Causal Strategic Linear Regression

Yonadav Shavit [1] **Benjamin L. Edelman** [1] **Brian Axelrod** [2]

Abstract

In many predictive decision-making scenarios, such as credit scoring and academic testing, a decision-maker must construct a model that accounts for agents' incentives to "game" by changing their features to receive better decisions. Whereas the strategic classification literature has previously assumed that agents' outcomes are not causally dependent on their features (and thus strategic behavior is a form of lying), we join concurrent work in modeling agents' outcomes as a function of their changeable attributes. Our work introduces the realizable linear regression setting, and is the first to incorporate a crucial phenomenon: when agents act to change observable features, they may as a side effect perturb hidden features that causally affect their true outcomes. As our main contribution, we provide the efficient algorithms for optimizing three distinct decision-making objectives: accurately predicting agents' post-gaming outcomes (*prediction risk minimization*), incentivizing agents to improve these outcomes (*agent outcome maximization*), and estimating the coefficients of the true underlying model (*parameter estimation*). Our algorithms circumvent the hardness result of Miller et al. (2020) by allowing the decision maker to test a sequence of decision rules and observe agents' responses, in effect performing causal interventions by varying the chosen rule.

1. Introduction

As individuals, we want algorithmic transparency in decisions that affect us. Transparency lets us audit models for fairness and correctness, and allows us to understand what changes we can make to receive a different decision. Why, then, are some models kept hidden from the view of those subject to their decisions?

Beyond setting-specific concerns like intellectual property theft or training-data extraction, the canonical answer is that transparency would allow strategic individual *agents* to "game" the model. These individual agents will act to change their features to receive a better *decision*. An accuracy-minded decision-maker, meanwhile, chooses a decision rule based on its predictiveness of individuals' true future *outcomes*. Strategic agents, the conventional wisdom goes, make superficial changes to their features that will not affect their true outcomes, reducing the decision rule's accuracy and harming the decision-maker. The field of strategic classification (Hardt et al., 2016) has until recently sought to design algorithms that are robust to such superficial changes. At their core, these algorithms treat transparency as a reluctant concession and propose ways for decision-makers to get by nonetheless.

But what if decision-makers could *benefit* from transparency? What if in some settings, gaming could help accomplish the decision-makers' goals, by causing agents to truly change their outcomes without loss of predictive accuracy?

Consider the case of car insurance companies, who wish to choose a pricing decision rule that charges a customer in line with that customer's expected cost of accidents. Insurers will often charge lower prices to drivers who have completed a "driver's ed" course which teaches comprehensive driving skills. In response, drivers often complete such courses to reduce their insurance costs. One view may be that only *ex ante* responsible drivers seek out such courses, and that were an unsafe driver to complete such a course it would not affect their expected cost of car accidents.

But another interpretation is that drivers in these courses learn safer driving practices, and truly become safer drivers *because* they took this course. In this case, a car insurer's decision rule *remains predictive* of accident probability when agents strategize, while also incentivizing the population of drivers to act in a way that truly makes them safer, allowing the insurer to reimburse fewer accidents.

This same dynamic appears in many decision settings where the decision-maker has a meaningful stake in the true future

[1]Harvard School of Engineering and Applied Sciences, Cambridge, MA, USA [2]Stanford Computer Science Department, Palo Alto, CA, USA. Correspondence to: Yonadav Shavit <yonadav@g.harvard.edu>.

Proceedings of the 37^{th} International Conference on Machine Learning, Vienna, Austria, PMLR 119, 2020.

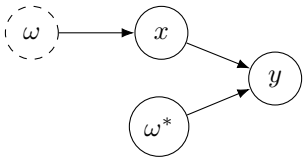

Figure 1. A causal graph illustrating that by intervening on the decision rule ω, a decision-maker can incentivize a change in x, enabling them to learn about how the agent outcome y is caused. We omit details of our setting for simplicity.

outcomes of its subject population, including credit scoring, academic testing, hiring, and online recommendation systems. In such scenarios, given the right decision rule, decision-makers can gain from transparency.

But how can we find such a decision rule that maximizes agents' outcomes if we do not know the effects of agents' actions? In recent work, Miller et al. (2020) argue that finding such "agent outcome"-maximizing decision rules requires solving a non-trivial causal inference problem. As we illustrate in Figure 1, the decision rule affects the agents' features, which causally affect agents' outcomes, and recovering these relationships from observational data is hard. We will refer to this setting as "causal strategic learning", after the causal relationship of decision rule on agent outcome.

The core insight of our work is that while we may not know how agents will respond to a decision rule, they will naturally respond to any rule we pick. Thus, as we test different decision rules and observe strategic agents' responses and true outcomes, we can improve our model over time. In the language of causality, by choosing a decision rule we are effectively launching an intervention that allows us to infer properties of the causal graph, circumventing the hardness result of (Miller et al., 2020).

In this work, we introduce the causal strategic linear regression setting in the realizable case and with norm-squared agent action costs. We propose algorithms for efficiently optimizing three possible objectives that decision-makers may maximize by leveraging strategic agents' gaming. *Agent outcome maximization* requires choosing a decision rule that will result in the highest expected outcome of agents who game that rule. *Prediction risk minimization* requires choosing a decision rule that accurately predicts agents' outcomes, even under agents' gaming in response to that decision rule. *Parameter estimation* involves accurately estimating the parameters of the true causal outcome-generating linear model. We show that these may be mutually non-satisfiable, and our algorithms maximize each objective independently (and jointly when possible).

We show that omitting unobserved outcome-affecting features from the decision rule has major consequences for causal strategic learning. Omitted variable bias in classic linear regression causes correlated-but-not-causal visible features, which are correlated with hidden causal features, to be rewarded in the learned predictor. In the strategic case, this backfires, as an agent's action may change a visible feature without changing the hidden feature in a way that breaks this correlation, undermining the naïvely-trained predictor. All of our methods are designed to succeed even when actions break the relationships between visible proxies and hidden causal features.

