior rule. Our interpolation construction that uses a simple convex combination on a low-dimensional path results in significant gain in the statistical effectiveness in attaining feasibility than updating high-dimensional pretrained model parameters.

3 Theoretical Guarantees for Feasibility Calibration

We first establish concretely the aforementioned nesting behavior of constraint satisfaction along the interpolation path. Leveraging this, we develop theoretical guarantees for our four feasibility calibrators, including statistical confidence on feasibility and joint optimality-feasibility guarantees.

3.1 Behavior of Interpolation Path

We assume the convexity of the decision space and the constraint, and the existence of the reference decision rule $\boldsymbol{w}_0(\cdot)$.

Assumption 1 (Convexity and existence of a safe reference decision)**.** *The decision space Ω is convex. For every $\boldsymbol{z} \in \mathcal{Z}$, $g(\cdot, \boldsymbol{z})$ is convex on Ω. Moreover, there exists a function $\boldsymbol{w}_0 : \mathcal{X} \to \Omega$ such that $g(\boldsymbol{w}_0(\boldsymbol{x}), \boldsymbol{z}) \le 0$ almost surely.*

Recall that within interpolation path, $\lambda = 0$ recovers the pretrained rule and $\lambda = 1$ recovers the reference rule. We call the path *nested* when $\{g(\boldsymbol{w}^{\text{tune}}_{\lambda}(\boldsymbol{x}), \boldsymbol{z}) \le 0\} \subseteq \{g(\boldsymbol{w}^{\text{tune}}_{\lambda'}(\boldsymbol{x}), \boldsymbol{z}) \le 0\}$ for every $\lambda \le \lambda'$. Under Assumption 1, we have the following nesting property.

Lemma 1 (Nesting along the interpolation path)**.** *Under Assumption 1, for candidates in Section 3.1, we have $\{g(\boldsymbol{w}_j(\boldsymbol{x}), \boldsymbol{z}) \le 0\} \subseteq \{g(\boldsymbol{w}_k(\boldsymbol{x}), \boldsymbol{z}) \le 0\}$ when $j < k$. Hence, $R_1 \le \cdots \le R_p = 1$.*

3.2 Theoretical Guarantees

We develop guarantees for our four calibrators in Section 2.1: expected violation for C, and high-confidence feasibility and pathwise optimality for D, N, and U.

Theorem 1 (Conformal expected violation calibration)**.** *Under Assumption 1, conditional on $\mathscr{T}$ almost surely,* $\mathbb{E}_{\mathcal{D}_{\text{val}}}[1 - R_{j_{\mathsf{C}}^*} \mid \mathscr{T}] \leq \alpha$.

This is an in-expectation guarantee: over the joint distribution of validation data and a fresh independent pair $(\boldsymbol{x}, \boldsymbol{z})$, the selected rule violates the constraint with probability at most α. Let $\widetilde{\alpha} := \min\{\alpha, 1-\alpha\}$. For universal constants $K_{\mathrm{G}}, k_{\mathrm{G}} > 0$, define

$$\mathfrak{E}_{\mathsf{D}} := 0, \qquad \mathfrak{E}_{\mathsf{U}} := K_{\mathrm{G}} \left(\frac{\log^7(pn_{\text{val}})}{n_{\text{val}}\widetilde{\alpha}} \right)^{1/6} + p \exp\left(-k_{\mathrm{G}} n_{\text{val}} \widetilde{\alpha}\right),$$

$$\mathfrak{E}_{\mathsf{N}} := K_{\mathrm{G}} \left\{ \left(\frac{\log^7(pn_{\text{val}})}{n_{\text{val}}\widetilde{\alpha}} \right)^{1/6} + \frac{\log^2(pn_{\text{val}})}{\sqrt{n_{\text{val}}\widetilde{\alpha}}} + p \exp\left(-k_{\mathrm{G}} \left(n_{\text{val}}\widetilde{\alpha}\right)^{2/3}\right) \right\} + p \exp(-k_{\mathrm{G}} n_{\text{val}} \widetilde{\alpha}).$$

Theorem 2 (Feasibility and pathwise optimality of validation selection)**.** *Fix* $\mathsf{m} \in \{\mathsf{D}, \mathsf{N}, \mathsf{U}\}$. *For* $\mathsf{m} = \mathsf{D}$*, suppose Assumption 1 holds; for* $\mathsf{m} \in \{\mathsf{N}, \mathsf{U}\}$, $pn_{\text{val}} \geq 2$. *Conditional on $\mathscr{T}$, we have:*

(i) Feasibility. $\mathbb{P}_{\mathcal{D}_{\text{val}}}\left(R_{j_{\mathsf{m}}^*} \geq 1 - \alpha \mid \mathscr{T}\right) \geq 1 - \beta - \mathfrak{E}_{\mathsf{m}}$.

(ii) Pathwise optimality. *Suppose* $|c(\boldsymbol{w}, \boldsymbol{z})| \leq B_c$ *for every* $\boldsymbol{w} \in \Omega$ *and* $\boldsymbol{z} \in \mathcal{Z}$. *Let* $\delta \in (0,1)$*, set* $\kappa_{\mathsf{D}} := 1$ *and* $\kappa_{\mathsf{N}} = \kappa_{\mathsf{U}} := p$. *If* $\eta_{\mathsf{m}}^R \leq \alpha$*, then*

$$\mathbb{P}_{\mathcal{D}_{\text{val}}}\left(C_{j_{\mathsf{m}}^*} \leq \min_{j:\, R_j \geq 1-\alpha+\eta_{\mathsf{m}}^R} C_j + \varepsilon_{\text{cost}} \,\middle|\, \mathscr{T} \right) \geq 1 - \delta,$$

where $\eta_{\mathsf{m}}^R := \sqrt{\frac{\log(\kappa_{\mathsf{m}}/\beta)}{2n_{\text{val}}}} + \sqrt{\frac{\log(2\kappa_{\mathsf{m}}/\delta)}{2n_{\text{val}}}}$, $\varepsilon_{\text{cost}} := 2B_c\sqrt{\frac{2\log(4p/\delta)}{n_{\text{val}}}}$.

Part (i) gives D a finite-sample feasibility guarantee at confidence level $1-\beta$; the guarantees for N and U include Gaussian-approximation errors. In part (ii), the margin η_{m}^R covers calibration and estimation errors; for D it does not depend on the number of candidates. Under the assumptions of part (ii), both guarantees hold simultaneously with probability at least $1 - \beta - \mathfrak{E}_{\mathsf{m}} - \delta$. The objective bound compares with the best path candidate whose safety probability exceeds the target by η_{m}^R. This margin ensures that the candidate passes validation with high confidence; $\varepsilon_{\text{cost}}$ accounts for objective statistical error. The guarantee is pathwise: its performance depends on the objective values available along the constructed interpolation path.

Choosing a practical calibrator. For practical use, we recommend normalized Gaussian calibration (N) when high-confidence feasibility guarantee is needed; it provides a favorable balance between objective performance and empirical feasibility in the experiments of Section 4. DKW calibration (D) is an appropriate choice when a finite-sample, distribution-agnostic high-confidence feasibility guarantee is needed. Conformal calibration (C) is appropriate when expected violation control is sufficient. It is generally less conservative, thus often selecting a smaller safety parameter in experiments, but it only attains a weaker in-expectation guarantee than high-confidence calibration.

4 Numerical Experiments

We evaluate our pretrain-then-calibrate framework by showing how calibration changes objective value and violation rate across pretrained decision-rule classes, and compare with existing methods. The main studies use synthetic and real-world newsvendor and portfolio data. Due to the page limitation, we provide more experimental details in Appendix D, including the design of the synthetic data-generating processes (DGP) in Appendix D.1, reference rules in Appendix D.2, and algorithmic implementation in Appendix D.3. We also report two contextual stochastic linear optimization problems, fractional-knapsack and alloy-allocation in Appendix E.

4.1 Experimental Setup

For newsvendor, the decision is an order quantity $\boldsymbol{w}(\boldsymbol{x})$ and $\boldsymbol{z}$ is demand. We minimize expected cost $c_{\text{nv}}(\boldsymbol{w}, \boldsymbol{z}) = (\boldsymbol{w} - \boldsymbol{z})^+ + 3(\boldsymbol{z} - \boldsymbol{w})^+$ subject to a constraint controlling the probability of

stockout (when demand exceeds the order), i.e., $\mathbb{P}\{\boldsymbol{z} > \boldsymbol{w}(\boldsymbol{x})\} \leq .10$. For portfolio optimization, the decision is a nonnegative weight vector summing to one. We maximize the expected return $\mathbb{E}[\boldsymbol{w}(\boldsymbol{x})^\top \boldsymbol{z}]$ subject to $\mathbb{P}\{\boldsymbol{w}(\boldsymbol{x})^\top \boldsymbol{z} < -.01\} \leq .10$. Synthetic portfolios use total returns and real-world portfolios use returns in excess of cash.

We use linear, neural, and random-forest-weighted sample average approximation (RF–SAA) pretrained decision-rule classes. We parameterize linear rules through affine scores of the context and neural rules through two-hidden-layer ReLU networks. We train both using Adam and projected dual ascent with a smoothed empirical constraint. RF–SAA uses forest weights to construct a conditional empirical distribution and obtain an application-specific decision. These classes cover differentiable parametric rules and a nonparametric forest-based rule. All four calibrators evaluate the same 101 path candidates, with $\lambda \in \{0, .01, \ldots, 1\}$. The nominal confidence is 95% for D, N, U; C controls expected violation.

Both objective value and violation rate are evaluated on test data. A synthetic repetition passes when its violation rate on test data is at most 10%. The test-pass rate summarizes these repetitions and is a finite-test diagnostic of population feasibility. Real data studies calibrate on earlier observations and evaluate on later test data.

4.2 Synthetic Experiments

For each application, we generate two DGPs: one is the control version with homoskedastic noise, the other is the stress version with the same context-dependent signal and context-dependent dispersion. Within each DGP, every pretrained rule uses 3,600 training observations. We draw 1,200 validation and 10,000 test observations for each of 200 repetitions per DGP. We evaluate the performance of different validation procedures across decision rules and applications in Table 1 aggregate the two DGPs, giving 400 repetitions per row.

Table 1: Synthetic test results: objective value / mean violation rate (%) / test-pass rate (%). Lower cost is better for newsvendor; higher return (%) is better for portfolio. Raw denotes the pretrained rule. Bold marks the best objective with at least 95% passes and mean violation at most 10%.

Application	Rule class	Raw	D	N	U	C
Newsvendor	Linear	2.840 / 11.95 / 0	3.279 / 6.11 / 100	3.135 / 7.52 / 99.5	**3.115** / 7.76 / 98.5	2.982 / 9.48 / 70.75
	Neural	1.553 / 9.49 / 50	1.688 / 5.58 / 100	1.633 / 6.89 / 100	1.616 / 7.33 / 100	**1.584** / 8.27 / 95.25
	RF–SAA	1.569 / 21.47 / 0	2.080 / 5.91 / 100	**1.979** / 7.23 / 100	1.990 / 7.05 / 100	1.864 / 9.21 / 77.75
Portfolio	Linear	**1.46** / 8.90 / 99.5	1.21 / 6.41 / 100	1.35 / 7.85 / 100	1.38 / 8.15 / 100	1.45 / 8.84 / 99.75
	Neural	1.45 / 8.68 / 59.5	1.20 / 6.35 / 100	1.31 / 7.42 / 99.25	**1.32** / 7.60 / 98	1.41 / 8.38 / 82.75
	RF–SAA	1.56 / 12.20 / 0	1.05 / 6.39 / 100	1.15 / 7.85 / 98.5	**1.17** / 8.04 / 97.75	1.33 / 9.88 / 55.5

The uncalibrated pretrained decision rules, although trained by smooth constraint Lagrangian, might have poor feasibility guarantees as we mention previously. DKW passes all 2,400 repetitions across applications and classes. Both Gaussian calibrators improve the objective over DKW in every row, with test-pass rates of 98.5–100% for N and 97.75–100% for U. Conformal calibration has overall pass rates of 81.3% for newsvendor and 79.3% for portfolio under its expected-violation target. Figure 1 shows how each pretrained rule determines the available optimality–feasibility tradeoff and how calibrators choose an adjustment along its path. The detailed implementation configuration values and computational time are reported in Appendix D.4. Here, complete calibration takes .002–.865 seconds, compared with 1.047–28.794 seconds for one pretraining run. In Appendix D.6, we further include complete-pipeline repetitions that redraw training, validation, and test data yield similar feasibility comparisons when pretraining variability is included.

We also extend our framework by fitting a *context-dependent* safety parameter $\lambda(\boldsymbol{x})$ on the validation data. It improves mean objective in all six application–class rows, with higher mean test violation rates (still under 10%) and 2,392/2,400 test passes. Appendix D.7 gives the procedure and results.

4.3 Real Data Experiments

The newsvendor study uses Bakery, Restaurant, and Store Item Demand (SID) data (Buttler et al., 2022; D3 Group, 2022), with contexts available at decision time and 28 test days after calibration.

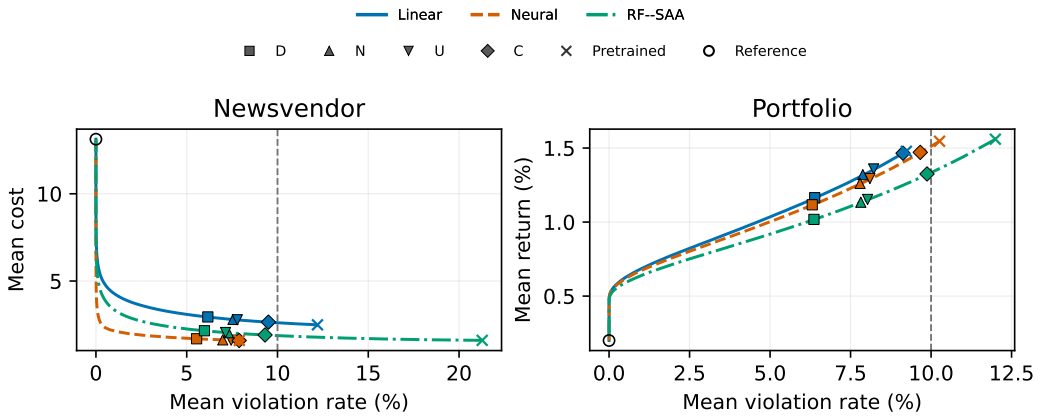


Figure 1: Interpolation paths under stress DGP averaged over 200 repetitions. Colors identify pretrained decision-rule classes; markers show pretrained and reference rules and four calibrators. High-confidence calibrators use 95% confidence. Dashed line is the 10% violation target.

The portfolio study uses four Fama–French panels (French, 2024), with lagged factors as contexts (Fama & French, 1993). Annual refits use 240 months of training data and 240 disjoint months of validation data; testing spans July 2003–December 2018.

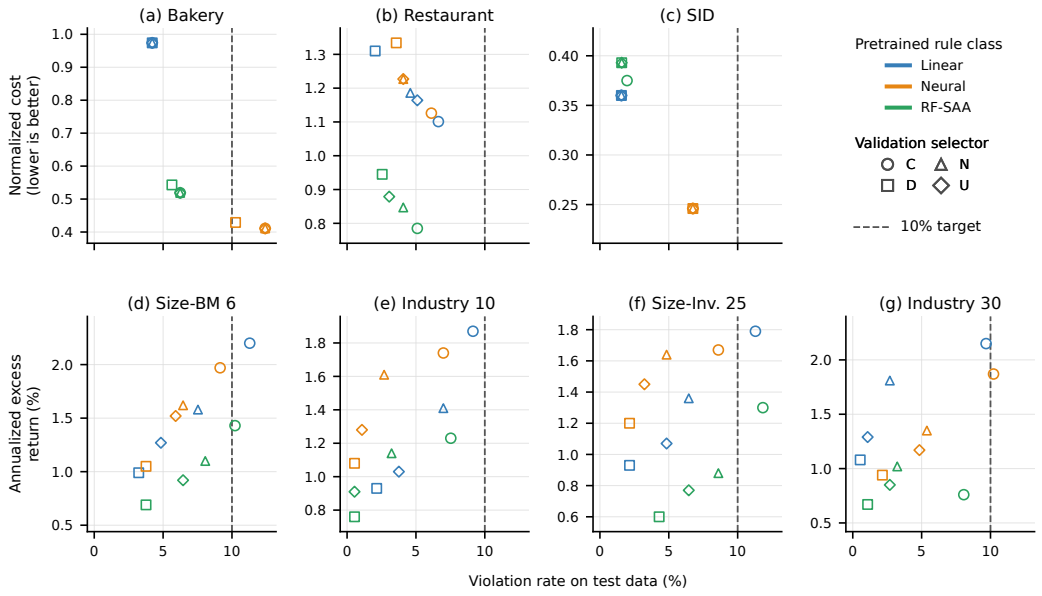


Figure 2: Real data objective values and violation rates. Colors identify pretrained decision-rule classes and marker shapes identify calibrators. Lower cost is better in (a)–(c); higher return is better in (d)–(g). The dashed line is the 10% target. Superimposed hollow markers show identical reported outcomes. The detailed values are provided in Table 4 of Appendix D.5.

