A fundamental practical barrier to deploying personalized data-driven decision-making systems is the inability to credibly and safely evaluate performance on previously-collected real-world data without live deployment. For example in healthcare, an important goal is to learn from electronic health records, which are widely available unlike costly randomized-controlled-trial data, in order to *estimate and optimize* the value of a personalized treatment policy. Unlike classical operational settings where decisions are made on objects with known dynamics, in human-centered systems such as e-commerce and healthcare, the personalized causal **effects** of actions are unknown and need to be statistically estimated. Although there has been great progress at the interface of causal inference and machine learning, incorporating these techniques requires certain assumptions that are widely known to *not* hold in practice. My work on **credible methodology** bridges the gap between theory and practice by empowering analysts and practitioners to optimize robust decisions or report bounds on inferential parameters under realistic and practitioner-tunable violations of assumptions. Credible performance assessment is also key for emerging applications in algorithmic fairness. Recent controversies about the use of machine learning for risk assessment in criminal justice, lending in financial services, and the provision of social services, emphasize the importance of stronger individual- or subgroup-level performance guarantees, although these same application areas pose practical barriers to measuring disparities to ensure **equitable** performance.

More broadly, my research interests are at the interface of statistical machine learning and operations research. In my dissertation work, I also draw on and contribute to causal inference. My ultimate goals are to develop reliable, **effective, credible, and equitable personalized data-driven decision-making** so that machine learning can be deployed in important applications for beneficial impacts on firms, individuals, and society. To do so, I (1) identify common statistical structure and desiderata from challenging and motivating applications, (2) establish *the theoretical foundations and performance guarantees for novel methodology*, (3) develop *practical algorithmic frameworks*, and (4) *articulate managerial insights* to illustrate the relevance of my "effective and credible" lens for researchers and practitioners working in these motivating settings. Methodologically, I often design robust and credible estimators and algorithms under a unifying viewpoint of optimization under ambiguity and prove statistical convergence guarantees by leveraging optimization structure.

Below, I describe my work in more detail along the parallel methodological threads of robust personalization from observational data and credible fairness and impact assessment.

Robust personalization from observational data ([2, 8, 6]). One line of work focuses on learning to improve personalized decisions from observational data in the presence of unobserved confounders. In "Minimax-Optimal Policy Learning Under Unobserved Confounding" (accepted at *Management Science*) [2], I provide a practical algorithmic framework and statistical theoretical guarantees.

Motivation and problem setting. Statistical confounding occurs when historical decisions depend on confounders which also affect the outcome. As a result, naive estimation of causal effects on a historical dataset is conditional on the historical selection pattern, and causal effect estimation is biased. Methods from causal inference adjust for this selection bias due to *observed confounders* by assuming *selection on observables* or *unconfoundedness*: that outcomes are conditionally independent of treatment assignment upon adjusting for observed covariates. Unconfoundedness is standard to justify most causal inference adjustment methods, and all recent progress in causal inference and machine learning assumed it.

However, unconfoundedness is often violated in practice by design of the operational environment. For example, in healthcare, physicians made their decisions based on additional information that is not recorded, such as intuition about patient presentation. This is true more broadly of decisions made by experts, or decisions that reflect prior optimization. In our paper, we illustrate the relevance of our approach with an extensive case study built on the Women's Health Initiative parallel observational study, which was subject to unobserved confounding from self-selection into elective treatment, and clinical trial, which would otherwise be the gold-standard for causal inference. The observational study originally suggested treatment might be beneficial for chronic disease prevention; while the clinical trial had to be halted early due to the increased incidence of deaths. A common explanation was that unobserved confounding led to overall healthier women being enrolled in the treatment in the observational study.

Contributions and challenges. In a setting with unobserved confounders, the data analyst only observes i.i.d draws from the joint distribution of covariate, treatment, and outcome data, (X, T, Y), for observed treatment tuples, but the underlying data, (U, X, T, Y), was generated with an additional unobserved confounder, U, which influences treatment and outcome. *If* we had access to the full underlying data, it would be sufficient to learn $\mathbb{P}(T \mid X, U)$, the true probability of treatment assignment (propensity score) in order to statistically adjust for selection bias via the likelihood ratio, e.g. by importance sampling. A data-driven approach recognizes that we may estimate $\mathbb{P}(T \mid X)$, which adjusts for some but not all confounding. I develop a nonparametric minimax approach to learn a machine learning policy with the best worst-case guarantee over an ambiguity set on the inverse propensity weights. The ambiguity set $\mathcal{U}$ restricts deviations of the true underlying inverse propensity weight $W^* = {}^1\!/_{\mathbb{P}(T|X,U)}$ from what we *can* estimate from observed data, $W = {}^1\!/_{\mathbb{P}(T|X)}$. These restrictions are parametrized according to the marginal sensitivity model from causal inference which translates to an interval uncertainty set of W^* with respect to W. The overall "size" of the set is parameterized by a scalar parameter, Γ: an analyst presumably specifies bounds on plausible ranges on the extent of residual unobserved confounding.