As much of the prior literature has focused on the case of binary classification, it's worth meditating on why we focus on regression. Many decisions, such as loan terms or insurance premiums, are not binary "accept/reject"s but rather lie somewhere on a continuum based on a prediction of a real-valued outcome. Furthermore, many ranking decisions, like which 10 items to recommend in response to a search query, can instead be viewed as real-valued predictions that are post-processed into an ordering.

1.1. Summary of Results

In Section 2, we introduce a setting for studying the performance of linear models that make continuous decisions about strategic agents. Our methodology incorporates the realities that agents' actions causally affect their eventual outcomes, that a decision-maker can only observe a subset of agents' features, and that agents' actions are constrained to a subspace of the feature space. We assume no prior knowledge of the agent feature distribution, or of the actions available to strategic agents, and require no knowledge of the true outcome function beyond that it is itself a noisy linear function of the features.

In Section 3, we propose an algorithm for efficiently learning a decision rule that maximizes agents' expected future outcomes. This method applies even when the decision-maker never observes the agents' available actions, so long as the decision-maker is willing to deploy a series of suboptimal decision rules.

In Section 4, we observe that under certain checkable conditions the prediction risk objective can be minimized using a convex optimization. We also provide a useful decomposition of prediction risk, and suggest how prediction risk and agent outcomes may be jointly optimized.

In Section 5, we show that in the case where all causally-outcome-affecting features are visible to the decision-maker, one can substantially improve the estimate of the true model parameters governing the outcome. At a high-level, this is because by incentivizing agents to change their features in certain directions, we are able to increase the variance along dimensions of the feature space that had little variance before gaming. For example, if two features were perfectly correlated in the initial agent feature distribution, incentiviz-

ing agents to increase only one of these features will allow us to disambiguate between the causal effect of each on the outcome.

1.2. Related Work

This paper is closely related to several recent and concurrent papers that study different aspects of causal strategic learning. Most of these works focus on one of our three objectives:

Agent outcomes. Our setting is partially inspired by Kleinberg & Raghavan (2019). In their setting, as in ours, an agent chooses an action vector in order to maximize the score they receive from a decision-maker. The action vector is mapped to a feature vector by an *effort conversion matrix*, and the decision-maker publishes a mechanism that maps the feature vector to a score. However, their decision-maker does not face a learning problem: the effort conversion matrix is given as input, agents do not have differing initial feature vectors, and there is no outcome variable. Moreover, there are no hidden features. In a variation on the agent outcomes objective, their decision-maker's goal is to incentivize agents to take a particular action vector. Their main result is that whenever a monotone mechanism can incentivize a given action vector, a linear mechanism suffices. Alon et al. (2020) analyze a multi-agent extension of this model.

In another closely related work, Miller et al. (2020) bring a causal perspective (Pearl, 2000; Peters et al., 2017) to the strategic classification literature. Whereas prior strategic classification works mostly assumed agents' actions have no effect on the outcome variable and are thus pure *gaming*, this paper points out that in many real-life strategic classification situations, the outcome variable is a descendant of some features in the causal graph, and thus actions may lead to genuine *improvement* in agent outcomes. Their main result is a reduction from the problem of orienting the edges of a causal graph to the problem of finding a decision rule that incentivizes net improvement. Since orienting a causal graph is a notoriously difficult causal inference problem given only observational data, they argue that this provides evidence that incentivizing improvement is hard. In this paper we we point out that improving agent outcomes may not be so difficult after all because the decision-maker does not need to rely only on observational data—they can perform causal interventions through the decision rule.

Haghtalab et al. (2020) study the agent outcomes objective in a linear setting that is similar to ours. A significant difference is that, while agents do have hidden features, they are never incentivized to change their hidden features because there is no effort conversion matrix. This, combined with the use of a Euclidean norm action cost (we, in contrast, use a Euclidean squared norm cost function), makes finding the optimal linear regression parameters trivial. Hence, they mainly focus on approximation algorithms for finding an optimal linear *classifier*.

Tabibian et al. (2020) consider a variant of the agent outcomes objective in a classification setting: the outcome is only "realized" if the agent receives a positive classification, and the decision-maker pays a cost for each positive classification it metes out. The decision-maker knows the dependence of the outcome variable on agent features a priori, so there is no learning.

Prediction risk. Perdomo et al. (2020) define *performative prediction* as any supervised learning scenario in which the model's predictions cause a change in the distribution of the target variable. This includes causal strategic learning as a special case. They analyze the dynamics of *repeated retraining*—repeatedly gathering data and performing empirical risk minimization—on the prediction risk. They prove that under certain smoothness and strong convexity assumptions, repeated retraining (or repeated gradient descent) converges at a linear rate to a near-optimal model.

Liu et al. (2020) introduce a setting where each agent responds to a classifier by intervening directly on the outcome variable, which then affects the feature vector in a manner depending on the agent's population subgroup membership.

Parameter estimation. Bechavod et al. (2020) study the effectiveness of repeated retraining at optimizing the parameter estimation objective in a linear setting. Like us, they argue that the decision-maker's control over the decision rule can be conducive to causal discovery. Specifically, they show that if the decision-maker repeatedly runs least squares regression (with a certain tie-breaking rule in the rank-deficient case) on batches of fresh data, the true parameters will eventually be recovered. Their setting is similar to ours but does not include an effort conversion matrix (nor hidden features, which we also omit from our parameter estimation section).