All calibrators meet the stockout target on Restaurant and SID (Figure 2). Bakery neural outcomes have 10.26–12.41% violation in the subsequent test window. In portfolio, D, N, U meet the target in all twelve panel–class combinations, with N giving the highest return among them in every case. C meets the target in seven combinations. Together with the synthetic results, these findings support N as a practical default; the relative performance of two Gaussian calibrators varies across applications.

We also examine how the pretraining objective affects calibration (Table 5 in Appendix D.6). Across eight portfolio panel–class combinations, every pretrained rule exceeds the 10% violation target under both chance-constrained training and training on the objective alone. DKW calibration meets the target on test data in all eight combinations for each training objective, with mean annualized excess returns of 1.02% and .93%, respectively. These findings illustrate how the same validation procedure can be applied to rules obtained with different training objectives.

Comparisons with existing methods. Table 2 compares with Conformal-Predict-Then-Optimize (CPO) (Patel et al., 2024), Learning Robust Optimization (LRO) (Wang et al., 2023), Conformal Predictive Programming (CPP) (Zhao et al., 2024), and probability-trimming conditional DRO (Trim–DRO) (Esteban-Pérez & Morales, 2022; 2023). CPO and LRO learn uncertainty regions for robust decisions; CPP uses a conformal quantile reformulation; Trim–DRO optimizes over a conditional ambiguity set. Our comparison uses N with RF–SAA for newsvendor and neural rules for portfolio. Our method gives the lowest newsvendor cost on Bakery and Restaurant and ties with CPO on SID at the displayed precision. It gives the highest portfolio return on three panels; LRO gives the highest return on Size–BM 6.

Table 2: Real data baseline test results: objective value / violation rate (%). Objectives are normalized cost for newsvendor and annualized excess return (%) for portfolio. Bold marks the best objective with violation at most 10%; n/a records the computational limit.

Application	Dataset / panel	Ours (N)	CPO	LRO	CPP	Trim–DRO
Newsvendor (RF–SAA)	Bakery	**.519** / 6.24	.676 / 3.68	.887 / 3.95	1.015 / 3.53	*n/a*
	Restaurant	**.847** / 4.08	.945 / 3.06	.945 / 1.53	1.015 / 7.14	1.214 / 2.55
	SID	**.393** / 1.58	**.393** / 4.02	.483 / .20	.591 / .16	*n/a*
Portfolio (neural)	Size–BM 6	1.62 / 6.45	.00 / .00	**3.19** / .00	.87 / 2.69	1.07 / 5.38
	Industry 10	**1.61** / 2.69	.00 / .00	−.73 / 1.08	.96 / 3.23	1.23 / 4.30
	Size–Investment 25	**1.64** / 4.84	.00 / .00	1.19 / .00	1.24 / 2.69	1.26 / 3.23
	Industry 30	**1.35** / 5.38	.00 / .00	1.29 / 1.08	1.14 / 2.69	1.30 / 5.38

Additional applications. Appendix E evaluates the framework on contextual fractional-knapsack and alloy-allocation problems using linear pretrained rules. We compare with SPO-RC+, a contextual stochastic linear optimization method that combines learned coefficient and cost predictions with a conformal uncertainty set and a downstream robust optimization problem (Im et al., 2025). Across the two applications and DGPs, the raw pretrained rules have mean violation rates ranging from 17.1% to 41.2%, while all four calibrators reduce the mean violation rate below the 10% target. In fractional knapsack, the calibrated rules attain higher reward than SPO-RC+ while using more of the available violation budget. In alloy allocation, SPO-RC+ is more conservative, whereas conformal calibration attains a lower cost with mean violation near the target. These results illustrate how calibration adapts the optimality-feasibility tradeoff across different decision structures and constraint boundaries.

5 Conclusion

We develop a model-agnostic and easy-to-implement pretrain-then-calibrate framework that aims to enhance optimality-feasibility balance for safe contextual optimization. The framework constructs an interpolation path that links a pretrained rule with a safe reference rule, and reduces calibration to selecting a scalar safety parameter with validation data. Compared to existing literature, our approach bypasses the difficulty in obtaining rigorous feasibility via direct decision training, and the low dimensionality of the interpolation path makes the calibration statistically efficient. We establish expected-violation guarantees for conformal calibration and high-confidence feasibility and pathwise optimality guarantees for DKW and Gaussian calibrators. We evaluate the framework across synthetic and real-world newsvendor and portfolio problems, with additional fractional-knapsack and alloy-allocation experiments. The results show that calibration substantially reduces violation rates while preserving useful objective performance across different pretrained decision-rule classes. Future work includes context-dependent and multidimensional interpolation paths and learning reference rules while accounting for their feasibility uncertainty.

AI USE STATEMENT

We used OpenAI Codex to provide feedback on research methodology and experiment design; implement, debug, and edit experimental code, including code used to generate synthetic samples from author-specified data-generating processes; interpret experimental results; prepare and revise tables and figures; and polish parts of the manuscript for organization, readability, and consistency. The authors reviewed all AI-assisted work, verified the computational implementation and results, and rewrote the entire manuscript to ensure clarity. We take responsibility for the final content of this work, including text, claims, code, and artifacts produced with the aid of generative AI.

REFERENCES

N. H. Agnew, R. A. Agnew, J. Rasmussen, and K. R. Smith. An application of chance constrained programming to portfolio selection in a casualty insurance firm. *Management Science*, 15(10): B–512–B–520, 1969. doi: 10.1287/mnsc.15.10.B512.

Anastasios N. Angelopoulos, Stephen Bates, Adam Fisch, Lihua Lei, and Tal Schuster. Conformal risk control. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=33XGfHLtZg.

Aaron Babier, Timothy C. Y. Chan, Adam Diamant, and Rafid Mahmood. Learning to optimize contextually constrained problems for real-time decision generation. *Management Science*, 71(2): 1165–1186, 2024. doi: 10.1287/mnsc.2020.03565. URL https://doi.org/10.1287/mnsc.2020.03565.

Aharon Ben-Tal, Arkadi Nemirovski, and Laurent El Ghaoui. Robust optimization. 2009.

Dimitris Bertsimas, David B Brown, and Constantine Caramanis. Theory and applications of robust optimization. *SIAM review*, 53(3):464–501, 2011.

Simone Buttler, Andreas Philippi, Nikolai Stein, and Richard Pibernik. A meta analysis of data-driven newsvendor approaches. In *ML Evaluation Standards Workshop at ICLR 2022*, 2022. URL https://ml-eval.github.io/assets/pdf/ICLR22_Workshop_ML_Eval_DDNV.pdf.

Marco C Campi and Simone Garatti. A sampling-and-discarding approach to chance-constrained optimization: feasibility and optimality. *Journal of optimization theory and applications*, 148(2): 257–280, 2011.

Marco C. Campi and Simone Garatti. *Introduction to the Scenario Approach*. MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics, 2018. doi: 10.1137/1.9781611975444.

Luiz Chamon and Alejandro Ribeiro. Probably approximately correct constrained learning. *Advances in Neural Information Processing Systems*, 33:16722–16735, 2020.

Abraham Charnes, William W Cooper, and Gifford H Symonds. Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil. *Management Science*, 4(3): 235–263, 1958. doi: 10.1287/mnsc.4.3.235.

Haoxian Chen, Ziyi Huang, Henry Lam, Huajie Qian, and Haofeng Zhang. Learning prediction intervals for regression: Generalization and calibration. In *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics*, volume 130 of *Proceedings of Machine Learning Research*, pp. 820–828. PMLR, 2021. URL https://proceedings.mlr.press/v130/chen21b.html.

Shivam R. Chenreddy, Chaithanya Bandi, and Erick Delage. Data-driven conditional robust optimization. In *Advances in Neural Information Processing Systems*, volume 35, 2022. doi: 10.52202/068431-0692. URL https://proceedings.neurips.cc/paper_files/paper/2022/hash/3df874367ce2c43891aab1ab23ae6959-Abstract-Conference.html.

Victor Chernozhukov, Denis Chetverikov, and Kengo Kato. Central limit theorems and bootstrap in high dimensions. *The Annals of Probability*, 45(4):2309–2352, 2017. doi: 10.1214/16-AOP1113.

Ya-Chi Chu, Alkiviades Boukas, and Madeleine Udell. Snarenet: Flexible repair layers for neural networks with hard constraints. *arXiv preprint arXiv:2602.09317*, 2026. doi: 10.48550/arXiv.2602.09317. URL https://arxiv.org/abs/2602.09317.

Gonzalo E Constante-Flores, Hao Chen, and Can Li. Enforcing hard linear constraints in deep learning models with decision rules. *arXiv preprint arXiv:2505.13858*, 2025. doi: 10.48550/arXiv.2505.13858. URL https://arxiv.org/abs/2505.13858.

D3 Group. A structured evaluation of data-driven newsvendor approaches. GitHub repository, 2022. URL https://github.com/d3group/A-structured-evaluation-of-data-driven-newsvendor-approaches. Commit eb6a3893e70d550ae8f44952bb00fe9ef791f1e3; accessed August 10, 2026.

Othman El Balghiti, Adam N. Elmachtoub, Paul Grigas, and Ambuj Tewari. Generalization bounds in the predict-then-optimize framework. *Mathematics of Operations Research*, 48(4):2043–2065, 2023. doi: 10.1287/moor.2022.1330.

Adam N Elmachtoub and Paul Grigas. Smart "predict, then optimize". *Management Science*, 68(1): 9–26, 2022. doi: 10.1287/mnsc.2020.3922.

Adrián Esteban-Pérez and Juan M. Morales. Distributionally robust stochastic programs with side information based on trimmings. *Mathematical Programming*, 195:1069–1105, 2022. doi: 10.1007/s10107-021-01724-0. URL https://doi.org/10.1007/s10107-021-01724-0.

Adrián Esteban-Pérez and Juan M. Morales. Distributionally robust optimal power flow with contextual information. *European Journal of Operational Research*, 306(3):1047–1058, 2023. doi: 10.1016/j.ejor.2022.10.024. URL https://doi.org/10.1016/j.ejor.2022.10.024.

Eugene F. Fama and Kenneth R. French. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56, 1993. doi: 10.1016/0304-405X(93)90023-5.

Ferdinando Fioretto, Pascal Van Hentenryck, Terrence WK Mak, Cuong Tran, Federico Baldo, and Michele Lombardi. Lagrangian duality for constrained deep learning. In *Machine Learning and Knowledge Discovery in Databases: Applied Data Science and Demo Track*, pp. 118–135. Springer, 2020.

Kenneth R. French. U.S. Historical Research Returns as of December 2024. Kenneth R. French Data Library, Dartmouth College, 2024. URL https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library_202412_archive.html. Accessed August 10, 2026.

Simone Garatti and Marco C Campi. Risk and complexity in scenario optimization. *Mathematical Programming*, 191(1):243–279, 2022.

L Jeff Hong, Zhiyuan Huang, and Henry Lam. Learning-based robust optimization: Procedures and statistical guarantees. *Management Science*, 67(6):3447–3467, 2021. doi: 10.1287/mnsc.2020.3640.

Yang Hu, Jieren Tan, Changliang Zou, Yajie Bao, and Haojie Ren. Conformal robustness control: A new strategy for robust decision. In *The Fourteenth International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/a80188f9b1246d4b95d396165cf99207-Abstract-Conference.html.

Hyungki Im, Wyame Benslimane, and Paul Grigas. Smart surrogate losses for contextual stochastic linear optimization with robust constraints. In *Advances in Neural Information Processing Systems*, volume 38, 2025. doi: 10.52202/085713-2523. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/6cc2e83abc493d8f7db72c3da1feccb8-Abstract-Conference.html.

Garud Iyengar, Henry Lam, and Tianyu Wang. Hedging against complexity: Distributionally robust optimization with parametric approximation. In *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics*, volume 206 of *Proceedings of Machine Learning Research*, pp. 9976–10011. PMLR, 2023. URL https://proceedings.mlr.press/v206/iyengar23a.html. Extended version: arXiv:2212.01518.

Simge Küçükyavuz and Ruiwei Jiang. Chance-constrained optimization under limited distributional information: A review of reformulations based on sampling and distributional robustness. *EURO Journal on Computational Optimization*, 10:100030, 2022. doi: 10.1016/j.ejco.2022.100030.

Henry Lam and Huajie Qian. Combating conservativeness in data-driven optimization under uncertainty: A solution path approach. *arXiv preprint arXiv:1909.06477*, 2019. doi: 10.48550/arXiv.1909.06477. URL https://arxiv.org/abs/1909.06477.

Miguel A. Lejeune and Janne Kettunen. Managing reliability and stability risks in forest harvesting. *Manufacturing & Service Operations Management*, 19(4):620–638, 2017. doi: 10.1287/msom.2017.0626.

Hongmin Li, Stephen C. Graves, and Woonghee Tim Huh. Optimal capacity conversion for product transitions under high service requirements. *Manufacturing & Service Operations Management*, 16(1):46–60, 2013. doi: 10.1287/msom.2013.0445.

Yongzhen Li, Jia Shu, Miao Song, Jiawei Zhang, and Huan Zheng. Multisourcing supply network design: Two-stage chance-constrained model, tractable approximations, and computational results. *INFORMS Journal on Computing*, 29(2):287–300, 2017. doi: 10.1287/ijoc.2016.0730.

Enming Liang, Minghua Chen, and Steven H Low. Homeomorphic projection to ensure neural-network solution feasibility for constrained optimization. *Journal of Machine Learning Research*, 25(329):1–55, 2024. URL https://www.jmlr.org/papers/v25/23-1577.html.

James Luedtke and Shabbir Ahmed. A sample approximation approach for optimization with probabilistic constraints. *SIAM Journal on Optimization*, 19(2):674–699, 2008. doi: 10.1137/070702928.

James Luedtke, Shabbir Ahmed, and George L. Nemhauser. An integer programming approach for linear programs with probabilistic constraints. *Mathematical Programming*, 122(2):247–272, 2010. doi: 10.1007/s10107-008-0247-4.

Pascal Massart. The tight constant in the Dvoretzky–Kiefer–Wolfowitz inequality. *The Annals of Probability*, 18(3):1269–1283, 1990. doi: 10.1214/aop/1176990746.

David Mildebrath, Taewoo Lee, Saumya Sinha, Andrew J. Schaefer, and A. Osama Gaber. Characterizing rational transplant program response to outcome-based regulation. *Operations Research*, 72(4):1421–1437, 2023. doi: 10.1287/opre.2018.0721.

Arkadi Nemirovski and Alexander Shapiro. Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, 17(4):969–996, 2006. doi: 10.1137/050622328.

Bernardo K. Pagnoncelli, Shabbir Ahmed, and Alexander Shapiro. Sample average approximation method for chance constrained programming: Theory and applications. *Journal of Optimization Theory and Applications*, 142(2):399–416, 2009. doi: 10.1007/s10957-009-9523-6.

Yash P Patel, Sahana Rayan, and Ambuj Tewari. Conformal contextual robust optimization. In *Proceedings of the 27th International Conference on Artificial Intelligence and Statistics*, volume 238 of *Proceedings of Machine Learning Research*, pp. 2485–2493. PMLR, 2024. URL https://proceedings.mlr.press/v238/patel24a.html.

Alejandra Peña-Ordieres, James R Luedtke, and Andreas Wachter. Solving chance-constrained problems via a smooth sample-based nonlinear approximation. *SIAM Journal on Optimization*, 30(3):2221–2250, 2020. doi: 10.1137/19M1261985.

Chun Peng, Erick Delage, and Jinlin Li. Probabilistic envelope constrained multiperiod stochastic emergency medical services location model and decomposition scheme. *Transportation Science*, 54(6):1471–1494, 2020. doi: 10.1287/trsc.2019.0947.

Álvaro Porras, Concepción Domínguez, Juan Miguel Morales, and Salvador Pineda. Tight and compact sample average approximation for joint chance-constrained problems with applications to optimal power flow. *INFORMS Journal on Computing*, 35(6):1454–1469, 2023. doi: 10.1287/ijoc.2022.0302.

Meng Qi, Paul Grigas, and Zuo-Jun Shen. Integrated conditional estimation-optimization. *Operations Research*, 74(3):1604–1625, 2026.

Junhao Qin, Ziyi Zou, Liu Liu, and Yongchao Liu. Piecewise affine decision rules for contextual chance-constrained stochastic programming. *Operations Research Letters*, 61:107296, 2025. doi: 10.1016/j.orl.2025.107296. URL https://doi.org/10.1016/j.orl.2025.107296.