A personalized decision policy $\pi(X)$ maps covariates to the probability of an action. For the purposes of interpretability and generalization, we optimize personalized decision policies over a parametrized function class such as the family of decision rules based on a linear index, $\pi(X) = \mathbb{I}[\beta^\top X > 0]$, or logistic assignment. When the treatment contrast is coded $T \in \{-1, +1\}$, the standard *policy learning* problem solves $\max_\pi \mathbb{E}[YT\pi W]$, e.g. by analogy to weighted classification. We solve a *minimax* problem, optimizing over the robust worst-case regret relative to a baseline policy π_0 over the ambiguity set, to find a confounding-robust policy:

$$\min_{\pi} \max_{W \in \mathcal{U}} \frac{\mathbb{E}[YT(\pi - \pi_0)W]}{\mathbb{E}[W]}$$

The baseline policy is simple, such as all-treat or all-control, and improves our safety guarantees. We leverage the special linear-fractional optimization structure to develop a computationally efficient algorithmic framework for learning a robust policy. We take a robust gradient descent approach: for parametrized policies, we solve the inner optimization to full optimality via a ternary search procedure and evaluate gradients at the worst-case solution.

Our setting introduces technical challenges. In contrast to robust decision-making, unobserved confounders result in ambiguity about *both* the underlying generating distribution and the realizations that appear in our dataset. Obtaining statistical convergence results of our robust machine-learning decision rule requires problem-specialized structural characterization of the optimal subproblem solution or otherwise proving stability of the optimization problem. Estimated nuisance parameters introduce perturbations to *left-hand side* constraints.

Our theoretical guarantees recover the same rate of convergence of the previous non-robust approach with our novel safety/improvement guarantees: with high probability, the sample-minimax optimal policy deviates from the population-minimax optimal policy by a vanishing $O_p(n^{-\frac{1}{2}})$ term, where n is the number of confounded samples. Therefore, our results show that researchers, analysts, and practitioners can develop confounding-robust personalized decision policies at no great cost *computationally* nor *statistically.*

Follow-up work: infinite-horizon reinforcement learning. While minimax-policy learning studies the single-stage decision case, in healthcare there is increasing interest in managing chronic health conditions over time, such as insulin dosing in diabetic control, which requires offline reinforcement learning and off-policy evaluation in the sequential setting. However, data collected from electronic medical records in this setting is often observational and hence subject to problems of unobserved confounders. We study robust policy evaluation in [6] in the infinite horizon reinforcement learning setting. Instead of a direct generalization of the robust estimator approach used in [2], which would become "exponentially robust" in the time horizon, we build on a recently proposed estimating equation for the stationary distribution density ratio. Assuming i.i.d unobserved confounders, we optimize robust bounds on an estimating equation for the density ratio of state occupancy measures.

Credible impact evaluation in algorithmic fairness ([7, 3, 4, 5]). In another line of work, I develop robust, credible methodology for disparity assessment for the growing area of **algorithmic fairness** and impact assessment in consequential settings more broadly. A key question in fair data-driven decision-making is to trade-off and assess disparities in the performance of decision rules, in particular along the dimensions of a "protected attribute" upon which discrimination is prohibited, such as race or sex. I develop tools and crisp managerial insights to align the performance benchmarks, by which we measure machine learning progress and ultimately justify the deployment of algorithms, with the real-world actual operating conditions and impacts of algorithms in consequential settings.

In "Assessing Disparities with Unobserved Protected Class" (accepted at Management Science) [7], we develop methodology for assessing bounds on disparities in settings where the protected attribute of interest is actually not recorded in the dataset of decision recommendations, decisions, and covariates $(Y, \hat{Y}, Z)$, but auxiliary data on the protected attribute and covariates (A, Z) is available. This is a common setting in practice, such as in financial services, where even regulators rely on *proxy methods* that estimate $\mathbb{P}(A \mid Z)$ but require untestable assumptions. Thus, disparities crucially rely on the *unobserved* conditional joint distribution although we only have access to the *observed* marginal distributions. We take a partial identification approach and estimate the sets that correspond to all possible disparities that are consistent with the observed data. Disparities are functionals that can be represented as marginal averages of the *unknown* $\mathbb{P}(Y, \hat{Y} \mid A, Z)$ with respect to the *estimable* measure $\mathbb{P}(A \mid Z)$. While $\mathbb{P}(Y, \hat{Y} \mid A, Z)$ is an unknown quantity, it satisfies properties of all probability distributions, such as boundedness and the law of total probability, and its marginalization over A, $\mathbb{P}(Y, \hat{Y} \mid Z)$, *is* estimable. We therefore optimize over the *ambiguous* probability distribution $\mathbb{P}(Y, \hat{Y} \mid A, Z)$ *subject to these constraints which arise from the underlying probability distribution structure.* A unifying optimization framework allows for adding additional structural assumptions and developing a general-purpose algorithm for recovering the set of disparities which are supported by the observed data in increasingly complex settings. We use a support function representation of the partial identification set in order to computationally represent the convex hull of the disparity set. Our work provides tools to assess the basic limits of what can be concluded from the data; and can be used to either verify that assessed disparities remain (regardless of the assumptions induced by any further estimation procedure), or that investment in further refinement of auxiliary data is necessary.