Non-causal strategic classification. The primary goal of the decision-maker in much of the classic strategic classification literature is robustness to gaming; the target measure is typically prediction risk. Our use of a Euclidean squared norm cost function is shared by the first paper in a strategic classfication setting (Brückner & Scheffer, 2011). Other works use a variety of different cost functions, such as the *separable* cost functions of Hardt et al. (2016). The online setting was introduced by Dong et al. (2018) and has also been studied by Chen et al. (2020), both with the goal of

minimizing "Stackelberg regret".[1] A few papers (Milli et al., 2019; Hu et al., 2019) show that accuracy for the decision-maker can come at the expense of increased agent costs and inequities. Braverman & Garg (2020) argue that random classification rules can be better for the decision-maker than deterministic rules.

Economics. Related problems have long been studied in information economics, specifically in the area of contract theory (Salanié, 2005; Laffont & Martimort, 2002). In *principal-agent problems* (Hölmstrom, 1979; Grossman & Hart, 1983; Holmstrom & Milgrom, 1991; Ederer et al., 2018), also known as *moral hazard* or *hidden action* problems, a decision-maker (called the *principal*) faces a challenge very similar to the agent outcomes objective. Notable differences include that the decision-maker can only observe the outcome variable, and the decision-maker must pay the agent. In a setting reminiscent of strategic classification, (Frankel & Kartik, 2020) prove that the fixed points of retraining can be improved in terms of accuracy if the decision-maker can commit to underutilizing the available information. Ball (2020) introduces a three-party model in which an intermediary scores the agent and a decision-maker makes a decision based on the score.

Ethical dimensions of strategizing. Ustun et al. (2019) and Venkatasubramanian & Alfano (forthcoming) argue that it is normatively good for individuals subject to models to have "recourse": the ability to induce the model to give a desired prediction by changing mutable features. Ziewitz (2019) discusses the shifting boundaries between morally "good" and "bad" strategizing in the context of search engine optimization.

Other strategic linear regression settings. A distinct literature on strategic variants of linear regression (Perote & Perote-Peña, 2004; Dekel et al., 2010; Chen et al., 2018; Ioannidis & Loiseau, 2013; Cummings et al., 2015) studies settings in which agents can misreport their y values to maximize their privacy or the model's prediction on their data point.

2. Problem Setting

Our setting is defined by the interplay between two parties: *agents*, who receive decisions based on their features, and a *decision-maker*, who chooses the decision rule.[2] We visualize our setting in Figure 2.

[1] See Bambauer & Zarsky (2018) for a discussion of online strategic classification from a legal perspective.

[2] In the strategic classification literature, these are occasionally referred to as the "Jury" and "Contestant".

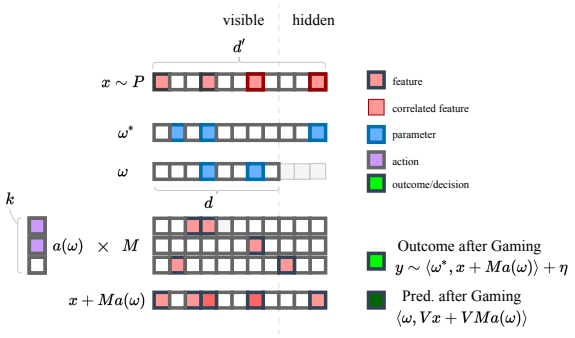

Figure 2. Visualization of the linear setting. Each box corresponds to a real-valued scalar. The two boxes with dark red outlines represent features that are correlated.

Each agent is described by a feature vector $x \in \mathbb{R}^{d'}$,[3] initially drawn from a distribution $P \in \Delta(\mathbb{R}^{d'})$ over the feature-space with covariance matrix Σ. Agents can choose an action vector $a \in \mathbb{R}^k$ to change their features from x to x_g, according to the following update rule: $x_g = x + Ma$ where the *action-effect matrix* $M \in \mathbb{R}^{d \times k}$ has an (i, j)th entry corresponding to the change in the ith feature of x as a result of spending one action unit along the jth direction of the action space. Each action dimension can affect multiple features simultaneously. For example, in the context of car insurance, a prospective customer's action might be "buy a new car", which can increase both the safety rating of the vehicle and the potential financial loss from an accident. The car-buying action might correspond to a column $M_1 = (2, 10000)^T$, in which the two entries represent the action's marginal impact on the car's safety rating and cost-to-refund-if-damaged respectively. M can be rank-deficient, meaning some feature directions are not independently perturbable.

Let y be a random variable representing an agent's true outcome, which we assume is decomposable into a noisy linear combination of the features $y := \langle \omega^*, x_g \rangle + \eta$, where $\omega^* \in \mathbb{R}^{d'}$ is the true *parameter vector*, and η is a subgaussian noise random variable with variance σ. Note that ω_i^* can be understood as the causal effect of a change in feature i on the outcome y, in expectation. Neither the decision-maker nor the agent knows ω^*. Overall, this setting captures any linear structural equation model with perfectly observed features, so long as the outcome itself does not affect the features.

To define the decision-maker's behavior, we must introduce an important aspect of our setting: the decision-maker never observes an agent's complete feature vector x_g, but only

[3] x, and all subsequent notation, uses homogenous coordinates for simplicity.

a subset of of those features Vx_g, where V is a diagonal projection matrix with 1s for the d visible features and 0s for the hidden features.

Now, our decision-maker assigns decisions $\langle \omega, Vx_g \rangle$, where $\omega \in \mathbb{R}^{d'}$ is the *decision rule*. Note that because the hidden-feature-dimensions of ω are never used, we will define them to be 0, and thus ω is functionally defined in a d-dimensional subspace.

For convenience, we define the matrix $G = MM^TV$ as shorthand. (We will see that G maps ω to the movement in agents' feature vectors it incentivizes.)

Agents incur a cost $C(a)$ based on the action they chose. Throughout this work this cost is quadratic $C(a) = \frac{1}{2}\|a\|_2^2$. This corresponds to a setting with increasing costs to taking any particular action.