Hamed Rahimian and Bernardo Pagnoncelli. Data-driven approximation of contextual chance-constrained stochastic programs. *SIAM Journal on Optimization*, 33(3):2248–2274, 2023. doi: 10.1137/22M1528045. URL https://doi.org/10.1137/22M1528045.

Henry W. J. Reeve. A short proof of the Dvoretzky–Kiefer–Wolfowitz–Massart inequality. *arXiv preprint arXiv:2403.16651*, 2024. doi: 10.48550/arXiv.2403.16651.

Gian-Carlo Rota. The number of partitions of a set. *The American Mathematical Monthly*, 71(5): 498–504, 1964. doi: 10.1080/00029890.1964.11992270.

Utsav Sadana, Abhilash Chenreddy, Erick Delage, Alexandre Forel, Emma Frejinger, and Thibaut Vidal. A survey of contextual optimization methods for decision-making under uncertainty. *European Journal of Operational Research*, 320(2):271–289, 2025.

Philipp J Schneider and Daniel Kuhn. Soft-radial projection for constrained end-to-end learning. *arXiv preprint arXiv:2602.03461*, 2026. doi: 10.48550/arXiv.2602.03461. URL https://arxiv.org/abs/2602.03461.

Natasja Sluijk, Alexandre M. Florio, Joris Kinable, Nico Dellaert, and Tom Van Woensel. A chance-constrained two-echelon vehicle routing problem with stochastic demands. *Transportation Science*, 57(1):252–272, 2022. doi: 10.1287/trsc.2022.1162.

Aad W. van der Vaart and Jon A. Wellner. *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer, 1996. doi: 10.1007/978-1-4757-2545-2.

Irina Wang, Bart Van Parys, and Bartolomeo Stellato. Learning decision-focused uncertainty sets in robust optimization. *arXiv preprint arXiv:2305.19225*, 2023. doi: 10.48550/arXiv.2305.19225. URL https://arxiv.org/abs/2305.19225.

Tianyu Wang, Ningyuan Chen, and Chun Wang. Contextual optimization under covariate shift: A robust approach by intersecting wasserstein balls. *arXiv preprint arXiv:2406.02426*, 2024. doi: 10.48550/arXiv.2406.02426. URL https://arxiv.org/abs/2406.02426.

Christopher Yeh, Nicolas Christianson, Adam Wierman, and Yisong Yue. Conformal risk training: End-to-end optimization of conformal risk control. In *Advances in Neural Information Processing Systems*, volume 38. Curran Associates, Inc., 2025. doi: 10.52202/085713-2355. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/6559542f75b4452ebaaf82094c7defb7-Abstract-Conference.html.

Yiyang Zhang, Junyi Liu, and Xiaobo Zhao. Data-driven piecewise affine decision rules for stochastic programming with covariate information. *Operations Research*, 2026. doi: 10.1287/opre.2023.0175. URL https://pubsonline.informs.org/doi/10.1287/opre.2023.0175.

Yiqi Zhao, Xinyi Yu, Matteo Sesia, Jyotirmoy V. Deshmukh, and Lars Lindemann. Conformal predictive programming for chance constrained optimization. *arXiv preprint arXiv:2402.07407*, 2024. doi: 10.48550/arXiv.2402.07407. URL https://arxiv.org/abs/2402.07407.

Wenbin Zhou and Shixiang Zhu. Calibrating decision robustness via inverse conformal risk control. *arXiv preprint arXiv:2510.07750*, 2026. doi: 10.48550/arXiv.2510.07750. URL https://arxiv.org/abs/2510.07750.

A Additional Related Work

Contextual optimization methods. Contextual optimization methods obtain decisions from observed contexts through prediction-and-optimization procedures or direct decision-rule learning. Decision-focused learning aligns predictions with downstream decision loss (Elmachtoub & Grigas, 2022; El Balghiti et al., 2023). Zhang et al. (2026) study piecewise-affine decision-rule empirical risk minimization and its nonconvex, nonsmooth algorithms. Qin et al. (2025) optimize a piecewise-affine rule under the same population chance-constraint formulation and analyze value and feasible-solution approximation; Babier et al. (2024) study real-time generators under context-dependent deterministic feasible regions. We calibrate pretrained rules obtained from different learning and optimization procedures, and provide complementary generalization results for the empirical constrained training under the empirical chance constraint counterparts.

Chance-constrained optimization. Classical chance-constrained optimization selects a fixed decision before uncertainty is realized. Scenario methods enforce constraints on sampled realizations, while sample-average approximations replace the satisfaction probability by its empirical counterpart (Campi & Garatti, 2018; Luedtke & Ahmed, 2008; Pagnoncelli et al., 2009). Related work develops conservative convex approximations, mixed-integer formulations, and smooth sample-based approximations to address the statistical and computational difficulties of chance constraints (Nemirovski & Shapiro, 2006; Luedtke et al., 2010; Peña-Ordieres et al., 2020; Küçükyavuz & Jiang, 2022). However, these methods do not use observed contextual information and produce a single decision to be applied across all contexts. Contextual chance-constrained optimization instead estimates the conditional distribution of uncertainty from context-weighted observations and solves a context-specific problem (Rahimian & Pagnoncelli, 2023), but generally requires resolving a constrained optimization problem whenever a new context is queried. We instead learn a decision rule that maps contexts directly to decisions and calibrate it along a one-dimensional interpolation path using validation data. Once calibrated, the same rule can be deployed across new contexts without resolving a chance-constrained program, while providing finite-sample feasibility control and retaining the objective performance of the pretrained rule.

Robust, distributionally robust, and conformal optimization. Robust and distributionally robust approaches protect an uncertainty set or ambiguity model (Hong et al., 2021; Esteban-Pérez & Morales, 2022; Iyengar et al., 2023; Wang et al., 2024). In contextual optimization, enforcing the constraint over a contextual uncertainty set transfers its coverage guarantee to the decision. The set's shape and size influence objective performance and the resulting optimization problem. Decision-focused set learning (Wang et al., 2023) incorporates the downstream objective into set construction. Conformal contextual robust optimization, conformal predictive programming, and conformal robustness control calibrate related objects while retaining a robust or reformulated downstream problem (Patel et al., 2024; Zhao et al., 2024; Hu et al., 2026). Im et al. (2025) pair contextual uncertainty sets for uncertain constraints with decision-focused training of cost predictors, including a truncation-and-reweighting variant. Zhou & Zhu (2026) study the miscoverage–regret tradeoff along robust policies indexed by nested uncertainty sets and use an independent data split to recertify a selected robustness level. Their candidate policies require downstream robust solves. We construct candidates by interpolating a pretrained rule and a reference rule, and calibrate their constraint violations directly. good optimality

Constrained learning. Constrained learning studies population requirements through empirical Lagrangians and probably approximately correct (PAC) generalization (Fioretto et al., 2020; Chamon & Ribeiro, 2020). Projection, repair, and feasible architectures enforce deterministic constraints for each input by construction (Liang et al., 2024; Constante-Flores et al., 2025; Schneider & Kuhn, 2026; Chu et al., 2026). Our chance constraint permits an α failure rate, and validation of operational failures determines the adjustment toward the reference rule.

B Training-Stage Generalization

B.1 Empirical Constrained Training and Generalization

For the training results in this appendix, assume that the pairs in $\mathcal{D}_n$ are i.i.d. In addition to the standing assumptions in the main text, we impose the following reference-closure condition on the decision-rule class.

Assumption 2 (Reference closure of the decision-rule class)**.** *The safe reference rule satisfies $w_0 \in \mathcal{H}$, and $\mathcal{H}$ is star-shaped with respect to $\boldsymbol{w}_0$. That is, for every $\boldsymbol{w} \in \mathcal{H}$ and every $\lambda \in [0,1]$,*

$$(1-\lambda)\boldsymbol{w} + \lambda \boldsymbol{w}_0 \in \mathcal{H}.$$

Assumption 2 ensures that the reference rule is feasible for the empirical program and that interpolation toward the reference rule remains within $\mathcal{H}$. We can obtain the pretrained empirical constrained decision rule $\boldsymbol{w}^{\text{pre}}$ by solving the empirical counterpart of (1). That is:

$$\begin{aligned} \min_{\boldsymbol{w}(\cdot)\in\mathcal{H}} \quad & \frac{1}{n}\sum_{i=1}^{n} c\left(\boldsymbol{w}\left(\boldsymbol{x}_i\right), \boldsymbol{z}_i\right) \\ \text{s.t.} \quad & \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\left\{g\left(\boldsymbol{w}\left(\boldsymbol{x}_i\right), \boldsymbol{z}_i\right) \le 0\right\} \ge 1-\alpha+t. \end{aligned} \tag{6}$$

The tightening margin $t \in [0, \alpha]$ accounts for statistical error in estimating the safety probability over $\mathcal{H}$. Under Assumption 2, the reference rule makes the empirical feasible set nonempty.

We study the statistical consequences of solving the empirical program (6). Empirical feasibility with an appropriate tightening margin provides population feasibility. A small optimization gap further yields a population cost guarantee. The result is conditional on these properties of the training output: the optimization tolerance records how accurately training solves the constrained problem. For a decision rule $\boldsymbol{w} \in \mathcal{H}$, write $R(\boldsymbol{w}) := \mathbb{P}\{g(\boldsymbol{w}(\boldsymbol{x}), \boldsymbol{z}) \le 0\}$ for its safety probability and $C(\boldsymbol{w}) := \mathbb{E}[c(\boldsymbol{w}(\boldsymbol{x}), \boldsymbol{z})]$ for its expected cost. Their empirical counterparts are $R_n(\boldsymbol{w}) := n^{-1}\sum_{i=1}^{n} \mathbf{1}\{g(\boldsymbol{w}(\boldsymbol{x}_i), \boldsymbol{z}_i) \le 0\}$ and $C_n(\boldsymbol{w}) := n^{-1}\sum_{i=1}^{n} c(\boldsymbol{w}(\boldsymbol{x}_i), \boldsymbol{z}_i)$. To relate these empirical quantities to their population counterparts uniformly over $\mathcal{H}$, we control the complexity of the decision-rule class through its metric entropy.

Definition 1 (Finite metric entropy)**.** For decision rules $\boldsymbol{w}, \widetilde{\boldsymbol{w}} \in \mathcal{H}$, set $\|\boldsymbol{w} - \widetilde{\boldsymbol{w}}\|_\infty := \sup_{\boldsymbol{x}\in\mathcal{X}} \|\boldsymbol{w}(\boldsymbol{x}) - \widetilde{\boldsymbol{w}}(\boldsymbol{x})\|_2$. The class $\mathcal{H}$ has *finite metric entropy* if its internal covering number $\mathcal{N}(\varepsilon, \mathcal{H}, \|\cdot\|_\infty)$ is finite for every $\varepsilon > 0$.

We impose Lipschitz regularity on the constraint and cost, a local probability regularity condition at the constraint boundary, and boundedness of the cost.

Assumption 3 (Constraint regularity)**.** *$g(\cdot, \boldsymbol{z})$ is L_w^g-uniformly Lipschitz for some $L_w^g > 0$.*

Assumption 4 (Cost function regularity)**.** *There exist $L_w^c, B_c > 0$ such that $c(\cdot, \boldsymbol{z})$ is L_w^c-Lipschitz and $|c(\boldsymbol{w}, \boldsymbol{z})| \le B_c$ for every $\boldsymbol{w}$ and $\boldsymbol{z}$.*

For affine or piecewise-affine constraint and cost functions whose coefficients are bounded on the support of $\boldsymbol{z}$, a bounded decision set Ω provides finite constants in Assumptions 3 and 4.

Assumption 5 (Boundary-mass regularity)**.** *There exist $\bar{d} < \infty$ and $r_{\text{loc}} > 0$ such that $\max\{\mathbb{P}(-r < g(\boldsymbol{w}(\boldsymbol{x}), \boldsymbol{z}) \le 0), \mathbb{P}(0 < g(\boldsymbol{w}(\boldsymbol{x}), \boldsymbol{z}) \le r)\} \le \bar{d}r$ for $\boldsymbol{w} \in \mathcal{H}$ and $r \in [0, r_{\text{loc}}]$.*

Under these conditions, for $\delta \in (0,1)$, we define deviation radii for the safety probability and cost

$$\phi(n, \mathcal{H}, \delta) := \inf_{0<\varepsilon\le r_{\text{loc}}/(2L_w^g)} \left\{ 2\bar{d}L_w^g\varepsilon + \sqrt{\frac{\log\left(2\mathcal{N}(\varepsilon, \mathcal{H}, \|\cdot\|_\infty)/\delta\right)}{2n}} \right\},$$

$$\psi(n, \mathcal{H}, \delta) := \inf_{\varepsilon>0} \left\{ 2L_w^c\varepsilon + 2B_c\sqrt{\frac{\log\left(2\mathcal{N}(\varepsilon, \mathcal{H}, \|\cdot\|_\infty)/\delta\right)}{2n}} \right\}.$$

For $a \in [0,1]$, define the population value $V(a) := \inf_{\boldsymbol{w}\in\mathcal{H}:\, R(\boldsymbol{w})\geq 1-a} C(\boldsymbol{w})$. The value is $+\infty$ if the feasible set is empty. Under Assumption 2 and the bounded-cost part of Assumption 4, it is finite for every $a \in [0,1]$.

For a margin $r_n \in [0,\alpha]$ and tolerance $\zeta_n \geq 0$, a rule $\widehat{\boldsymbol{w}} \in \mathcal{H}$ is *ζ_n-optimal at margin r_n* if

$$R_n(\widehat{\boldsymbol{w}}) \geq 1-\alpha+r_n, \qquad C_n(\widehat{\boldsymbol{w}}) \leq \inf_{\substack{\boldsymbol{w}\in\mathcal{H}:\\ R_n(\boldsymbol{w})\geq 1-\alpha+r_n}} C_n(\boldsymbol{w}) + \zeta_n. \tag{7}$$

The reference rule makes the empirical feasible set nonempty. The basic optimality guarantee compares the learned rule with the population optimum under a slightly stricter chance constraint. To measure excess cost relative to the original target $V(\alpha)$, we use the following stability condition.

Assumption 6 (Local value stability). *Assume $V(\alpha-t) \leq V(\alpha) + K_V t$ for every $t \in [0,\bar{t}_V]$ for some $K_V < \infty$ and $\bar{t}_V \in (0,\alpha]$.*

Appendix B.5 gives a transparent sufficient condition based on a strictly feasible decision rule.

Theorem 3 (Training with an explicit feasibility margin). *Suppose $\mathcal{H}$ satisfies Definition 1 and Assumptions 3, 4, and 5 hold. Fix a tolerance $\zeta_n \geq 0$. Let $\delta_1, \delta_2 \in (0,1)$ satisfy $\delta_1+\delta_2 < 1$ and write $r_n := \phi(n,\mathcal{H},\delta_1)$, $s_n := \psi(n,\mathcal{H},\delta_2)$. If $r_n \leq \alpha$, set the tightening margin in the empirical program (6) to $t = r_n$, and let $\widehat{\boldsymbol{w}}$ be any rule that is ζ_n-optimal in the sense of (7). Then*

$$\mathbb{P}_{\mathcal{D}_n}\left(R(\widehat{\boldsymbol{w}}) \geq 1-\alpha\right) \geq 1-\delta_1.$$

If, in addition, Assumption 6 holds and $2r_n \leq \bar{t}_V$, then

$$\mathbb{P}_{\mathcal{D}_n}\left(C(\widehat{\boldsymbol{w}}) - V(\alpha) \leq 2K_V r_n + 2s_n + \zeta_n\right) \geq 1-\delta_1-\delta_2.$$

Theorem 3 separates statistical errors through r_n and s_n from optimization error through ζ_n. Empirical feasibility at margin r_n suffices for its feasibility conclusion; the small optimization gap is needed for near-optimality. When training meets these conditions with useful bounds, it already provides the desired population guarantees. Establishing a small gap for an expressive rule class can be difficult, and the uniform statistical margin may be large. Section 3 gives calibration guarantees that condition on the trained candidates and use validation data to determine the safety adjustment. In the following, we provide the proof of Theorem 3 in Appendices B.2 and B.3, examples that satisfy these assumptions in Appendices B.4 and B.5. Furthermore, we provide a VC/Rademacher extension for classes such as trees with continuously varying splits that generalizes the definition of metric entropy in Appendix B.6.