Other fairness work In [3], we show that in some consequential settings where fairness in machine learning has been of concern – lending, criminal justice, and social services – all machine learning methods are necessarily trained on data censored by previous decisions. A key insight is that not *only* does the standard machine learning assumption of representative data fail, its failure is an *inevitable* consequence of practical settings, and attempts to mitigate disparities could *inadvertently continue to disadvantage* those who were harmed by non-representation under the guise of fairness. Other work proposes and studies principled evaluation metrics that better align with stated stakeholder desiderata as in the xAUC metric [5] to deliver crisp insights to inform practitioners in these areas, or develops robust evaluation for more challenging settings such as causal interventions [4]. These insights are directly targeted towards researchers and practitioners; the xAUC metric was recently used in a large-scale study of fairness in clinical risk scores [9].

I am particularly interested in studying practical opportunities to improve reliability and outcomes from machine learning methods without relying on approaches that penalize model performance. During my research internship with Miro Dudik and Jenn Wortman Vaughan, motivated by a study of what business practitioners do in practice to improve model disparities [1], I developed theory for and empirically studied when the common practical intervention of collecting more data may improve fairness properties of regression models by developing approximations to group-conditional finite-sample error which an analyst may use to guide the collection of additional data to improve disparities most efficiently.

Ongoing and future research: In ongoing research, I am continuing work on off-policy learning in the sequential setting. The key insight is that commonly in operations research, a subset of the state variable undergoes known dynamics, conditional on a (possibly contextual or covariate-conditional) effect. Therefore the global state variable (e.g. inventory or resource consumption) introduces sequential dependence although the contextual personalized effect is time-invariant. This setting bridges the single-stage and dynamic settings and highlights the interface of statistical machine learning and operations research: the structure of OR problems provides opportunities to leverage optimization structure in order to improve upon off-policy learning algorithms and results that rely solely on statistical estimation.

In recently submitted work, I have developed a framework to study questions of algorithmic fairness in relation to personalized pricing. The setting of personalized pricing can illustrate the benefits of greater personalization for expanding access directing resources to those who “value” them the most, but many real-world challenges, including the joint covariance structure of groups and covariates or finite-sample uncertainty can lead to inequities in resulting allocations. An important generalization for further work is studying the question of fairness in decision utility of the *contextual* version of classical operations research stochastic optimization problems which newly incorporate machine-learned predictors.

For future research, I plan on continuing my research on reliable, robust, and trustworthy machine learning more broadly, with attention to the interface of optimization and estimation for making better decisions from data. For example, jointly considering the robust bias guarantees of the sensitivity parameter alongside the estimation variance can lead to end-to-end guarantees for accounting for both ambiguity and finite-sample uncertainty.

Other research directions continue to develop the interface between methodological theory and practical settings and particularly benefit from the operations perspective. One research direction recognizes that, as my work in fairness in pricing argues, in broader application settings, short-term fairness or welfare considerations may be of interest because of longer-term considerations such as consumer retention or future value. It would be important to study this empirically as well as develop algorithmic personalization schemes that better manage long-term system performance. Another research direction recognizes that while my previous work considered the case of *unobserved confounders*, the opposite challenge for algorithmic decision support is understanding how “humans-in-the-loop” *should* use additional information available to decision-makers at the time of deployment. This is a fundamental problem that limits the practical impact of prescriptions from algorithms in practice, is not considered by conventional theory, and relates to statistical bias-variance tradeoffs.

Over the long term, I plan on continuing to tailor statistical machine learning methodological development to the challenges of emerging application areas.

References

[1] K. Holstein, J. Wortman Vaughan, H. Daumé III, M. Dudik, and H. Wallach. Improving fairness in machine learning systems: What do industry practitioners need? In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–16, 2019.

[2] N. Kallus and A. Zhou. Minimax-optimal policy learning under unobserved confounding. *Management Science (Accepted), supersedes Neurips 2018 version.*

[3] N. Kallus and A. Zhou. Residual unfairness in fair machine learning from prejudiced data. *ICML*, pages 2439–2448, 2018.

[4] N. Kallus and A. Zhou. Assessing disparate impact of personalized interventions: Identifiability and bounds. *Neurips*, pages 3426–3437, 2019.

[5] N. Kallus and A. Zhou. The fairness of risk scores beyond classification: Bipartite ranking and the xauc metric. *Neurips*, pages 3438–3448, 2019.

[6] N. Kallus and A. Zhou. Confounding-robust policy evaluation in infinite-horizon reinforcement learning. *Neurips*, 2020.

[7] N. Kallus, X. Mao, and A. Zhou. Assessing algorithmic fairness with unobserved protected class using data combination. *Management Science (Accepted). A preliminary version appeared at FaCCT 2020 as an extended abstract.*

[8] N. Kallus, X. Mao, and A. Zhou. Interval estimation of individual-level causal effects under unobserved confounding. *AISTATS*, pages 2281–2290, 2019.

[9] S. R. Pfohl, A. Foryciarz, and N. H. Shah. An empirical characterization of fair machine learning for clinical risk prediction. *arXiv preprint arXiv:2007.10306*, 2020.