Importantly, we assume that agents will best-respond to the published decision rule by choosing whiehver action $a(\omega)$ maximizes their utility, defined as their received decision minus incurred action cost:

$$a(\omega) = \arg\max_{a' \in \mathbb{R}^k} \left[\langle \omega, V(x + Ma) \rangle - \frac{1}{2}\|a\|^2 \right] \quad (1)$$

However, to take into account the fact that not all agents will in practice study the decision rule to maximize their utility, we further assume that only a p fraction of agents game the decision rule,[4] while a $1 - p$ fraction remain at their initial feature vector.

Now, the interaction between agents and the decision-maker proceeds in a series of rounds, where a single round *i* consists of the sequence described in the following figure:

For round $t \in \{1, \ldots, r\}$:

1. The decision-maker publishes a new decision rule ω_t.
2. A new set of n agents arrives: $\{x \sim P\}$.
   Each agent games w.r.t. ω_t; i.e. $x + Ma(\omega_t)$.
3. The decision-maker observes the post-gaming visible features $V(x + Ma)$ for each agent.
   Agents receive decisions $\omega_t^T V(x + Ma)$.
4. The decision-maker observes $y \sim (\omega^*)^T(x + Ma) + \eta$ for each agent.

In general, we will assume that the decision-maker cares more about the number of rounds required for an algorithm than the number of agent samples collected in each round.

We now turn to the three objectives that decision-makers may wish to optimize.

Objective 1. The *agent outcomes* objective is the average outcome over the agent population after gaming:

$$\mathbb{E}_{x \sim P, \eta} \left[\langle \omega^*, x + Ma(\omega) \rangle + \eta \right] \quad (2)$$

In subsequent discussion we will restrict ω to the ℓ_2 unit ball, as an infinitely large ω can cause infinitely positive outcomes.

A decision-maker might care about maximizing agent outcomes if it is in their interest for agents to achieve outcomes that are as positive as possible. For example, a teacher formulating a test for their students may care more about incentivizing the students to learn the material than they care about accurately stratifying students based on their knowledge of the material.

Objective 2. *Prediction risk* captures how close the output of the model is to the true outcome. It is measured in terms of expected squared error:

$$\mathbb{E}_{x \sim P, \eta} \left[\left(\langle \omega^*, x + Ma(\omega) \rangle + \eta - \langle \omega, V(x + Ma(\omega)) \rangle \right)^2 \right] \quad (3)$$

A decision-maker cares about minimizing prediction risk if they want the scores they assign to individuals to be as predictive of their true outcomes as possible. For example, insurers' profitability is contingent on neither over- nor under-estimating client risk.

In the realizable linear setting, there is a natural third objective:

Objective 3. *Parameter estimation* measures how close the decision rule's coefficients are to the visible coefficients of the underlying linear model:

$$\|V(\omega - \omega^*)\|_2 \quad (4)$$

Below, we show that these objectives may be mutually non-satisfiable. A natural question is whether we can optimize a weighted combination of these objectives. In Section 4, we outline an algorithm for optimizing a weighted combination of prediction risk and agent outcomes. Our parameter recovery algorithm will only work in the fully-visible ($V = I$) case; in this case, all three objectives are jointly satisfied by $\omega = \omega^*$, though each objective implies a different type of approximation error and thus requires a different algorithm.

2.1. Illustrative example

To illustrate the setting, and demonstrate that in certain cases these objectives are mutually non-satisfiable, we provide a toy scenario, visualized in Figure 3. Imagine a car insurer that predicts customers' expected accident costs given access to three features: (1) whether the customer drives

[4] Note that this means that all strategic agents will, for a given decision rule ω, choose the same gaming action $a(\omega)$.

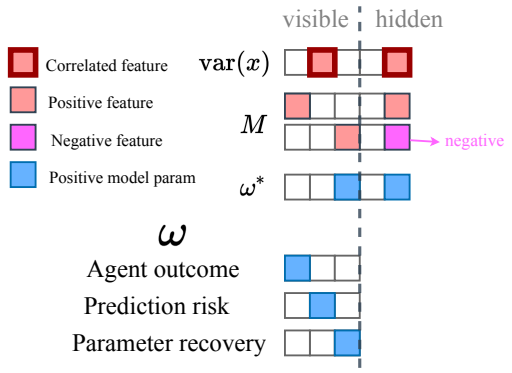

Figure 3. A toy example in which the objectives are mutually non-satisfiable. Each ω optimizes a different objective.

a minivan, (2) whether they own their own car, and (3) whether they have a motorcycle license. There is a single hidden, unmeasured feature: (4) how defensive a driver they are. Let $\omega^* = (0, 0, 1, 1)$, i.e. of these features only knowing how to drive a motorcycle and being a defensive driver actually reduce the expected cost of accidents. Let the initial data distribution have "owning a minivan" correlate with defensive driving (because minivan owners are often parents worried about their kids). Let the first action effect column $M_1 = (1, 0, 0, 2)$ be the action of purchasing a new car, which also makes the customer a much more defensive driver to protect their investment. Let the second action-effect column $M_2 = (0, 0, 1, -2)$ be the action of learning to ride a motorcycle, which slightly improves the likelihood of safe driving by understanding how motorcyclists will react, while making the customer themselves substantially more thrill-seeking and thus reckless in their driving and prone to accidents.

How should the car insurer choose a decision rule, to maximize each objective? Minivans (1) may be a useful predictor because of a historical correlation (good for prediction risk), but anyone buying a minivan specifically to reduce insurance payments will not be a safer driver (unhelpful for agent outcomes) nor does minivan ownership truly cause lower accident costs (bad for parameter estimation). If the rule rewards (2) customers who own their own car, this will make those customers more responsible (good for agent outcomes), but will cause the decision-maker to be inaccurate on the $(1-p)$-fraction of non-gaming agents who already had old cars they worried about less (bad for prediction risk), and owning a car does not itself reduce expected accident costs (bad for parameter estimation). Finally, if the decision rule rewards customers who (3) have a motorcycle license, this does reflect the fact that possessing a motorcycle license itself does reduce a driver's expected accident cost (good for parameter estimation), but an agent acquiring a license will do more harm than good to their overall likelihood of an accident due to the action's side effects (bad for agent outcomes), and rewarding this feature in the decision rule will lead to poor predictions as it is anti-correlated with expected accident cost once the agents have acted (bad for prediction risk).