B.2 Uniform-Deviation Results

Lemma 2 (Uniform deviations in safety probability and cost). *Suppose $\mathcal{H}$ satisfies Definition 1. For every $\delta \in (0,1)$, the following statements hold.*

(i) Under Assumptions 3 and 5, with probability at least $1-\delta$,

$$\sup_{\boldsymbol{w}\in\mathcal{H}} |R_n(\boldsymbol{w}) - R(\boldsymbol{w})| \leq \phi(n,\mathcal{H},\delta).$$

(ii) Under Assumption 4, with probability at least $1-\delta$,

$$\sup_{\boldsymbol{w}\in\mathcal{H}} |C_n(\boldsymbol{w}) - C(\boldsymbol{w})| \leq \psi(n,\mathcal{H},\delta).$$

Proof. (i) Safety probability. Fix $0 < \varepsilon \leq r_{\mathrm{loc}}/(2L_w^g)$ and let $\mathcal{H}_\varepsilon$ be an ε-net of cardinality $\mathcal{N}(\varepsilon,\mathcal{H},\|\cdot\|_\infty)$. Write $G_{\boldsymbol{w}} := g(\boldsymbol{w}(\boldsymbol{x}),\boldsymbol{z})$ and let $\mathbb{P}_n$ denote empirical probability. Set

$$\eta_\varepsilon := \sqrt{\frac{\log\left(2\mathcal{N}(\varepsilon,\mathcal{H},\|\cdot\|_\infty)/\delta\right)}{2n}}.$$

One-sided Hoeffding inequalities and a union bound over the two relevant thresholds for each net rule give, with probability at least $1-\delta$, simultaneously for every $\boldsymbol{w}^{\mathrm{net}} \in \mathcal{H}_\varepsilon$,

$$\mathbb{P}_n\left(G_{\boldsymbol{w}^{\mathrm{net}}} \leq L_w^g\varepsilon\right) \leq \mathbb{P}\left(G_{\boldsymbol{w}^{\mathrm{net}}} \leq L_w^g\varepsilon\right) + \eta_\varepsilon,$$

$$\mathbb{P}_n\left(G_{\boldsymbol{w}^{\mathrm{net}}} \leq -L_w^g \varepsilon\right) \geq \mathbb{P}\left(G_{\boldsymbol{w}^{\mathrm{net}}} \leq -L_w^g \varepsilon\right) - \eta_\varepsilon.$$

On this event, fix any $\boldsymbol{w} \in \mathcal{H}$ and choose $\boldsymbol{w}^{\mathrm{net}} \in \mathcal{H}_\varepsilon$ with $\|\boldsymbol{w} - \boldsymbol{w}^{\mathrm{net}}\|_\infty \leq \varepsilon$. Lipschitz continuity gives $|G_{\boldsymbol{w}} - G_{\boldsymbol{w}^{\mathrm{net}}}| \leq L_w^g \varepsilon$ almost surely. Applying the positive and negative threshold inequalities, respectively, and then the boundary-mass condition gives

$$|R_n(\boldsymbol{w}) - R(\boldsymbol{w})| \leq 2\bar{d} L_w^g \varepsilon + \eta_\varepsilon.$$

Taking the supremum over $\boldsymbol{w}$ proves the bound for every fixed ε. Finally, take a sequence of these bounds decreasing to their infimum and use continuity of probability for the associated decreasing events.

(ii) Cost. Let $\mathcal{H}_\varepsilon$ be an ε-net of cardinality $\mathcal{N}(\varepsilon, \mathcal{H}, \|\cdot\|_\infty)$. For each $\boldsymbol{w} \in \mathcal{H}$, choose $\boldsymbol{w}^{\mathrm{net}} \in \mathcal{H}_\varepsilon$ with $\|\boldsymbol{w} - \boldsymbol{w}^{\mathrm{net}}\|_\infty \leq \varepsilon$. The cost Lipschitz assumption gives

$$|C_n(\boldsymbol{w}) - C(\boldsymbol{w})| \leq 2L_w^c \varepsilon + \left|C_n\left(\boldsymbol{w}^{\mathrm{net}}\right) - C\left(\boldsymbol{w}^{\mathrm{net}}\right)\right|.$$

Hoeffding's inequality and a union bound imply, with probability at least $1 - \delta$,

$$\sup_{\boldsymbol{w}^{\mathrm{net}} \in \mathcal{H}_\varepsilon} \left|C_n\left(\boldsymbol{w}^{\mathrm{net}}\right) - C\left(\boldsymbol{w}^{\mathrm{net}}\right)\right| \leq 2B_c \sqrt{\frac{\log\left(2\mathcal{N}(\varepsilon, \mathcal{H}, \|\cdot\|_\infty)/\delta\right)}{2n}}.$$

Taking the supremum over $\boldsymbol{w}$ and then passing to a sequence of bounds decreasing to their infimum proves the result. □

B.3 Oracle Reduction

Lemma 3 (Oracle reduction from uniform deviations). *Fix $r_n, s_n, \zeta_n \geq 0$, let $\delta_1, \delta_2 \in (0,1)$ satisfy $\delta_1 + \delta_2 < 1$, and suppose*

$$\mathbb{P}\left(\sup_{\boldsymbol{w} \in \mathcal{H}} |R_n(\boldsymbol{w}) - R(\boldsymbol{w})| \leq r_n\right) \geq 1 - \delta_1, \qquad \mathbb{P}\left(\sup_{\boldsymbol{w} \in \mathcal{H}} |C_n(\boldsymbol{w}) - C(\boldsymbol{w})| \leq s_n\right) \geq 1 - \delta_2.$$

Assume $2r_n \leq \alpha$. Every rule $\widehat{\boldsymbol{w}}$ that is ζ_n-optimal at margin r_n satisfies

$$\mathbb{P}_{\mathcal{D}_n}\left(R(\widehat{\boldsymbol{w}}) \geq 1 - \alpha\right) \geq 1 - \delta_1, \qquad \mathbb{P}_{\mathcal{D}_n}\left(C(\widehat{\boldsymbol{w}}) \leq V(\alpha - 2r_n) + 2s_n + \zeta_n\right) \geq 1 - \delta_1 - \delta_2.$$

Proof of Lemma 3. Let $\mathcal{E}_R$ and $\mathcal{E}_C$ denote the two displayed deviation events. On $\mathcal{E}_R$, empirical feasibility of $\widehat{\boldsymbol{w}}$ implies

$$R(\widehat{\boldsymbol{w}}) \geq R_n(\widehat{\boldsymbol{w}}) - r_n \geq 1 - \alpha.$$

Moreover, every $\boldsymbol{w} \in \mathcal{H}$ with $R(\boldsymbol{w}) \geq 1 - \alpha + 2r_n$ satisfies

$$R_n(\boldsymbol{w}) \geq R(\boldsymbol{w}) - r_n \geq 1 - \alpha + r_n$$

on $\mathcal{E}_R$, so it is a feasible empirical comparator. Hence, on $\mathcal{E}_R \cap \mathcal{E}_C$, empirical optimality and two applications of the cost deviation give

$$C(\widehat{\boldsymbol{w}}) \leq C_n(\widehat{\boldsymbol{w}}) + s_n \leq \inf_{\boldsymbol{w} \in \mathcal{H}:\, R(\boldsymbol{w}) \geq 1 - \alpha + 2r_n} C_n(\boldsymbol{w}) + s_n + \zeta_n \leq V(\alpha - 2r_n) + 2s_n + \zeta_n.$$

The last step remains valid when the infimum is not attained: apply the uniform cost bound to an approximating sequence of population comparators. Finally, $\mathbb{P}(\mathcal{E}_R) \geq 1 - \delta_1$, and the union bound gives $\mathbb{P}(\mathcal{E}_R \cap \mathcal{E}_C) \geq 1 - \delta_1 - \delta_2$. □

Proof of Theorem 3. Parts (i)–(ii) of Lemma 2 verify the two deviation hypotheses of Lemma 3 with

$$r_n = \phi(n, \mathcal{H}, \delta_1), \qquad s_n = \psi(n, \mathcal{H}, \delta_2).$$

On the safety-probability deviation event, $R(\widehat{\boldsymbol{w}}) \geq R_n(\widehat{\boldsymbol{w}}) - r_n \geq 1 - \alpha$, which proves feasibility when $r_n \leq \alpha$. Under Assumption 6, the lemma's cost bound and its defining inequality at $t = 2r_n$ give the stated excess-cost bound. □

B.4 Metric Entropy Examples

The examples below verify the uniform-norm entropy condition for finite, parametric, and smooth rule classes, and show why hard-split trees require the VC route. As in Definition 1, covering centers belong to the class being covered.

Example 1 (Finite classes)**.** If $|\mathcal{H}| = M < \infty$, then the class itself is an ε-net for every $\varepsilon > 0$, and

$$\log \mathcal{N}\left(\varepsilon, \mathcal{H}, \|\cdot\|_\infty\right) \le \log M.$$

For bounded finite-dimensional parameterizations, the following result transfers a parameter-space cover to the rule class. It applies directly to vector-valued rules under the uniform Euclidean norm.

Proposition 1 (Parameter-to-function entropy transfer)**.** *Let $\Theta \subseteq \mathbb{R}^P$ be contained in the Euclidean ball of radius B, and let $\mathcal{H} = \{\boldsymbol{w}_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \Theta\}$. Suppose, for some $L_\Theta > 0$,*

$$\|\boldsymbol{w}_{\boldsymbol{\theta}} - \boldsymbol{w}_{\boldsymbol{\theta}'}\|_\infty \le L_\Theta \|\boldsymbol{\theta} - \boldsymbol{\theta}'\|_2 \qquad \text{for every } \boldsymbol{\theta}, \boldsymbol{\theta}' \in \Theta.$$

Then, for every $\varepsilon > 0$,

$$\log \mathcal{N}\left(\varepsilon, \mathcal{H}, \|\cdot\|_\infty\right) \le P \log\left(1 + \frac{2BL_\Theta}{\varepsilon}\right).$$

Proof. Let $\Theta_0 \subseteq \Theta$ be an ε/L_Θ-net. Its image $\{\boldsymbol{w}_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \Theta_0\} \subseteq \mathcal{H}$ is an ε-net by the displayed Lipschitz condition. A maximal r-separated subset of Θ, with $r = \varepsilon/L_\Theta$, is an r-net. The Euclidean balls of radius $r/2$ around its points are disjoint and lie in the ball of radius $B + r/2$. Comparing volumes bounds its cardinality by $(1 + 2B/r)^P$. Taking logarithms gives the result. □

Example 2 (Linear, generalized linear, and basis rules)**.** If $\sup_{\boldsymbol{x} \in \mathcal{X}} \|\boldsymbol{x}\|_2 \le B_x$ and $\|\boldsymbol{\theta}\|_2 \le B$, the linear rule $\boldsymbol{w}_{\boldsymbol{\theta}}(\boldsymbol{x}) = \boldsymbol{\theta}^\top \boldsymbol{x}$ has $L_\Theta = B_x$. With an L_σ-Lipschitz scalar link, the rule $\boldsymbol{w}_{\boldsymbol{\theta}}(\boldsymbol{x}) = \sigma\left(\boldsymbol{\theta}^\top \boldsymbol{x}\right)$ has $L_\Theta = L_\sigma B_x$.

For basis rules, let $\Phi(\boldsymbol{x}) = (\phi_1(\boldsymbol{x}), \ldots, \phi_D(\boldsymbol{x}))$ satisfy $\sup_{\boldsymbol{x} \in \mathcal{X}} \|\Phi(\boldsymbol{x})\|_2 \le B_\Phi$. The class $\boldsymbol{w}_a(\boldsymbol{x}) = a^\top \Phi(\boldsymbol{x})$ with $\|a\|_2 \le B$ has $L_\Theta = B_\Phi$ and $P = D$. This includes polynomial bases, a fixed collection of random contexts, and spline bases with fixed knots. In each case, Cauchy–Schwarz verifies the Lipschitz condition of Proposition 1.

Example 3 (Bounded neural networks with a fixed architecture)**.** Consider a fixed feedforward architecture with parameter vector $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^P$. Assume that $\mathcal{X}$ is bounded, every activation is Lipschitz, and every weight matrix and bias is restricted to a bounded set. Then there is a finite architecture-dependent constant L_{net} such that

$$\|\boldsymbol{w}_{\boldsymbol{\theta}} - \boldsymbol{w}_{\boldsymbol{\theta}'}\|_\infty \le L_{\text{net}} \|\boldsymbol{\theta} - \boldsymbol{\theta}'\|_2. \tag{8}$$

Consequently, if Θ is contained in a radius-B ball, then

$$\log \mathcal{N}\left(\varepsilon, \mathcal{H}, \|\cdot\|_\infty\right) \le P \log\left(1 + \frac{2BL_{\text{net}}}{\varepsilon}\right).$$

This statement covers, for example, a fixed ReLU, sigmoid, or $\tanh$ architecture. It does not assert a dimension-free bound: the parameter count P and L_{net} may both grow when the architecture grows.

To verify (8), write a layer as $h_\ell = \sigma_\ell(W_\ell h_{\ell-1} + b_\ell)$. Bounded inputs and parameters give a uniform bound M_ℓ on $\|h_\ell\|$ by induction. If the activation at layer ℓ has Lipschitz constant L_ℓ and $\|W_\ell\| \le S_\ell$, then

$$\left\|h_\ell^{\boldsymbol{\theta}} - h_\ell^{\boldsymbol{\theta}'}\right\| \le L_\ell \left\{S_\ell \left\|h_{\ell-1}^{\boldsymbol{\theta}} - h_{\ell-1}^{\boldsymbol{\theta}'}\right\| + M_{\ell-1} \|W_\ell - W_\ell'\| + \|b_\ell - b_\ell'\|\right\}.$$

Iterating this finite recursion proves (8), and Proposition 1 completes the entropy calculation.

Example 4 (Hölder balls)**.** Let $\mathcal{X} = [0,1]^d$, fix $s > 0$, and set $r := \lceil s \rceil - 1$, $\gamma := s - r \in (0, 1]$. This convention is important at integer smoothness: when $s = m \in \mathbb{N}$, it requires derivatives of order $m - 1$ to be Lipschitz. For $L > 0$, define

$$\mathcal{H}^s(L) := \left\{f : [0,1]^d \to \mathbb{R} : \max_{|\nu| \le r} \|D^\nu f\|_\infty + \max_{|\nu| = r} \sup_{\boldsymbol{x} \ne \boldsymbol{x}'} \frac{|D^\nu f(\boldsymbol{x}) - D^\nu f(\boldsymbol{x}')|}{\|\boldsymbol{x} - \boldsymbol{x}'\|_2^\gamma} \le L\right\}.$$

There is a constant $C_{d,s} < \infty$, depending only on d and s, such that, for $0 < \varepsilon \leq L$,

$$\log \mathcal{N}\left(\varepsilon, \mathcal{H}^s\left(L\right), \|\cdot\|_\infty\right) \leq C_{d,s}\left(\frac{L}{\varepsilon}\right)^{d/s}. \tag{9}$$

The bound (9) is the standard uniform-norm entropy estimate for Hölder balls; see Theorem 2.7.1 of van der Vaart & Wellner (1996). If the covering centers in that result lie outside $\mathcal{H}^s\left(L\right)$, apply it at radius $\varepsilon/2$ and replace each nonempty covering ball by one member of the class. This gives an internal ε-cover with the same entropy order.

Example 5 (Hard-split trees: infinite uniform-norm entropy)**.** Finite VC dimension does not imply finite uniform-norm entropy. The one-dimensional stump class $\mathcal{T} := \{f_t\left(x\right) = \mathbf{1}\left\{x \leq t\right\} : t \in [0,1]\}$ has VC dimension one, but $\|f_s - f_t\|_\infty = 1$ whenever $s \neq t$. Thus $\mathcal{T}$ contains an uncountable 1-separated subset, and $\mathcal{N}\left(\varepsilon, \mathcal{T}, \|\cdot\|_\infty\right) = \infty$ for every $\varepsilon < 1/2$. Any tree class containing these stumps inherits the obstruction. Accordingly, a tree with continuously varying hard thresholds is outside the metric-entropy argument used in the main text unless its thresholds are discretized or its splits are smoothed; the VC analysis later in the appendix is the appropriate alternative for hard trees.

B.5 Verifying Local Value Stability

Assumption 6 is the only value-sensitivity condition used by the main generalization theorem. The following proposition gives a sufficient condition tailored to the same convex interpolation used by our interpolation path. It requires that the infimum $V\left(\alpha\right)$ is attained by some $\boldsymbol{w}^* \in \mathcal{H}$; write $G^* := g\left(\boldsymbol{w}^*\left(\boldsymbol{x}\right), \boldsymbol{z}\right)$.

Proposition 2 (Local value stability from a strictly feasible reference rule)**.** *Suppose Assumptions 1, 2, 4 hold, and the reference rule has a strict margin:*

$$g\left(\boldsymbol{w}_0\left(\boldsymbol{x}\right), \boldsymbol{z}\right) \leq -\xi_0 \qquad \textit{almost surely}$$

for some $\xi_0 > 0$. Assume $D_0 := \|\boldsymbol{w}_0 - \boldsymbol{w}^*\|_\infty < \infty$ and, for some $\kappa, \bar{r} > 0$,

$$\mathbb{P}\left(0 < G^* \leq r\right) \geq \kappa r \qquad \textit{for } r \in [0, \bar{r}].$$

Then, for $0 \leq t \leq \min\left\{\alpha, \kappa\bar{r}\right\}$,

$$V\left(\alpha - t\right) - V\left(\alpha\right) \leq L_w^c D_0 \frac{t}{\kappa\xi_0 + t} \leq \frac{L_w^c D_0}{\kappa\xi_0} t.$$

Thus Assumption 6 holds with $K_V = L_w^c D_0/(\kappa\xi_0)$ and $\bar{t}_V = \min\left\{\alpha, \kappa\bar{r}\right\}$.