3. Agent Outcome Maximization

In this section we will propose an algorithm for choosing a decision rule that will maximize agents' outcomes by incentivizing agents to choose actions that maximally increase their outcome. Throughout this section, we will assume that w.l.o.g. $p = 1$. If only a subset of agents are strategic, the non-strategic agents cannot affect their own outcome and can thus be safely ignored. [5]

$$\omega_{imp} = \underset{\omega \in \mathbb{R}^d, \|\omega\|_2 \leq 1}{\arg\max} \; \mathbb{E}_{x \sim P} \left[\omega^{*T} (x + Ma(\omega)) \right] \tag{5}$$

In our car insurance example, this means choosing a decision rule that causes drivers to behave the most safely, regardless of whether the decision rule accurately predicts accident probability.

Theorem 1. *Suppose the covariance matrix Σ has largest eigenvalue $\leq \lambda_{max}$, and suppose the outcome noise η is 1-subgaussian. Then Algorithm 1 learns an approximation of ω_{imp} with squared ℓ_2 error at most ϵ in $d+1$ rounds with $O(\epsilon^{-1}\lambda_{max} d + 1)$ samples in each round. Furthermore, Algorithm 1 is non-adaptive. It succeeds and achieves the desired error with high probability.*

Algorithm 1 Agent Outcome Maximization

Input: scalar parameters λ_{max}, ϵ, matrix V
Let $n = O(\epsilon^{-1}\lambda_{max} d)$
Sample $x_1 \dots x_n$ with $\omega = 0$.
Let $\hat{\mu} = \frac{1}{n}\sum x_i$
for $i = 1$ **to** d **do**
  Sample $x_1 \dots x_n$ with $\omega = e_i$
  Let $\hat{\omega}_i = \frac{1}{n}\sum x_i$
end for
if $||\hat{\omega}|| > 1$ **then**
  Let $\hat{\omega} = \frac{\hat{\omega}}{||\hat{\omega}||}$
end if
Return $\hat{\omega}$

Proof Idea. First we note that it is straightforward to compute the action that each agent will take. The agent maximizes $\omega^T V(x + Ma) - \frac{1}{2}\|a\|^2$ over $a \in \mathbb{R}^m$. Note that

[5] We will require, for the related algorithms, that $\|\omega\|_2 \leq 1$ because without a bound on ω, the decision-maker could provide arbitrarily large incentives to agents, who would then take actions of arbitrary magnitude.

$\nabla_a(\omega^T V M a - \frac{1}{2}\|a\|^2) = M^T V \omega - a$. Thus,

$$\begin{aligned}&\arg\max_a \omega^T V(x + Ma) - \frac{1}{2}\|a\|^2 \\ &= \arg\max_a \omega^T V M a - \frac{1}{2}\|a\|^2 \\ &= M^T V \omega\end{aligned}$$

That is, every agent chooses $x_g = x + MM^T V\omega = x + G\omega$. This means that if the decision-maker publishes ω, the resulting expected agent outcome is ${\omega^*}^T(x + G\omega)$. Hence, the optimal value of ω for the decision–maker to choose is

$$\frac{G^T\omega^*}{\|G^T\omega^*\|_2}$$

Imagine for a moment that there were no outcome noise η. We could directly make linear measurements of perfect quality by choosing ω's that form a basis, as the agent outcomes corresponding to each basis vector decision rule would differ from the $\omega = 0$ case[6] by, on average, $e_i^T G^T \omega^*$. Algorithm 1 is a robust version of this basic intuition. We leave a complete proof to the appendix. □

Note that this procedure will learn to reward any visible features which, even if they do not truly affect the outcome, cause agents to choose actions which do increase the true outcome.

This algorithm has several desirable characteristics. First, the decision-maker who implements the algorithm does not need to have any knowledge of M or even of the number of hidden features $d' - d$. Second, the algorithm is non-adaptive, in that the published decision rule in each round does not depend on the results of previous rounds. Hence, the algorithm can be parallelized by simultaneously applying d separate decision rules to d separate subsets of the agent population and simultaneously observing the results. Finally, by using decision rules as causal interventions, this procedure resolves the challenge associated with the hardness result of (Miller et al., 2020).

4. Prediction Risk Minimization

Low prediction risk is important in any setting where the decision-maker wishes for a decision to accurately match the eventual outcome. For example, consider an insurer who wishes to price car insurance exactly in line with drivers' expected costs of accident reimbursements. Pricing too low would make them unprofitable, and pricing too high would allow a competitor to undercut them.

[6]We can modify this same algorithm to find an "outcome-maximizing adjustment" to some pre-existing decision rule ω' by replacing the base outcome estimate with ω', and adding each basis vector to ω'.

Specifically, we want to minimize expected squared error when predicting the true outcomes of agents, a p-fraction of whom have gamed with respect to ω:

$$\begin{aligned}Risk(\omega) = \mathbb{E}_{x\sim P,\eta}\Bigg[&(1-p)\left(\omega^T V x - (\omega^*)^T V x\right)^2 \\ &+ p\left(\omega^T V A(x,\omega) - \left({\omega^*}^T A(x,\omega) + \eta\right)\right)^2\Bigg]\end{aligned} \tag{6}$$

We begin by noting a useful decomposition of accuracy in the linear setting:

Lemma 1. *Let ω be the linear decision rule and let a be the action taken by agents in response to h. Suppose that Ma and x are uncorrelated. Then the expected squared error of a decision rule ω on the gamed distribution can be decomposed as the sum of the following two positive terms (plus a constant offset):*

1. *The risk of ω on the original distribution*
2. *The squared error of h in predicting the expected impact (on agents' outcomes) of a, weighted by p.*

That is,

$$\begin{aligned}Risk(\omega) \propto\ &\mathbb{E}_{x\sim P}\left[\left((V\omega - \omega^*)^T x\right)^2\right] \\ &+ p\cdot\mathbb{E}_{x\sim P}\left[\left((V\omega - \omega^*)^T (Ma)\right)^2\right]\end{aligned} \tag{7}$$

Note that this lemma holds regardless of the choice of agent action model.

The proof appears in the appendix.

This decomposition illustrates that minimizing prediction risk requires balancing two competing phenomena. First, one must minimize the risk associated with the original (ungamed) agent distribution by rewarding features that are *correlated with outcome* in the original data. Second one must minimize error in predicting the effect of agents' gaming on their outcomes by rewarding features in accordance with *the true change in outcome*. The relative importance of these two phenomena depends on p, the fraction of agents who game.

Unfortunately, minimizing this objective not be straightforward. Even with just squared action cost (with actions $a(\omega)$ linear in ω), the objective becomes a non-convex quartic. However, we will show that in cases where the naive gaming-free predictor *overestimates* the impact of gaming, this quartic can be minimized efficiently.

Remark 1. *Let $\omega_{\text{no-gaming}}$ be the decision rule that minimizes the prediction risk without gaming, and let agent*

action costs be quadratic. If $\omega_{\text{no-gaming}}$ overestimates *the change in outcome as a result of the agent action, then we can find an approximate prediction-risk-minimizing decision rule in $O(\text{poly}(d))$ rounds via a gradient-free convex optimization algorithm.*

Proof Idea. As shown in Lemma 1, the prediction risk consists of two terms: the prediction risk before gaming (a convex quadratic), and the error in estimating the effect of gaming on the outcome, which can be written out as $\left((V\omega - \omega^*)^T(MM^TV\omega^T)\right)^2$ using our equation for $a(\omega)$ in the quadratic action cost case. We'll refer to this second term as the "gaming error", and study its geometry.

The matrix VMM^TV is symmetric and positive-semidefinite, so the gaming error will be a normal quartic as a function of ω. More specifically, there is an affine transformation of ω into a vector z such that this quartic decomposes into the square of the squared magnitude of z minus a constant: $(z_1^2 + z_2^2 + \cdots + z_d^2 - t)^2$ for some $t > 0$. The zero of this equation defines an ellipse. Every ω on the ellipse makes the gaming-error quartic's value 0. Each such point corresponds to perfect prediction of gaming's effects. For any ω corresponding to a point in the interior of the ellipse, ω^TVMa *underestimates* the true effect of gaming $(\omega^*)^TMa$. Conversely, for any ω corresponding to a point outside the ellipse, the decision rule *over-estimates* the true change in outcome. This entire region is convex by composition of a convex monotonic function (the square outside the ellipse) and a convex function (the internal quadratic).

If $\omega_{\text{no-gaming}}$ happens to fall in this convex "overestimate" region, then $Risk(\omega)$ is convex in the neighborhood of this no-gaming decision rule, since it is the sum of two convex functions.

Furthermore, we will show that some global prediction-risk-minimizing ω falls in a convex set which includes $\omega_{\text{no-gaming}}$, and that the prediction risk objective is convex on this set. To see why, assume that the opposite is true. Then any global minimizer ω' has the property that somewhere on the slice of the quartic from $\omega_{\text{no-gaming}}$ to ω' the objective $Risk(\omega)$ is not convex. This is only possible if this slice intersects the zero-ellipse. (This is because the line's starting point is in the convex region, and all non-convex regions are contained within the ellipse.) But at the slice's first intersection with the zero-ellipse (starting from $\omega_{\text{no-gaming}}$) which we call ω_e, the value of the gaming-effect component is 0. Additionally, the gaming-free risk component at ω_e is at least as small as that at ω', because ω_e is strictly closer to $\omega_{\text{no-gaming}}$ than ω'. Thus the value at ω_e is itself a global minimizer with only convex points on the line from it to $\omega_{\text{no-gaming}}$, leading to a contradiction. Therefore, there exists a global minimizer lies in a convex region which also includes the no-gaming optimal decision rule, and for which the prediction risk objective is convex everywhere.

We now have a sketch of our prediction-risk-minimization algorithm: start by collecting data without publishing a decision rule (i.e. publishing the rule $\omega = \vec{0}$), and learn the no-gaming risk-minimizing decision rule $\omega_{\text{no-gaming}}$. Publish this decision rule, and observe empirically whether the predicted decisions over-estimate the true outcomes. If they do, we now know that the global minimum lies in a convex set around our point, and that our risk is convex on this set. We now use a derivative-free convex optimization algorithm to propose a sequence of decision rules, until our procedure converges. Note that we must ensure the queries must remain within the convex region with high probability. □

This raises an interesting observation: in our scenario it is easier to recover from an initial over-estimate of the effect of agents' gaming on the outcome (by reducing the weights on over-estimated features) than it is to recover from an under-estimate (which requires increasing the weight of each feature by an unknown amount).

We make one further observation:

Remark 2. *The procedure described in Remark 1 can also be used to minimize weighted sum of the outcome-maximization and prediction-risk-minimization objectives.*

This follows from the fact that the outcome-maximization objective is linear in ω, and therefore adding it to the prediction-risk objective preserves the convexity/concavity of each of the different regions of the objective. Thus, if a credit scorer wishes to find the optimum of a weighted sum of their predictive accuracy at assigning loans, and the fraction of their customers who successfully repay (according to some weighting), this provides a method for doing so under certain initial conditions.