Proof. For $\lambda \in [0,1]$, let $\boldsymbol{w}_\lambda = (1-\lambda)\boldsymbol{w}^* + \lambda\boldsymbol{w}_0$. By Assumption 2, $\boldsymbol{w}_\lambda \in \mathcal{H}$. Convexity gives

$$g\left(\boldsymbol{w}_\lambda\left(\boldsymbol{x}\right), \boldsymbol{z}\right) \leq (1-\lambda)G^* - \lambda\xi_0.$$

Choose $\lambda = t/(\kappa\xi_0 + t)$, so that $\lambda\xi_0/(1-\lambda) = t/\kappa$. The boundary-mass condition then gives $R\left(\boldsymbol{w}_\lambda\right) \geq 1 - \alpha + t$, while Lipschitz continuity of the cost gives $C\left(\boldsymbol{w}_\lambda\right) \leq V\left(\alpha\right) + L_w^c D_0 \lambda$. Combining the two displays proves the result. □

Convexity is only one part of the proposition. Affine constraints and maxima of affine constraints are convex and therefore meet that part automatically, but the conclusion also requires a strict feasibility margin for the reference rule, closure of $\mathcal{H}$ under the interpolation, finite D_0, and positive probability mass immediately above the active boundary. The last condition holds, for example, when G^* has a density bounded below by a positive constant on $[0, \bar{r}]$.

B.6 VC/Rademacher Feasibility and Optimality Guarantees

The VC/Rademacher route controls the scalar satisfaction and cost classes induced by decision rules. The tree calculations below verify the satisfaction-class complexity for hard-split trees; the excess-cost conclusion additionally requires control of the composed cost Rademacher complexity.

Let $\boldsymbol{U} := (\boldsymbol{x}, \boldsymbol{z})$, and write $\mathbb{P}f := \mathbb{E}\left[f\left(\boldsymbol{U}\right)\right]$ and $\mathbb{P}_n f := n^{-1}\sum_{i=1}^n f\left(\boldsymbol{U}_i\right)$. Define

$$\mathcal{C}_{\mathcal{H}} := \left\{\boldsymbol{U} \mapsto c\left(\boldsymbol{w}\left(\boldsymbol{x}\right), \boldsymbol{z}\right) : \boldsymbol{w} \in \mathcal{H}\right\}, \qquad \mathcal{G}_{\mathcal{H}} := \left\{\boldsymbol{U} \mapsto \mathbf{1}\left\{g\left(\boldsymbol{w}\left(\boldsymbol{x}\right), \boldsymbol{z}\right) \leq 0\right\} : \boldsymbol{w} \in \mathcal{H}\right\}.$$

For a scalar class $\mathcal{F}$, define its expected absolute Rademacher complexity by

$$\mathfrak{R}_n(\mathcal{F}) := \mathbb{E}_{\boldsymbol{U}_1^n,\sigma_1^n}\left[\sup_{f\in\mathcal{F}}\left|\frac{1}{n}\sum_{i=1}^n \sigma_i f(\boldsymbol{U}_i)\right|\right],$$

where the σ_i are independent Rademacher variables.

Lemma 4 (Rademacher cost deviation)**.** *Under Assumption 4, for every $\delta \in (0,1)$, with probability at least $1-\delta$,*

$$\sup_{\boldsymbol{w}\in\mathcal{H}} |C_n(\boldsymbol{w}) - C(\boldsymbol{w})| \le \psi_{\mathrm{Rad}}(n, \mathcal{C}_{\mathcal{H}}, \delta) := 2\mathfrak{R}_n(\mathcal{C}_{\mathcal{H}}) + 2B_c\sqrt{\frac{\log(2/\delta)}{2n}}.$$

Proof. The symmetrization inequality gives

$$\mathbb{E}\left[\sup_{f\in\mathcal{C}_{\mathcal{H}}} |\mathbb{P}_n f - \mathbb{P} f|\right] \le 2\mathfrak{R}_n(\mathcal{C}_{\mathcal{H}}).$$

Changing one observation changes the supremum by at most $2B_c/n$. McDiarmid's inequality therefore bounds its deviation above the mean by $2B_c\sqrt{\log(2/\delta)/(2n)}$ with probability at least $1-\delta$. Combining the two displays proves the result. □

Lemma 5 (VC convergence of safety probability)**.** *Suppose $\mathcal{G}_{\mathcal{H}}$ has VC dimension $v_g < \infty$. There is a universal constant $K_{\mathrm{VC}} > 0$ such that, for every $\delta \in (0,1)$, with probability at least $1-\delta$,*

$$\sup_{\boldsymbol{w}\in\mathcal{H}} |R_n(\boldsymbol{w}) - R(\boldsymbol{w})| \le \phi_{\mathrm{VC}}(n, v_g, \delta) := K_{\mathrm{VC}}\sqrt{\frac{v_g \vee 1}{n}} + 2\sqrt{\frac{\log(2/\delta)}{2n}}.$$

Proof. Applying the symmetrization and bounded-difference argument from Lemma 4 to the binary class $\mathcal{G}_{\mathcal{H}}$ yields

$$\sup_{h\in\mathcal{G}_{\mathcal{H}}} |\mathbb{P}_n h - \mathbb{P} h| \le 2\mathfrak{R}_n(\mathcal{G}_{\mathcal{H}}) + 2\sqrt{\frac{\log(2/\delta)}{2n}}.$$

If $v_g \ge 1$, the VC covering lemma gives, uniformly over discrete probability measures Q and $0 < \varepsilon \le 1$,

$$\mathcal{N}(\varepsilon, \mathcal{G}_{\mathcal{H}}, L_2(Q)) \le (A_0/\varepsilon)^{2v_g}$$

for a universal A_0. Dudley's entropy integral then yields $\mathfrak{R}_n(\mathcal{G}_{\mathcal{H}}) \le K_0\sqrt{v_g/n}$. If $v_g = 0$, all functions in $\mathcal{G}_{\mathcal{H}}$ coincide. For their common representative h_0, conditional Jensen's inequality gives

$$\mathbb{E}_\sigma\left|\frac{1}{n}\sum_{i=1}^n \sigma_i h_0(\boldsymbol{U}_i)\right| \le \frac{1}{n}\sqrt{\sum_{i=1}^n h_0(\boldsymbol{U}_i)^2} \le \frac{1}{\sqrt{n}}.$$

Thus $\mathfrak{R}_n(\mathcal{G}_{\mathcal{H}}) \le K_0\sqrt{(v_g \vee 1)/n}$ in both cases. Absorbing $2K_0$ into K_{VC} proves the claim. □

Corollary 1 (VC/Rademacher generalization guarantee)**.** *Let $\delta_1, \delta_2 \in (0,1)$ satisfy $\delta_1 + \delta_2 < 1$, let $\zeta_n \ge 0$, and let $\widehat{\boldsymbol{w}}$ be a ζ_n-optimal rule in the sense of (7), with margin $r_n = \phi_{\mathrm{VC}}(n, v_g, \delta_1)$; in particular,*

$$R_n(\widehat{\boldsymbol{w}}) \ge 1 - \alpha + \phi_{\mathrm{VC}}(n, v_g, \delta_1).$$

Under the assumptions of Lemmas 4 and 5, if $r_n \le \alpha$, then

$$\mathbb{P}_{\mathcal{D}_n}(R(\widehat{\boldsymbol{w}}) \ge 1 - \alpha) \ge 1 - \delta_1.$$

If, in addition, Assumption 6 holds and $2r_n \le \bar{t}_V$, then

$$\mathbb{P}_{\mathcal{D}_n}(C(\widehat{\boldsymbol{w}}) - V(\alpha) \le 2K_V r_n + 2\psi_{\mathrm{Rad}}(n, \mathcal{C}_{\mathcal{H}}, \delta_2) + \zeta_n) \ge 1 - \delta_1 - \delta_2.$$

Proof. If $r_n \le \alpha$, the safety-probability deviation event gives $R(\widehat{\boldsymbol{w}}) \ge R_n(\widehat{\boldsymbol{w}}) - r_n \ge 1 - \alpha$. For $2r_n \le \bar{t}_V \le \alpha$, apply Lemma 3 with $s_n = \psi_{\mathrm{Rad}}(n, \mathcal{C}_{\mathcal{H}}, \delta_2)$. The lemma gives $C(\widehat{\boldsymbol{w}}) \le V(\alpha - 2r_n) + 2s_n + \zeta_n$ on the joint deviation event. Assumption 6 at $t = 2r_n$ gives the stated bound. □

B.6.1 Composition through bounded partition complexity

We verify the VC premise for piecewise-constant rules by separating the constraint geometry from how the rules partition contexts. Define the action-indexed satisfaction class

$$\mathcal{G}_0 := \{\boldsymbol{z} \mapsto \mathbf{1}\{g(\boldsymbol{w}, \boldsymbol{z}) \le 0\} : \boldsymbol{w} \in \Omega\}.$$

For any binary class $\mathcal{A}$, let $\Pi_{\mathcal{A}}(m)$ denote its shatter coefficient, the largest number of label vectors induced on m points. For contexts $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$, each $\boldsymbol{w} \in \mathcal{H}$ induces a partition of $[m]$ according to equality of the action values $\boldsymbol{w}(\boldsymbol{x}_i)$. Let $N_{\mathcal{H}}(m)$ be the maximum number of such partitions over all context samples.

Theorem 4 (Composition with bounded partition complexity)**.** *Suppose $v := \mathrm{VCdim}(\mathcal{G}_0) < \infty$, $L \in \mathbb{N}$, $A \ge 1$, and $\nu \ge 0$. Assume every $\boldsymbol{w} \in \mathcal{H}$ takes at most L distinct action values on every finite context sample and $N_{\mathcal{H}}(m) \le Am^{\nu}$. If $v \ge 1$, then, for every $m \ge 1$,*

$$\Pi_{\mathcal{G}_{\mathcal{H}}}(m) \le N_{\mathcal{H}}(m)\,[e(1 + m/v)]^{Lv} \le Am^{\nu}\,[e(1 + m/v)]^{Lv}.$$

Consequently,

$$\mathrm{VCdim}(\mathcal{G}_{\mathcal{H}}) = O((\nu + Lv)\log(2 + \nu + Lv) + \log(2A)).$$

If $v = 0$, all members of $\mathcal{G}_0$ coincide and $\Pi_{\mathcal{G}_{\mathcal{H}}}(m) = 1$.

Proof. For $v \ge 1$, fix $S = \{(\boldsymbol{x}_i, \boldsymbol{z}_i)\}_{i=1}^{m}$ and condition on one induced partition $I_1, \dots, I_k$, where $k \le L$. On block I_ℓ, $\boldsymbol{w}(\boldsymbol{x}_i)$ is constant, so the labels are generated by $\mathcal{G}_0$ on $|I_\ell|$ outcome points. Dropping the restriction that the block actions arise from a common rule, the number of labels for this partition is at most

$$\prod_{\ell=1}^{k} \Pi_{\mathcal{G}_0}(|I_\ell|) \le \prod_{\ell=1}^{k} [e(1 + |I_\ell|/v)]^{v} \le [e(1 + m/v)]^{Lv},$$

where the first inequality is Sauer–Shelah. Multiplying by the number of induced partitions proves the growth bound. For $m \ge v$, set $q = \nu + Lv$ and bound the right-hand side by $A(2e)^q m^q$. If m is shattered, $2^m \le A(2e)^q m^q$; standard inversion gives $m \le K_{\mathrm{poly}}\{q\log(2q) + \log(2A)\}$ for a universal K_{poly}. Values $m < v$ have the same order. The case $v = 0$ is immediate. □

Corollary 2 (Axis-aligned regression trees)**.** *Consider vector-valued axis-aligned regression trees on $\mathbb{R}^{d_x}$ with a shared partition and at most L leaves. Then*

$$N_{\mathcal{H}}(m) \le \mathrm{Bell}(L)\,(8d_x)^{L-1}\,m^{L-1},$$

where $\mathrm{Bell}(L)$ counts the partitions of L leaves (Rota, 1964). Thus Theorem 4 applies with $\nu = L - 1$. Bounded leaf count alone is not sufficient for an arbitrary split family; the conclusion uses controlled axis-aligned splits in fixed context dimension.

Proof. A binary tree with L leaves has at most $L - 1$ internal nodes and at most 4^{L-1} rooted binary shapes. On a sample of size m, an axis-aligned node has at most $d_x(m + 1) \le 2d_x m$ distinct splits. Finally, equality among the values assigned to different leaves can coarsen the leaf partition in at most $\mathrm{Bell}(L)$ ways. Multiplying these bounds proves the display. Separate coordinate-specific trees can be handled by multiplying their partition bounds, with the corresponding larger exponent. □

B.6.2 Action-indexed satisfaction classes

The universal VC premise is simply $\mathrm{VCdim}(\mathcal{G}_{\mathcal{H}}) < \infty$. For the bounded-partition reduction above, it remains to control $v = \mathrm{VCdim}(\mathcal{G}_0)$. The following examples cover the affine constraints used in our newsvendor and portfolio formulations, as well as a finite maximum used to encode a joint constraint.

Example 6 (Affine constraint)**.** If $\Omega \subseteq \mathbb{R}^{d_w}$ and $g(\boldsymbol{w}, \boldsymbol{z}) = a(\boldsymbol{z})^{\top}\boldsymbol{w} - b(\boldsymbol{z})$, then $\mathcal{G}_0$ is a subclass of the pullback of homogeneous halfspaces in $\mathbb{R}^{d_w+1}$ under the context map $\boldsymbol{z} \mapsto (a(\boldsymbol{z}), -b(\boldsymbol{z}))$. Hence $v \le d_w + 1$.

Example 7 (Maximum of affine constraints)**.** If $g(\boldsymbol{w}, \boldsymbol{z}) = \max_{j\in[M]} \left\{ a_j(\boldsymbol{z})^\top \boldsymbol{w} - b_j(\boldsymbol{z}) \right\}$, then satisfaction is an intersection of M affine satisfaction events. Allowing a separate parameter for each event only enlarges the class, so the product of the M halfspace growth functions gives

$$v = O\left((d_w + 1) M \log(2M)\right).$$

Each conclusion follows because pullbacks cannot increase the shatter coefficient. In the max-affine case, put $d = d_w + 1$; Sauer–Shelah gives $\Pi_{\mathcal{G}_0}(m) \le [e(1 + m/d)]^{Md}$, and inversion yields the displayed order.

C Proofs in Section 3

Before proving Theorems 1 and 2, we provide the proof of Lemma 1.