5. Parameter Estimation

Finally, we provide an algorithm for estimating the causal outcome-generating parameters ω^*, specifically in the case where the features are fully visible ($V = I$).[7] Estimating the causal parameters is desirable both because they deepen our understanding of the outcome-generating phenomenon. Furthermore, when used as a decision rule in the fully-visible case, parameter estimation ensures accuracy and good performance even when the distribution of agents P or their actions M shifts.

Theorem 2. *(Informal) Given $V = I$ (all dimensions are visible) and $\Sigma + \lambda MM^T$ is full rank for some λ (that is, there exist actions that will allow change in the full feature*

[7] For simplicity, we also assume $p = 1$, though any lower fraction of gaming agents can be accomodated by scaling the samples per round.

space), we can estimate ω^* to arbitrary precision. We do so by computing an ω that results in more informative samples, and then gathering samples under that ω. The procedure requires $\tilde{O}(d)$ rounds. See the appendix for details.)

The algorithm that achieves this result is actually simple to sketch. It consists of the following steps:

1. Estimate the covariance of the initial agent feature distribution before strategic behavior Σ by initially not disclosing any decision rule to the agents, and observing their features.
2. Estimate parameters of the Gramian of the action matrix $G = MM^T$ by incentivizing agents to vary each feature sequentially.
3. Use this information to learn the decision function ω which will yield the most informative samples in identifying ω^*, via a convex optimization.
4. Use the new, more informative samples in order to run OLS to compute an estimate of the causally precise regression parameters ω^*.

At its core, this can be understood as running OLS after acquiring a better dataset via the smartest choice of ω (which is again, surprisingly, unique!). Whereas the convergence of OLS without gaming would be controlled by the minimum eigenvalue of the second moment, convergence of our method is governed by the minimum eigenvalue of following matrix:

$$\mathbb{E}[(x + G\omega)(x + G\omega)^T] = \Sigma + 2\mu\omega^T G^T + G\omega\omega^T G^T$$

Our method learns a value of ω that results in the above matrix having a larger minimum eigenvalue, potentially resolving issues of rank and improving the convergence rate.

The proof and complete algorithm description is left to the appendix.

6. Discussion

In this work, we have introduced a model and techniques for analyzing decision-making about strategic agents capable of changing their outcomes. We provide algorithms for leveraging agents' behaviors to maximize agent outcome, minimize prediction risk, and recover the true parameter vector.

Let us dwell on several real-world considerations that should inform the utility of these algorithms. First, while these algorithms eventually yield more desirable decision-making functions, they substantially reduce the decision-maker's accuracy in the short term while exploration is occurring, and this tradeoff should inform their use. In general, these procedures make the most sense in scenarios with a fairly small period of delay between decision and outcome (e.g. predicting short-term creditworthiness rather than long-term professional success), as at each new round the decision-maker must wait this length of time to receive the first samples gamed according to the new rule. That said, our algorithms are either non-adaptive procedures, or have very limited adaptivity. This allows them to be parallelized in a straightforward fashion by using different decision rules for different agents.

It is also important that these methods only be applied to agents whose actions do not change their state in irreversible and possible detrimental ways. In particular, in the case of repeated decision-making (as in a credit scoring setting), an agent may make changes to respond to a temporary decision rule, only to realize their features leave them in a worse position when the decision rule changes as part of the type of algorithm we've described. Only if agents have no future expectation of the consistency of the decision rule, or if they receive a decision only once (as in college admissions), can we be certain that the exploration induced by the decision rule is not exploitative. (After all, agents will only incur action cost if they actually benefit from the decision they'll receive.)

As we have mentioned, our model has several notable social implications. First, in many settings, our results show that decision-makers are incentivized not only to be fully transparent, but to be actively informative. Sharing details about workings of their algorithm can potentially maximize both the decision-maker's and agents' utilities. More radically, agents may actually be incentivized to join together to construct a decision-maker if one does not exist. The agents themselves may wish to know ω^*, and the way to do that is to aggregate their data, and have the agent-led decision-maker reward different agents for gaming in different directions, in order to more quickly identify the true causal parameters, as in Section 5.

Our work opens up several avenues for future work. First, one can explore algorithms that work under a more general set of action cost models, or where different agents have different action costs, or where the actions available to agents are state-dependent. One can explore methods for minimizing the regret of decision-makers during the exploration phase of our algorithms. One could also explore the dynamics of decision rules when agents have persistent states across multiple decisions.

One could also explore improving the efficiency of the algorithms we proposed. Finally, one could extend these results to the setting where both the decision rule, and the true outcome-producing function, are non-linear.

Acknowledgements

The authors would like to thank Suhas Vijaykumar, Cynthia Dwork, Christina Ilvento, Anying Li, Pragya Sur, Shafi Goldwasser, Zachary Lipton, Hansa Srinivasan, Preetum Nakkiran, Thibaut Horel, and our anonymous reviewers for their helpful advice and comments. Ben Edelman was partially supported by NSF Grant CCF-15-09178. Brian Axelrod was partially supported by NSF Fellowship Grant DGE-1656518, NSF Grant 1813049, and NSF Award IIS-1908774.

References

Alon, T., Dobson, M., Procaccia, A., Talgam-Cohen, I., and Tucker-Foltz, J. Multiagent Evaluation Mechanisms. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(02):1774–1781, April 2020.

Ball, I. Scoring Strategic Agents. Manuscript at https://drive.google.com/file/d/1Vlr1MFe_JhL7_IHAGEABhKmEIfZXpDF_/view, 2020.

Bambauer, J. and Zarsky, T. The Algorithm Game. *Notre Dame Law Review*, 94(1):1–48, 2018.

Bechavod, Y., Ligett, K., Wu, Z. S., and Ziani, J. Causal Feature Discovery through Strategic Modification. *arXiv:2002.07024*, February 2020.

Braverman, M. and Garg, S. The Role of Randomness and Noise in Strategic Classification. In Roth, A. (ed.), *1st Symposium on Foundations of Responsible Computing (FORC 2020)*, volume 156 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pp. 9:1–9:20, Dagstuhl, Germany, 2020. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.

Brückner, M. and Scheffer, T. Stackelberg games for adversarial prediction problems. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '11, pp. 547–555, San Diego, California, USA, August 2011. Association for Computing Machinery.

Chen, Y., Podimata, C., Procaccia, A. D., and Shah, N. Strategyproof Linear Regression in High Dimensions. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pp. 9–26, Ithaca NY USA, June 2018. ACM.

Chen, Y., Liu, Y., and Podimata, C. Learning Strategy-Aware Linear Classifiers. *arXiv:1911.04004*, June 2020.

Cummings, R., Ioannidis, S., and Ligett, K. Truthful Linear Regression. In *Conference on Learning Theory*, pp. 448–483, June 2015.

Dekel, O., Fischer, F., and Procaccia, A. D. Incentive compatible regression learning. *Journal of Computer and System Sciences*, 76(8):759–777, December 2010.

Dong, J., Roth, A., Schutzman, Z., Waggoner, B., and Wu, Z. S. Strategic Classification from Revealed Preferences. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, EC '18, pp. 55–70, Ithaca, NY, USA, June 2018. Association for Computing Machinery.

Ederer, F., Holden, R., and Meyer, M. Gaming and strategic opacity in incentive provision. *The RAND Journal of Economics*, 49(4):819–854, 2018.

Frankel, A. and Kartik, N. Improving Information from Manipulable Data. Manuscript at http://arxiv.org/abs/1908.10330, April 2020.

Grossman, S. J. and Hart, O. D. An Analysis of the Principal-Agent Problem. *Econometrica*, 51(1):7–45, 1983.

Haghtalab, N., Immorlica, N., Lucier, B., and Wang, J. Maximizing Welfare with Incentive-Aware Evaluation Mechanisms. Manuscript at https://www.cs.cornell.edu/~nika/pubs/betterx2.pdf, 2020.

Hardt, M., Megiddo, N., Papadimitriou, C., and Wootters, M. Strategic Classification. In *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science*, ITCS '16, pp. 111–122, Cambridge, Massachusetts, USA, January 2016. Association for Computing Machinery.

Hölmstrom, B. Moral Hazard and Observability. *The Bell Journal of Economics*, 10(1):74–91, 1979.

Holmstrom, B. and Milgrom, P. Multitask Principal–Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *The Journal of Law, Economics, and Organization*, 7(special_issue):24–52, January 1991.

Hu, L., Immorlica, N., and Vaughan, J. W. The Disparate Effects of Strategic Manipulation. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, FAT* '19, pp. 259–268, Atlanta, GA, USA, January 2019. Association for Computing Machinery.

Ioannidis, S. and Loiseau, P. Linear Regression as a Non-cooperative Game. In Chen, Y. and Immorlica, N. (eds.), *Web and Internet Economics*, Lecture Notes in Computer Science, pp. 277–290, Berlin, Heidelberg, 2013. Springer.

Kleinberg, J. and Raghavan, M. How Do Classifiers Induce Agents to Invest Effort Strategically? In *Proceedings of the 2019 ACM Conference on Economics and Computation*, EC '19, pp. 825–844, Phoenix, AZ, USA, June 2019. Association for Computing Machinery.

Laffont, J.-J. and Martimort, D. *The Theory of Incentives*. Princeton University Press, 2002.

Liu, L. T., Wilson, A., Haghtalab, N., Kalai, A. T., Borgs, C., and Chayes, J. The disparate equilibria of algorithmic decision making when individuals invest rationally. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, FAT* '20, pp. 381–391, Barcelona, Spain, January 2020. Association for Computing Machinery.

Miller, J., Milli, S., and Hardt, M. Strategic Classification is Causal Modeling in Disguise. *arXiv:1910.10362*, February 2020.

Milli, S., Miller, J., Dragan, A. D., and Hardt, M. The Social Cost of Strategic Classification. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, FAT* '19, pp. 230–239, Atlanta, GA, USA, January 2019. Association for Computing Machinery.

Pearl, J. *Causality*. Cambridge University Press, Cambridge, U.K., second edition edition, 2000.

Perdomo, J. C., Zrnic, T., Mendler-Dünner, C., and Hardt, M. Performative Prediction. *arXiv:2002.06673*, April 2020.

Perote, J. and Perote-Peña, J. Strategy-proof estimators for simple regression. *Mathematical Social Sciences*, 47(2): 153–176, March 2004.

Peters, J., Janzing, D., and Schölkopf, B. *Elements of Causal Inference : Foundations and Learning Algorithms*. The MIT Press, 2017.

Salanié, B. *The Economics of Contracts: A Primer, 2nd Edition*. MIT Press, March 2005.

Tabibian, B., Tsirtsis, S., Khajehnejad, M., Singla, A., Schölkopf, B., and Gomez-Rodriguez, M. Optimal Decision Making Under Strategic Behavior. *arXiv:1905.09239*, February 2020.

Ustun, B., Spangher, A., and Liu, Y. Actionable Recourse in Linear Classification. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, FAT* '19, pp. 10–19, Atlanta, GA, USA, January 2019. Association for Computing Machinery.

Venkatasubramanian, S. and Alfano, M. The philosophical basis of algorithmic recourse. *Fairness, Accountability, and Transparency Conference 2020*, forthcoming.

Ziewitz, M. Rethinking gaming: The ethical work of optimization in web search engines. *Social Studies of Science*, 49(5):707–731, October 2019.