Proof of Lemma 1. For $j < k$, set $\theta_{jk} := (\lambda_k - \lambda_j)/(1 - \lambda_j) \in [0, 1]$. Then $\boldsymbol{w}_k = (1 - \theta_{jk})\boldsymbol{w}_j + \theta_{jk}\boldsymbol{w}_0$. Whenever $g(\boldsymbol{w}_j(\boldsymbol{x}), \boldsymbol{z}) \le 0$, convexity and the safety of $\boldsymbol{w}_0$ imply

$$g(\boldsymbol{w}_k(\boldsymbol{x}), \boldsymbol{z}) \le (1 - \theta_{jk})\, g(\boldsymbol{w}_j(\boldsymbol{x}), \boldsymbol{z}) + \theta_{jk} g(\boldsymbol{w}_0(\boldsymbol{x}), \boldsymbol{z}) \le 0.$$

Taking probabilities yields the ordering, and $R_p = 1$ because $\lambda_p = 1$. □

Proof of Theorem 1. Condition on $\mathscr{T}$. For the n_{val} validation pairs and one fresh pair, let $T_i \in [p]$ be the first-satisfaction index. These $n_{\text{val}} + 1$ indices are i.i.d., and $1 - Y_{ij} = \mathbf{1}\{T_i > j\}$. If $\alpha \ge 1/(n_{\text{val}} + 1)$, set $k := \lceil (n_{\text{val}} + 1)(1 - \alpha) \rceil \le n_{\text{val}}$. The conformal acceptance criterion is equivalent to requiring at least k of the validation indices to be no larger than j, so j^*_{C} is their kth order statistic. Exchangeability of the $n_{\text{val}} + 1$ indices gives

$$\mathbb{P}\left(T_{n_{\text{val}}+1} > j^*_{\mathsf{C}} \mid \mathscr{T}\right) \le \frac{n_{\text{val}} + 1 - k}{n_{\text{val}} + 1} \le \alpha.$$

The left-hand side equals $\mathbb{E}_{\mathcal{D}_{\text{val}}}\left[1 - R_{j^*_{\mathsf{C}}} \mid \mathscr{T}\right]$. If $\alpha < 1/(n_{\text{val}} + 1)$, the rule returns the reference rule and the same conclusion is immediate. □

Lemma 6 (First-satisfaction-index DKW bound)**.** *Suppose, conditional on $\mathscr{T}$, that $Y_{i1} \le \cdots \le Y_{ip} = 1$ almost surely. Then, conditional on $\mathscr{T}$ almost surely,*

$$\mathbb{P}_{\mathcal{D}_{\text{val}}}\left(R_j \ge \widehat{R}_j - M_j^{\mathsf{D}} \text{ for every } j \in [p] \,\middle|\, \mathscr{T} \right) \ge 1 - \beta.$$

Proof of Lemma 6. Define $T_i := \min\{j \in [p] : Y_{ij} = 1\}$. The nesting assumption makes these first-satisfaction indices well defined almost surely. Conditional on $\mathscr{T}$, they are i.i.d., $\widehat{R}_j$ is their empirical distribution function evaluated at j, and $R_j = \mathbb{P}(T_i \le j \mid \mathscr{T})$. Because $\beta < 1/2$, the sharp one-sided DKW inequality of Massart (1990) applies with the common margin M_j^{D} and gives

$$\mathbb{P}\left(\widehat{R}_j - R_j > M_j^{\mathsf{D}} \text{ for some } j \in [p] \,\middle|\, \mathscr{T} \right) \le \beta.$$

On the complementary event, acceptance implies $R_j \ge \widehat{R}_j - M_j^{\mathsf{D}} \ge 1 - \alpha$. □

Lemma 7 (Empirical-Gaussian maximum bound for bounded coordinates)**.** *Let $X_1, \ldots, X_n \in [0, 1]^p$ be i.i.d. with $pn \ge 2$, let $\mu_j = \mathbb{E}X_{1j}$, $\sigma_j^2 = \operatorname{var}(X_{1j})$, and $\bar{X}_j := n^{-1}\sum_{i=1}^n X_{ij}$, and form $\widehat{\Sigma}$, $\widehat{\sigma}_j$, $q^{\mathsf{U}}_{1-\beta}$, and $q^{\mathsf{N}}_{1-\beta}$ as in Section 2.1. Fix a coordinate set J such that $\sigma_j^2 \ge b$ for every $j \in J$, where $b \in (0, 1/4)$. Suppose there is an envelope constant $B_{\mathrm{G}} \ge 1$ such that, for every $j \in J$ and $k \in \{1, 2\}$,*

$$\mathbb{E}\exp\left\{ \frac{|X_{1j} - \mu_j|^2}{B_{\mathrm{G}}^2 \sigma_j^2} \right\} \le 2, \qquad \mathbb{E}\left\{ \frac{|X_{1j} - \mu_j|}{\sigma_j} \right\}^{2+k} \le B_{\mathrm{G}}^k.$$

There are universal constants $K_{\mathrm{G}}, k_{\mathrm{G}} > 0$ such that

$$\mathbb{P}\left(\mu_j \geq \bar{X}_j - \frac{q^{\mathsf{U}}_{1-\beta}}{\sqrt{n}} \text{ for all } j \in J\right) \geq 1 - \beta - K_{\mathrm{G}}\left(\frac{B_{\mathrm{G}}^2 \log^7(pn)}{n}\right)^{1/6},$$

$$\mathbb{P}\left(\mu_j \geq \bar{X}_j - \frac{q^{\mathsf{N}}_{1-\beta}\widehat{\sigma}_j}{\sqrt{n}} \text{ for all } j \in J\right) \geq 1 - \beta - K_{\mathrm{G}}\left\{\begin{array}{c}\left(\frac{B_{\mathrm{G}}^2 \log^7(pn)}{n}\right)^{1/6} + \frac{\log^2(pn)}{\sqrt{nb}} \\ + p\exp\left(-k_{\mathrm{G}} b B_{\mathrm{G}}^{2/3} n^{2/3}\right)\end{array}\right\}.$$

Proof of Lemma 7. The first display follows from Theorem EC.9 of Lam & Qian (2019), applied to the subvector indexed by J. The second is the bounded-coordinate conclusion of their Theorem EC.10, with the variance lower bound b. Those results are derived from the high-dimensional Gaussian and empirical-bootstrap approximations of Chernozhukov et al. (2017).

The cited theorems use conditional critical values for the maximum over J. The critical values $q^{\mathsf{U}}_{1-\beta}$ and $q^{\mathsf{N}}_{1-\beta}$ used here maximize over all eligible coordinates, so they are no smaller than their restrictions to J. Replacing $|J|$ by p in the error bounds and using these conservative critical values gives the two displays. □

Lemma 8 (Simultaneous Gaussian lower bounds)**.** *Let $\widetilde{\alpha} := \min\{\alpha, 1-\alpha\}$, suppose $pn_{\mathrm{val}} \geq 2$, and define*

$$\mathcal{J}_0 := \{j : \widetilde{\alpha}/2 < R_j < 1 - \alpha\}.$$

There are universal constants $K_{\mathrm{G}}, k_{\mathrm{G}} > 0$ such that, conditional on $\mathscr{T}$ almost surely,

$$\mathbb{P}_{\mathcal{D}_{\mathrm{val}}}\left(R_j \geq \widehat{R}_j - M_j^{\mathsf{U}} \text{ for all } j \in \mathcal{J}_0 \mid \mathscr{T}\right) \geq 1 - \beta - \Delta_{\mathsf{U}},$$

$$\mathbb{P}_{\mathcal{D}_{\mathrm{val}}}\left(R_j \geq \widehat{R}_j - M_j^{\mathsf{N}} \text{ for all } j \in \mathcal{J}_0 \mid \mathscr{T}\right) \geq 1 - \beta - \Delta_{\mathsf{N}},$$

where

$$\Delta_{\mathsf{U}} := K_{\mathrm{G}}\left(\frac{\log^7(pn_{\mathrm{val}})}{n_{\mathrm{val}}\widetilde{\alpha}}\right)^{1/6},$$

$$\Delta_{\mathsf{N}} := K_{\mathrm{G}}\left\{\left(\frac{\log^7(pn_{\mathrm{val}})}{n_{\mathrm{val}}\widetilde{\alpha}}\right)^{1/6} + \frac{\log^2(pn_{\mathrm{val}})}{\sqrt{n_{\mathrm{val}}\widetilde{\alpha}}} + p\exp\left(-k_{\mathrm{G}}\left(n_{\mathrm{val}}\widetilde{\alpha}\right)^{2/3}\right)\right\}.$$

Proof of Lemma 8. Conditional on $\mathscr{T}$, the vectors $(Y_{i1}, \ldots, Y_{ip})$ are i.i.d. in $[0,1]^p$. The concavity of $r \mapsto r(1-r)$ and the endpoints of $(\widetilde{\alpha}/2, 1-\alpha)$ give, for $j \in \mathcal{J}_0$,

$$\sigma_j^2 = R_j(1 - R_j) \geq \widetilde{\alpha}/4.$$

Moreover, $|Y_{ij} - R_j| \leq 1$. With $b = \widetilde{\alpha}/4$ and $B_{\mathrm{G}} = K_0\widetilde{\alpha}^{-1/2}$ for a sufficiently large universal constant K_0,

$$\mathbb{E}\exp\left\{\frac{|Y_{ij} - R_j|^2}{B_{\mathrm{G}}^2\sigma_j^2}\right\} \leq 2, \qquad \mathbb{E}\left\{\frac{|Y_{ij} - R_j|}{\sigma_j}\right\}^{2+k} \leq \sigma_j^{-k} \leq B_{\mathrm{G}}^k, \qquad k = 1, 2.$$

Thus all assumptions of Lemma 7 are verified. Substitution of b and B_{G}, and enlargement of universal constants, gives the two stated bounds. This argument also covers coordinates whose empirical variance happens to be zero through the explicit studentization-error terms. □

Proof of Theorem 2. (i) Feasibility. For $\mathsf{m} = \mathsf{D}$, let $\mathcal{E}_{\mathsf{D}}$ be the simultaneous DKW event in Lemma 6. On $\mathcal{E}_{\mathsf{D}}$, every accepted candidate has safety probability at least $1 - \alpha$. If no candidate is accepted, Algorithm 1 returns the reference rule. Therefore the selected rule is population feasible on $\mathcal{E}_{\mathsf{D}}$, which has conditional probability at least $1 - \beta = 1 - \beta - \mathfrak{E}_{\mathsf{D}}$.

Now fix $\mathsf{m} \in \{\mathsf{N}, \mathsf{U}\}$ and let $\mathcal{E}_{\mathsf{m}}$ be the simultaneous event in Lemma 8. Define

$$\mathcal{E}_{\mathrm{low}} := \left\{\widehat{R}_j < \widetilde{\alpha} + M_j^{\mathsf{m}} \text{ for every } j \text{ with } R_j \leq \widetilde{\alpha}/2\right\}.$$

Since $M_j^{\mathsf{m}} \geq 0$, Bernstein's inequality for a Bernoulli mean gives, for every such j,

$$\mathbb{P}\left(\widehat{R}_j \geq \widetilde{\alpha} + M_j^{\mathsf{m}} \mid \mathscr{T}\right) \leq \mathbb{P}\left(\widehat{R}_j \geq \widetilde{\alpha} \mid \mathscr{T}\right) \leq \exp\left(-k_{\mathrm{G}} n_{\mathrm{val}} \widetilde{\alpha}\right).$$

The union bound therefore gives $\mathbb{P}\left(\mathcal{E}_{\mathrm{low}}^c \mid \mathscr{T}\right) \leq p \exp\left(-k_{\mathrm{G}} n_{\mathrm{val}} \widetilde{\alpha}\right)$.

Consider any population-infeasible candidate, so $R_j < 1 - \alpha$. If $R_j > \widetilde{\alpha}/2$, then $j \in \mathcal{J}_0$ and, on $\mathcal{E}_{\mathsf{m}}$,

$$\widehat{R}_j - M_j^{\mathsf{m}} \leq R_j < 1 - \alpha,$$

so candidate j fails the acceptance test. If $R_j \leq \widetilde{\alpha}/2$, then, on $\mathcal{E}_{\mathrm{low}}$,

$$\widehat{R}_j - M_j^{\mathsf{m}} < \widetilde{\alpha} \leq 1 - \alpha,$$

so this candidate also fails. Thus no population-infeasible candidate is accepted on $\mathcal{E}_{\mathsf{m}} \cap \mathcal{E}_{\mathrm{low}}$.

If any candidates are accepted, the algorithm minimizes cost over population-feasible candidates only; otherwise it returns the reference rule. Hence the selected rule is population feasible on $\mathcal{E}_{\mathsf{m}} \cap \mathcal{E}_{\mathrm{low}}$. The two failure bounds and a union bound give

$$\mathbb{P}_{\mathcal{D}_{\mathrm{val}}}\left(R_{j_{\mathsf{m}}^*} \geq 1 - \alpha \mid \mathscr{T}\right) \geq 1 - \beta - \mathfrak{E}_{\mathsf{m}}.$$

(ii) Pathwise optimality. Allocate failure probability $\delta/2$ to each of the safety-probability and cost deviations. For the method m, define $\bar{m}_{\mathsf{m}} := \sqrt{\frac{\log(\kappa_{\mathsf{m}}/\beta)}{2n_{\mathrm{val}}}}$ and $\eta_{R,\mathsf{m}} := \sqrt{\frac{\log(2\kappa_{\mathsf{m}}/\delta)}{2n_{\mathrm{val}}}}$. For method D, $M_j^{\mathsf{D}} = \bar{m}_{\mathsf{D}}$. For method N, a Gaussian union bound gives $q_{1-\beta}^{\mathsf{N}} \leq \sqrt{2\log(p/\beta)}$, and $\widehat{\sigma}_j \leq 1/2$ gives $M_j^{\mathsf{N}} \leq \bar{m}_{\mathsf{N}}$. For method U, each Z_j has conditional variance at most $1/4$, so the same union bound gives $q_{1-\beta}^{\mathsf{U}} \leq \sqrt{\log(p/\beta)/2}$ and $M_j^{\mathsf{U}} \leq \bar{m}_{\mathsf{U}}$. These statements include $p = 1$ and degenerate Gaussian coordinates.

On any nested path, the first-satisfaction indices from Lemma 6 make $\widehat{R}_j$ and R_j an empirical and a population distribution function. The lower-tail one-sided DKW inequality (Reeve, 2024), applied to the reflected first-satisfaction indices, therefore gives the following event with conditional probability at least $1 - \delta/2$, simultaneously over all j and without a union bound, yielding $\eta_{R,\mathsf{D}}$. For the Gaussian calibrators on a finite candidate set, Hoeffding's inequality and a union bound give the same event with their respective $\eta_{R,\mathsf{m}}$:

$$\widehat{R}_j \geq R_j - \eta_{R,\mathsf{m}} \qquad \text{for all } j \in [p].$$

On this event, every j with $R_j \geq 1 - \alpha + \eta_{\mathsf{m}}^R$ is accepted because $\widehat{R}_j - M_j^{\mathsf{m}} \geq R_j - \eta_{R,\mathsf{m}} - \bar{m}_{\mathsf{m}} \geq 1 - \alpha$. Since costs lie in $[-B_c, B_c]$, Hoeffding's inequality and a union bound give

$$\max_{j \in [p]} \left|\widehat{C}_j - C_j\right| \leq \varepsilon_{\mathrm{cost}}/2.$$

The complementary event has conditional probability at most $2p \exp\left(-n_{\mathrm{val}} \varepsilon_{\mathrm{cost}}^2/(8B_c^2)\right) = \delta/2$. The comparator set is nonempty because $R_p = 1$ and $\eta_{\mathsf{m}}^R \leq \alpha$. On the intersection of the two concentration events, every comparator belongs to $\widehat{\mathcal{J}}_{\mathsf{m}}$, and empirical optimality gives, for each such comparator,

$$C_{j_{\mathsf{m}}^*} \leq \widehat{C}_{j_{\mathsf{m}}^*} + \varepsilon_{\mathrm{cost}}/2 \leq \widehat{C}_j + \varepsilon_{\mathrm{cost}}/2 \leq C_j + \varepsilon_{\mathrm{cost}}.$$

Taking the minimum and applying the union bound gives probability at least $1 - \delta$ in part (ii). For either Gaussian calibrator on a nested path, the same DKW step replaces $\eta_{R,\mathsf{m}}$ by $\sqrt{\log(2/\delta)/(2n_{\mathrm{val}})}$ in the pathwise bound and its condition $\eta_{\mathsf{m}}^R \leq \alpha$, while leaving $\bar{m}_{\mathsf{m}}$ unchanged. □

D Detailed Experimental Setups and Additional Results

D.1 Synthetic Data-Generating Processes in Section 4.2

Let $a := \sqrt{3}$. All context coordinates and primitive shocks below are independent $\mathrm{Unif}[-a, a]$ variables, and hence have mean zero and variance one. Both applications use five context coordinates

$\boldsymbol{x} = (x_1, \ldots, x_5)$. Define the common context-dependent dispersion factor

$$s(\boldsymbol{x}) := \frac{1}{2} + \frac{1}{1 + \exp\{-2x_2x_4\}},$$

which lies in $(1/2, 3/2)$. The control DGP uses shock multiplier one, and the stress DGP uses $s(\boldsymbol{x})$. Thus control and stress have the same signal functions and bounded primitive shocks and differ in conditional dispersion.

Newsvendor. With an independent shock ϵ, define

$$m_{\mathrm{nv}}(\boldsymbol{x}) := 8 + 1.5x_1 - x_2 + 1.2x_1x_3 + 0.8x_4^2 + \sin x_5,$$
$$z := [m_{\mathrm{nv}}(\boldsymbol{x}) + \kappa(\boldsymbol{x})\,\epsilon]_+,$$

where $\kappa(\boldsymbol{x}) = 1$ in the control DGP and $\kappa(\boldsymbol{x}) = s(\boldsymbol{x})$ in the stress DGP. The reference rule is the analytical demand-support cap. In particular, with

$$M := 8 + a\max\{|1.5 - 1.2a|, |1.5 + 1.2a|\} + a + 0.8a^2 + 1,$$

the constant caps $M + a$ and $M + 1.5a$ are valid for the control and stress DGPs, respectively; the DGP determines both caps.

Portfolio. The cash return is $z_0 = .002$. The four risky conditional means are

$$m_1(\boldsymbol{x}) = .0100 + .006\,(.9x_1 - .3x_2),$$
$$m_2(\boldsymbol{x}) = .0105 + .006\,(-.8x_1 + .7x_2),$$
$$m_3(\boldsymbol{x}) = .0110 + .006\,(.7x_1x_3 - .25x_4),$$
$$m_4(\boldsymbol{x}) = .0100 + .006\left(.6\left(x_4^2 - 1\right) + .55\sin x_5\right).$$

Using one common shock ϵ_0 and four idiosyncratic shocks $\epsilon_1, \ldots, \epsilon_4$, the risky returns are

$$z_j = m_j(\boldsymbol{x}) + \kappa(\boldsymbol{x})\,(.010\epsilon_0 + .018\epsilon_j), \qquad j = 1, \ldots, 4.$$

Again $\kappa(\boldsymbol{x}) = 1$ in control and $s(\boldsymbol{x})$ in stress. The all-cash rule serves as the reference rule because its return .002 always exceeds the downside floor $-.01$.

D.2 Reference-Rule Constructions

For newsvendor, the demand-support bounds above give the constant rules $\boldsymbol{w}_0(\boldsymbol{x}) = M + a$ for the control DGP and $\boldsymbol{w}_0(\boldsymbol{x}) = M + 1.5a$ for the stress DGP, each meeting every demand realization. For portfolio, the all-cash rule has return .002 in the synthetic DGP for both the control and stress cases, and zero return in excess of cash in the real data evaluation, meeting the downside floor $-.01$.

These reference rules give explicit feasible candidates and allow the same interpolation and validation procedures to be applied to the four problems.

D.3 Algorithmic Implementation

Decision-Rule Classes. The linear class uses an affine score in the context, and the neural class uses a fully connected ReLU network with two hidden layers of width 64. In the synthetic newsvendor study, the scalar score is mapped through a sigmoid and multiplied by the analytical demand cap. For synthetic portfolio, five scores are mapped through a softmax to enforce nonnegative weights summing to one.

The nonparametric RF–SAA class fits a 100-tree random forest with minimum leaf size five, square-root context subsampling, and at most 256 sampled training observations per tree. Forest co-occurrence weights define a local empirical distribution at each context. The newsvendor decision is its weighted 75th percentile; the portfolio decision selects the asset with the largest weighted conditional mean return. Calibration treats this pretrained rule as a black box.

Training and Calibration. The first stage fits $\boldsymbol{w}^{\text{pre}}(\cdot)$ directly for the downstream application. For the linear and neural classes parameterized by θ, we write the decision rule as $\boldsymbol{w}_\theta(\cdot)$ and approximate the violation indicator $\mathbf{1}\{u > 0\}$ by $S_{h_{\text{sm}}}(u) := \Phi(u/h_{\text{sm}})$, where $h_{\text{sm}} > 0$ is the smoothing bandwidth and Φ is the standard normal cumulative distribution function. We solve the smoothed empirical program

$$\min_{\theta} \quad C_n(\boldsymbol{w}_\theta) = \frac{1}{n}\sum_{i=1}^{n} c(\boldsymbol{w}_\theta(\boldsymbol{x}_i), \boldsymbol{z}_i)$$

$$\text{s.t.} \quad \widehat{h}(\theta) := \frac{1}{n}\sum_{i=1}^{n} S_{h_{\text{sm}}}(g(\boldsymbol{w}_\theta(\boldsymbol{x}_i), \boldsymbol{z}_i)) - \alpha \le 0.$$

Writing $\widehat{L}(\theta, \mu) := C_n(\boldsymbol{w}_\theta) + \mu\widehat{h}(\theta)$ for multiplier $\mu \ge 0$, we use the Adam optimizer on minibatches to decrease $\widehat{L}$ in θ and projected dual ascent $\mu \leftarrow [\mu + \eta_\mu \widehat{h}(\theta)]_+$ with step size $\eta_\mu > 0$. The synthetic bandwidths are $h_{\text{sm}} = .1$ for newsvendor and .005 for portfolio; the corresponding real-data values are .02 and .005. The RF–SAA extension evaluates the calibration interface with a nonparametric, nondifferentiable pretrained rule.

The second stage fixes the pretrained rule before using validation data. Across the synthetic study we evaluate conformal risk calibration (C), DKW calibration (D), and normalized and unnormalized Gaussian calibration (N/U). Here method C applies the corrected empirical binary risk directly to our interpolation path. For comparisons among calibrators, we use the same $\boldsymbol{w}^{\text{pre}}(\cdot)$, $\boldsymbol{w}_0(\cdot)$, and interpolation path. Gaussian critical values use 10,000 draws and a binomial-certified upper order statistic; their theoretical interpretation retains the approximation errors in part (i) of Theorem 2.

D.4 Running Time Discussion

Each synthetic application–class–DGP combination uses 3,600 observations for pretraining. Calibration evaluates the 101 path candidates on the safety-parameter grid $\lambda_j \in \{0, .01, \ldots, 1\}$ using validation data, and performance is evaluated on test data with 10,000 independent observations. The primary validation size is 1,200, with nested samples of sizes 300, 600, 800, and 1,000 used for the sample-size sensitivity. The four calibrators use exactly the same pretrained rule, reference rule, safety-parameter grid, validation data, and test data.

For each class and DGP, the primary study has 200 repetitions that keep the pretrained rule fixed and redraw validation and test data. Ten additional repetitions redraw all three samples and repeat pretraining. Paired random seeds align comparisons across calibrators and confidence levels. Pooled summaries give equal weight to the three classes and two DGPs; counts sum over the corresponding repetitions. Chance-constraint targets are marginal over $(\boldsymbol{x}, \boldsymbol{z})$.

Calibration overhead. Table 3 reports one measured pretraining time and the mean complete calibration-stage time. The latter includes validation prediction, summarizing all 101 path candidates, and applying the specified calibrator.

Table 3: Runtime (seconds) by application, DGP, and pretrained decision-rule class. Pretrain reports one fit; calibrator column averages 200 complete calibration runs. Methods D/N/U use 95% confidence.

Application	DGP	Rule class	Pretrain	C	D	N	U
Newsvendor	Control	Linear	1.722	.002	.002	.009	.012
		Neural	2.753	.003	.003	.007	.012
		RF–SAA	1.140	.159	.159	.164	.167
	Stress	Linear	1.415	.003	.003	.010	.014
		Neural	5.572	.004	.004	.007	.013
		RF–SAA	1.167	.169	.169	.174	.177
Portfolio	Control	Linear	4.340	.004	.004	.015	.019
		Neural	1.047	.004	.005	.015	.019
		RF–SAA	1.751	.616	.616	.631	.635
	Stress	Linear	28.794	.005	.005	.021	.024
		Neural	2.650	.006	.006	.024	.026
		RF–SAA	1.776	.838	.838	.862	.865

D.5 Complete Real Data Results

Table 4 gives all real data outcomes, including the uncalibrated pretrained rules. Within each dataset and class, the four calibrators use the same pretrained rule, validation observations, and 101 path candidates. Newsvendor calibration uses one binary stockout indicator per SKU–day.

For real data newsvendor, the reference order for each series is 1.5 times its largest training-period demand, with a minimum of one; known Restaurant closure days receive zero orders. This rule has zero violations on validation data and on the reported test windows: 0/2,660 for Bakery, 0/196 for Restaurant, and 0/14,000 for SID. The portfolio reference is all cash, giving zero excess return and 0/186 violations in each panel. Across all 21 dataset–class combinations, an observation-wise check of the 101 test candidates found zero transitions from constraint satisfaction to violation as λ increased. These real data studies provide empirical evidence; temporal dependence and the empirical demand bounds limit the applicability of the exchangeability and population-safety assumptions underlying the formal guarantees.

Table 4: Complete real data test results. Entries report objective value / violation rate (%). Objective value is normalized cost for newsvendor and annualized excess return (%) for portfolio. *Raw* is the uncalibrated pretrained rule. Bold marks best objective among entries with violation at most 10%.

Dataset / panel	Pretrained rule class	Raw	C	D	N	U
Newsvendor						
Bakery	Linear	**.974** / 4.21	**.974** / 4.21	**.974** / 4.21	**.974** / 4.21	**.974** / 4.21
	Neural	.386 / 17.14	.411 / 12.41	.429 / 10.26	.411 / 12.41	.411 / 12.41
	RF–SAA	**.495** / 7.22	.519 / 6.24	.543 / 5.64	.519 / 6.24	.519 / 6.24
Restaurant	Linear	**1.101** / 6.63	**1.101** / 6.63	1.310 / 2.04	1.186 / 4.59	1.164 / 5.10
	Neural	.975 / 10.71	**1.126** / 6.12	1.334 / 3.57	1.227 / 4.08	1.227 / 4.08
	RF–SAA	.570 / 23.47	**.785** / 5.10	.945 / 2.55	.847 / 4.08	.879 / 3.06
SID	Linear	**.360** / 1.56	**.360** / 1.56	**.360** / 1.56	**.360** / 1.56	**.360** / 1.56
	Neural	.193 / 27.51	**.246** / 6.75	**.246** / 6.75	**.246** / 6.75	**.246** / 6.75
	RF–SAA	.229 / 38.16	**.375** / 1.97	.393 / 1.58	.393 / 1.58	.393 / 1.58
Portfolio						
Size–BM 6	Linear	6.96 / 24.73	2.20 / 11.29	.99 / 3.23	**1.58** / 7.53	1.27 / 4.84
	Neural	4.83 / 13.98	**1.97** / 9.14	1.05 / 3.76	1.62 / 6.45	1.52 / 5.91
	RF–SAA	7.30 / 19.89	1.43 / 10.22	.69 / 3.76	**1.10** / 8.06	.92 / 6.45
Industry 10	Linear	7.54 / 25.27	**1.87** / 9.14	.93 / 2.15	1.41 / 6.99	1.03 / 3.76
	Neural	6.24 / 18.28	**1.74** / 6.99	1.08 / .54	1.61 / 2.69	1.28 / 1.08
	RF–SAA	5.84 / 26.88	**1.23** / 7.53	.76 / .54	1.14 / 3.23	.91 / .54
Size–Inv. 25	Linear	8.90 / 28.49	1.79 / 11.29	.93 / 2.15	**1.36** / 6.45	1.07 / 4.84
	Neural	7.52 / 21.51	**1.67** / 8.60	1.20 / 2.15	1.64 / 4.84	1.45 / 3.23
	RF–SAA	6.15 / 25.81	1.30 / 11.83	.60 / 4.30	**.88** / 8.60	.77 / 6.45
Industry 30	Linear	9.19 / 26.88	**2.15** / 9.68	1.08 / .54	1.81 / 2.69	1.29 / 1.08
	Neural	5.37 / 25.27	1.87 / 10.22	.94 / 2.15	**1.35** / 5.38	1.17 / 4.84
	RF–SAA	5.10 / 31.18	.76 / 8.06	.67 / 1.08	**1.02** / 3.23	.85 / 2.69

D.6 Sensitivity Analysis

Validation sample size and retraining. Figure 3 averages over the three retained classes and two DGPs, with 1,200 repetitions per application and validation size. As n_{val} increases from 300 to 1,200, the high-confidence calibrators become less conservative while retaining high test-pass rates. DKW passes every repetition at all five sizes; C has pass rates between 74% and 82%.

At $n_{\text{val}} = 1{,}200$, the ten complete-retraining repetitions per class and DGP yield 60 runs per application. DKW passes all 60 in both applications; each Gaussian procedure passes 60 newsvendor

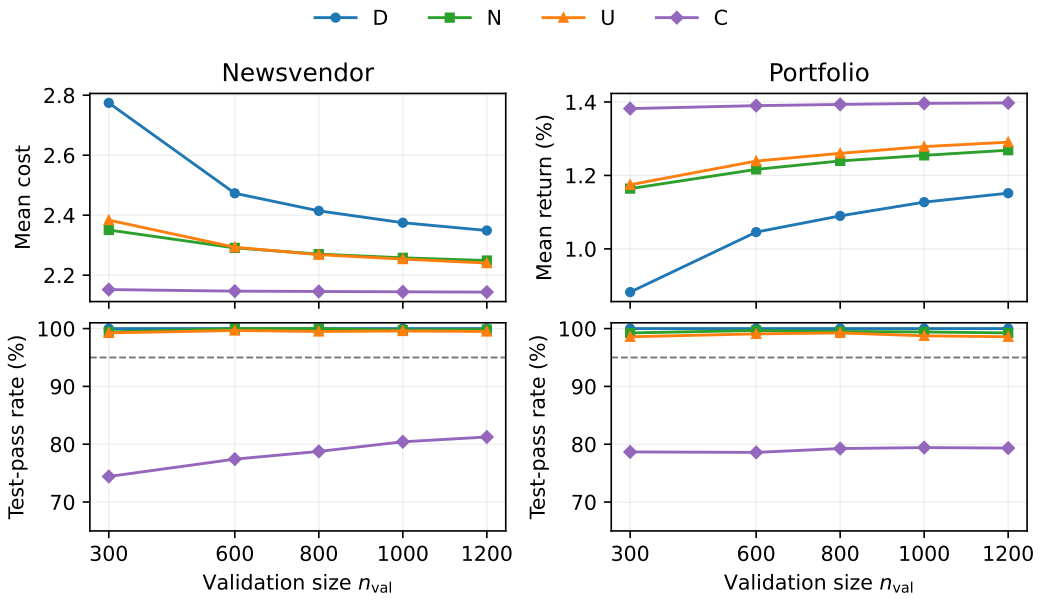


Figure 3: Validation-size sensitivity over linear, neural, and RF–SAA pretrained decision-rule classes and two DGPs. Objective is cost for newsvendor and return (%) for portfolio. The dashed line marks a 95% test-pass rate; each point pools 1,200 repetitions.

and 59 portfolio runs. Conformal calibration passes 49 and 47, respectively. The distinction between the procedures therefore also appears when training variability is included.

Pretraining objective. Table 5 compares chance-constrained and task-only pretraining for the linear and neural portfolio classes across four panels, giving eight panel–class combinations. Every uncalibrated rule violates the 10% target under both objectives. With DKW, all eight combinations meet the target, with mean return 1.02% after chance-constrained pretraining and .93% after task-only pretraining. Conformal calibration increases mean return to about 1.91% for both objectives and passes five and three combinations, respectively. In this ablation, the calibration choice accounts for more of the observed conservatism than the pretraining objective.

Table 5: Real world portfolio pretraining–calibration ablation at $\alpha = .10$, averaged over linear and neural rules and four panels. Return is annualized excess return, downside is the frequency of excess return below -1%, and cash is mean allocation; all use test data. Bold marks the highest return among procedures feasible in all eight combinations.

Pretraining	Calibration	Return (%)	Downside (%)	Cash (%)	Feasible cells
Chance-constrained	None	7.07	23.05	22.67	0/8
Chance-constrained	D	**1.02**	2.08	89.95	8/8
Chance-constrained	C	1.91	9.54	79.53	5/8
Task-only	None	9.12	27.42	4.74	0/8
Task-only	D	**.93**	2.49	89.90	8/8
Task-only	C	1.91	10.35	79.73	3/8

Changing α changes both pretraining and calibration. For linear and neural rules with DKW calibration, increasing α from .10 to .30 reduces mean cash from 89.95% to 59.24% and increases annualized excess return from 1.02% to 4.08%; downside frequency on test data rises from 2.08% to 18.21%. These are averages over the same eight panel–class combinations.

Calibration confidence. Figure 4 varies $1 - \beta$ while holding the pretrained rules, interpolation paths, validation size, and margin multiplier fixed. Across 70%, 80%, 90%, 95%, and 99% confidence, higher confidence makes each high-confidence calibrator more conservative and

generally raises its test-pass rate. The conformal calibrator appears as a horizontal reference because its criterion is invariant to β.

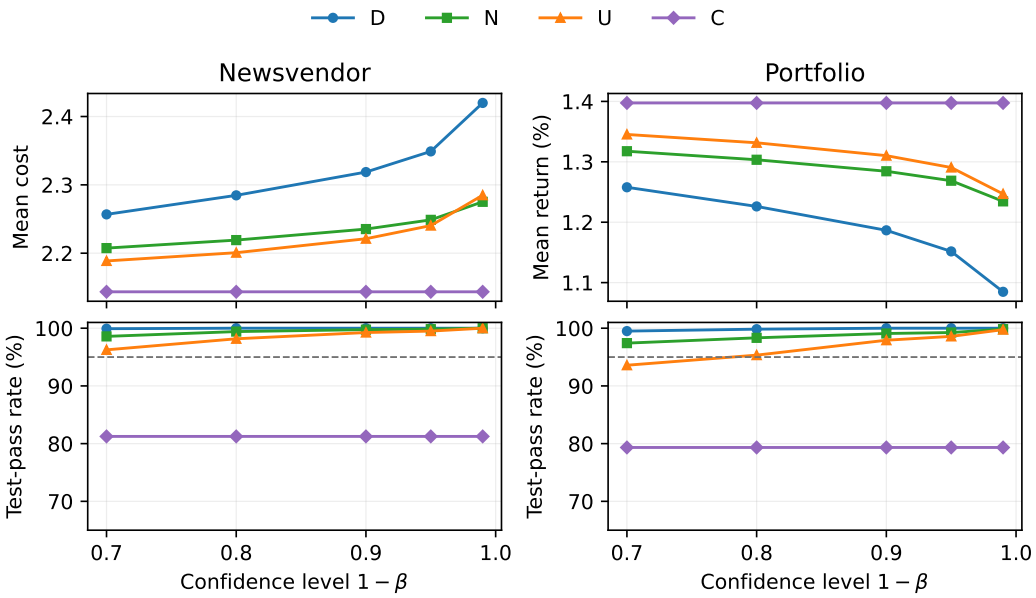


Figure 4: Synthetic confidence sensitivity at $n_{\text{val}} = 1{,}200$ and $n_{\text{te}} = 10{,}000$, pooled over three classes and two DGPs (1,200 repetitions per application, calibrator, and confidence level). Objective is cost for newsvendor and return (%) for portfolio. The dashed line marks a 95% test-pass rate.

The real data portfolio sensitivity shows a similar objective gain (Table 6). Linear and neural rules meet the 10% downside target on test data in all four panels under both confidence settings. For RF–SAA, 70% DKW and unnormalized Gaussian remain feasible in all panels; normalized Gaussian reaches 10.22% downside on Size–Investment 25.

Table 6: Real data portfolio confidence sensitivity, averaged over four panels. Return is annualized excess return, and downside is the test frequency of monthly excess return below -1%. Feas. counts panels meeting the 10% target. Bold marks the highest return among calibrators feasible in all four panels; both confidence levels are retrospective.

		95% confidence			70% confidence		
Rule class	Selector	Return (%)	Downside (%)	Feas.	Return (%)	Downside (%)	Feas.
Linear	D	.98	2.02	4/4	1.33	4.57	4/4
	N	**1.54**	5.91	4/4	**1.68**	7.53	4/4
	U	1.16	3.63	4/4	1.58	6.32	4/4
Neural	D	1.07	2.15	4/4	1.39	3.90	4/4
	N	**1.55**	4.84	4/4	**1.72**	6.32	4/4
	U	1.36	3.76	4/4	1.64	6.18	4/4
RF–SAA	D	.68	2.42	4/4	.92	4.57	4/4
	N	**1.03**	5.78	4/4	1.17	6.72	3/4
	U	.86	4.03	4/4	**1.13**	6.45	4/4

For the neural Bakery rule, lowering confidence from 95% to 70% increases DKW stockout from 10.26% to 12.41% and both Gaussian rules have 12.41% stockout at 95% confidence. We retain 95% as the default and report 70% as a sensitivity.

D.7 Context-dependent Safety Parameters

We examine whether the safety adjustment can vary with the context. For each validation sample, we form a standardized context vector $\phi(\boldsymbol{x})$ containing an intercept, the five context coordinates, their

squares, and their pairwise products (21 coefficients in total). We set $\lambda_\vartheta(\boldsymbol{x}) = \mathrm{sigmoid}(\vartheta^\top \phi(\boldsymbol{x}))$ and use the decision rule $\boldsymbol{w}_\vartheta(\boldsymbol{x}) = (1-\lambda_\vartheta(\boldsymbol{x}))\boldsymbol{w}^{\mathrm{pre}}(\boldsymbol{x}) + \lambda_\vartheta(\boldsymbol{x})\boldsymbol{w}_0(\boldsymbol{x})$. The parameter is initialized near the scalar D selection. Using the same 1,200 validation samples, we run 150 Adam steps with learning rate .05 to minimize

$$\widehat{C}_{\mathrm{val}}(\vartheta) + \rho\big[\widetilde{V}_\tau(\vartheta) + M^{\mathsf{D}} - \alpha\big]_+^2 + \gamma\|\vartheta_{-0}\|_2^2, \qquad \widetilde{V}_\tau(\vartheta) = \frac{1}{n_{\mathrm{val}}}\sum_{i=1}^{n_{\mathrm{val}}} \mathrm{sigmoid}\left(\frac{g(\boldsymbol{w}_\vartheta(\boldsymbol{x}_i^{\mathrm{val}}), \boldsymbol{z}_i^{\mathrm{val}})}{\tau}\right),$$

where $\widehat{C}_{\mathrm{val}}$ is mean validation cost, $M^{\mathsf{D}} = \sqrt{\log(1/\beta)/(2n_{\mathrm{val}})}$, $\alpha = .10$, and $\beta = .05$; τ is the smoothing bandwidth, and ρ, γ are penalty weights.

After each update, we evaluate the original violation indicator and objective on validation data. Here $\widehat{V}_{\mathrm{val}}(\vartheta)$ is the fraction of validation samples with $g(\boldsymbol{w}_\vartheta(\boldsymbol{x}_i^{\mathrm{val}}), \boldsymbol{z}_i^{\mathrm{val}}) > 0$. We retain the lowest-cost iterate satisfying $\widehat{V}_{\mathrm{val}}(\vartheta) + M^{\mathsf{D}} \le \alpha$ if it improves the scalar D rule's validation objective; otherwise we retain that scalar rule. The independent 10,000-observation test sample is used only for evaluation. The margin here matches the scalar D margin numerically.

Table 7: Context-dependent safety parameter versus scalar D at 95% confidence. Objective is mean test cost for newsvendor (lower is better) and mean test return (%) for portfolio (higher is better). Control and stress DGPs are pooled, with 400 paired repetitions per row. Violation is the mean test violation rate (%); the final column counts context-dependent test passes. Scalar D passes all 400 repetitions in every row.

		Objective		Violation (%)		
Application	Class	D	Context	D	Context	Passes / 400
Newsvendor	Linear	3.279	2.745	6.11	6.94	400
	Neural	1.688	1.612	5.58	7.26	399
	RF–SAA	2.080	1.737	5.91	6.96	395
Portfolio	Linear	1.206	1.339	6.41	7.18	400
	Neural	1.202	1.334	6.35	7.01	400
	RF–SAA	1.047	1.288	6.39	7.84	398

Table 7 uses one fixed pretrained rule per application–DGP–class combination and 200 paired validation/test repetitions for each DGP. Paired 95% intervals for the mean objective gain over scalar D exclude zero in all twelve underlying DGP-specific settings, conditional on those pretrained rules. The contextual rule also improves mean objective over N in ten of these settings, U in eight, and C in four; these comparisons describe different feasibility guarantees. In a separate check with five independently pretrained rules per setting and ten paired validation/test draws per rule, the mean gain over scalar D is positive in every setting. The contextual rule passes 599/600 tests in this check, versus 600/600 for scalar D. All pass counts use the finite-test violation threshold.

D.8 Global-VaR comparison

We consider a contextual portfolio problem with a single global cost threshold. For a return vector $\boldsymbol{z} \in \mathbb{R}^d$, let $c(\boldsymbol{w}, \boldsymbol{z}) := -\boldsymbol{z}^\top \boldsymbol{w}$. The decision rule chooses a long-only, fully invested portfolio and solves

$$\min_{\boldsymbol{w}(\cdot)\in\mathcal{H},\tau} \ \tau \quad \text{s.t.} \quad \mathbb{P}\{c(\boldsymbol{w}(\boldsymbol{x}), \boldsymbol{z}) \le \tau\} \ge 1-\alpha, \qquad \boldsymbol{w}(\boldsymbol{x}) \ge 0, \quad \mathbf{1}^\top \boldsymbol{w}(\boldsymbol{x}) = 1 \quad \text{a.s.}$$

For any fixed decision rule, the smallest feasible threshold τ is $\mathrm{VaR}_{1-\alpha}(c(\boldsymbol{w}(\boldsymbol{x}), \boldsymbol{z}))$, so this formulation minimizes global value-at-risk over the decision-rule class. We set $\alpha = .10$ and use the conditional Gaussian return model of Wang et al. (2023), with $d = 10$ assets and 20 context regimes sampled independently and uniformly. At $N = 1000$, the training data contain 240 fitting and 60 model-selection observations. Each method completes its training and model-selection procedure once. We then use 200 paired repetitions, each with 200 fresh validation samples and 500 test observations. All methods receive the same training data and validation/test draws in each repetition.

Our conformal and normalized Gaussian validation selectors C and N calibrate the global threshold for the same neural pretrained rule. Selector N uses $\beta = .05$. The portfolio network applies an

asset-scoring network with two 64-unit ReLU hidden layers and converts the scores to portfolio weights by softmax. Pretraining jointly learns this rule and a threshold through a smoothed chance constraint. Calibration keeps the pretrained rule fixed and selects τ from a solution path containing 100 training-loss quantiles and $+\infty$. Selector C chooses the first threshold satisfying its corrected empirical-risk criterion.

We compare with CPO (Patel et al., 2024), LRO (Wang et al., 2023), Hu-CRC (Hu et al., 2026), IDCC-E (Chenreddy et al., 2022), and conformal risk training (CRT) (Yeh et al., 2025). CPO learns a conditional generator, calibrates a conformal uncertainty region, and solves a worst-case epigraph problem. LRO learns a contextual ellipsoid through a robust solution map and an outer CVaR constraint; we retain its five-step training budget. Our global-τ Hu-CRC implementation learns an ellipsoid through smoothed robustness training and selects a radius using validation data. IDCC-E is our full-covariance ellipsoidal implementation of Integrated Deep Cluster then Classify: it learns contextual clusters and ellipsoids, calibrates regional coverage, and solves robust portfolios. CRT fine-tunes a task-trained portfolio network through conformal risk control and applies a fresh conformal correction to τ on each validation sample.

Table 8: Global-VaR comparison at $d = 10$ and $N = 1000$, averaged over 200 paired repetitions. $\bar{\tau}$ is the mean selected threshold; test VaR is the mean empirical 90th percentile of cost on test data. Lower is better for both. Passes count test samples with violation at most 10%. Time is a single measured wall time for training/model selection and all 200 validation/test pairs.

Method	$\bar{\tau}$	Test VaR	Violation (%)	Passes / 200	Time (s)
Ours, C	$\mathbf{-.654}$	$-.660$	9.687	125	4.6
Ours, N	$-.601$	$-.660$	5.753	191	16.1
CPO	.360	$-.551$	.003	200	127.4
LRO	$-.228$	$-.521$	1.029	200	29.6
Hu-CRC	$-.483$	$-.520$	7.967	189	60.3
IDCC-E	$-.053$	$-.710$	.004	200	34.0
Yeh CRT	$-.635$	$-.639$	9.936	109	24.3

Table 8 shows that C attains the lowest mean selected threshold, with 9.69% mean violation and 125/200 test passes. Selector N attains a lower mean selected threshold than CPO, LRO, Hu-CRC, and IDCC-E, with 5.75% mean violation and 191/200 test passes. The two selectors have identical test VaR because they use the same pretrained portfolio. IDCC-E has the lowest test VaR; its substantially higher selected threshold illustrates the conservativeness of its robust certificate. Both C and CRT control expected violation across validation samples. CRT has mean violation close to 10% and 109/200 test passes. For a continuous cost distribution, even the exact 90% VaR produces a test violation rate at most 10% in approximately 54% of 500-observation test samples. The pass count therefore describes a finite-test diagnostic for these different forms of risk control.

We measure complete pipeline times of 4.55 and 16.1 seconds for C and N, respectively. Complete calibration, including prediction, path construction, validation evaluation, and selection, takes .029 and 6.66 seconds across 200 repetitions. Each timed run reconstructs the same pretrained rule; the measured pretraining times are 1.71 and 5.65 seconds. These single-run measurements illustrate the low incremental cost of calibration.

E Fractional Knapsack and Alloy Allocation

We also evaluate two contextual stochastic linear optimization applications. For fractional knapsack, let $\boldsymbol{z} = (\boldsymbol{r}, \boldsymbol{a})$ denote the uncertain reward and item-weight vectors. The decision rule allocates at most one unit across five items and solves

$$\max_{\boldsymbol{w}(\cdot)\in\mathcal{H}} \quad \mathbb{E}\left[\boldsymbol{r}^\top \boldsymbol{w}(\boldsymbol{x})\right]$$
$$\text{s.t.} \quad \mathbb{P}\left\{\boldsymbol{a}^\top \boldsymbol{w}(\boldsymbol{x}) \le 1\right\} \ge .90, \qquad \boldsymbol{w}(\boldsymbol{x}) \ge 0, \quad \mathbf{1}^\top \boldsymbol{w}(\boldsymbol{x}) \le 1 \quad \text{a.s.}$$

Thus $c(\boldsymbol{w}, \boldsymbol{z}) = -\boldsymbol{r}^\top \boldsymbol{w}$ and $g(\boldsymbol{w}, \boldsymbol{z}) = \boldsymbol{a}^\top \boldsymbol{w} - 1$. The reference rule is

$$\boldsymbol{w}_0(\boldsymbol{x}) = \mathbf{0},$$

which has zero resource use and belongs to the decision set.

For alloy allocation, let $\boldsymbol{z} = (\boldsymbol{q}, \boldsymbol{A})$ denote procurement costs and the uncertain contents of two metals supplied by four suppliers, with $\boldsymbol{a}_j^\top$ denoting row j of $\boldsymbol{A}$. The problem is

$$\begin{aligned} \min_{\boldsymbol{w}(\cdot)\in\mathcal{H}} \quad & \mathbb{E}\left[\boldsymbol{q}^\top \boldsymbol{w}(\boldsymbol{x})\right] \\ \text{s.t.} \quad & \mathbb{P}\left\{\boldsymbol{a}_j^\top \boldsymbol{w}(\boldsymbol{x}) \geq 1\right\} \geq .90, \forall j = 1, 2, \qquad \boldsymbol{w}(\boldsymbol{x}) \geq 0, \quad \mathbf{1}^\top \boldsymbol{w}(\boldsymbol{x}) \leq 3 \quad \text{a.s.} \end{aligned}$$

Here $c(\boldsymbol{w}, \boldsymbol{z}) = \boldsymbol{q}^\top \boldsymbol{w}$ and $g(\boldsymbol{w}, \boldsymbol{z}) = \max_{j=1,2}\{1 - \boldsymbol{a}_j^\top \boldsymbol{w}\}$. The reference rule is

$$\boldsymbol{w}_0(\boldsymbol{x}) = (0, 0, 0, 1.5).$$

Since the fourth supplier contains at least .70 units of either metal per unit, purchasing 1.5 units from supplier four provides at least $1.5 \times .70 = 1.05$ units of each required metal, while respecting the total procurement limit of 3. Therefore, this rule satisfies both requirements for every observation.

Each application uses five bounded context coordinates. The control DGP has context-dependent coefficient means and constant noise amplitude; the stress DGP adds an $x_1 x_2$ interaction and scales the main coefficient-noise terms by 1.5. We use 3,600 pretraining observations for each linear and neural rule. We use 101 interpolation-path candidates, $\alpha = .10$, and five paired repetitions with 1,200 validation and 2,000 test observations each. The high-confidence calibrators use 95% confidence. For comparison, we implement the untruncated SPO-RC+ method of Im et al. (2025) with neural coefficient and cost predictors, an ℓ_1 split-conformal uncertainty set, and a robust linear program. Its uncertainty set is calibrated on the first validation draw and then held fixed; our calibrators use each repetition's validation draw.

Table 9: Additional synthetic test results: mean objective / mean violation rate (%). Higher reward is better for knapsack; lower cost is better for alloy. Lin. denotes a linear pretrained rule. Each SPO-RC+ row reports one separate robust prediction-and-optimization method.

Application	DGP	Rule/method	Raw	C	D	N	U
Knapsack	Control	Lin.	1.667/37.6	1.550/7.2	1.530/4.5	1.543/6.2	1.537/5.3
		Neural	1.686/39.2	1.568/8.0	1.551/5.5	1.555/5.9	1.551/5.5
		SPO-RC+			1.472/1.1		
	Stress	Lin.	1.671/40.0	1.517/9.9	1.490/6.6	1.500/7.7	1.494/6.9
		Neural	1.693/41.2	1.524/9.7	1.496/6.7	1.503/7.5	1.500/7.0
		SPO-RC+			1.389/1.3		
Alloy	Control	Lin.	1.251/17.1	1.356/9.8	1.412/6.2	1.386/8.0	1.388/7.8
		Neural	1.188/19.5	1.342/9.9	1.399/6.6	1.375/7.8	1.379/7.6
		SPO-RC+			1.380/1.4		
	Stress	Lin.	1.294/21.9	1.496/8.9	1.551/5.9	1.531/6.9	1.538/6.6
		Neural	1.228/22.6	1.464/9.6	1.531/5.9	1.500/7.5	1.510/6.9
		SPO-RC+			1.500/0.9		

The pretrained rules have mean violation rates of 17.1–41.2%, whereas all four calibrators bring the mean below 10% in these settings. For knapsack, the calibrated rules earn more reward than SPO-RC+ at higher mean violation rates. For alloy, SPO-RC+ has lower cost than DKW and substantially lower mean violation; conformal calibration achieves lower cost than SPO-RC+ with mean violation near the target. The results show the objective–violation tradeoff under capacity and joint composition constraints.