# Causal Modeling for Fairness in Dynamical Systems: A Case Study in Lending

**Anonymous authors**
Paper under double-blind review

## Abstract

In many application areas—lending, education, and online recommenders, for example—fairness and equity concerns emerge when a machine learning system interacts with a dynamically changing environment to produce both immediate and long-term effects for individuals and demographic groups. This paper investigates the benefits of *causal reasoning* and the role of modeling *interventions* in these settings. Through a detailed case study, we illustrate how causal assumptions enable simulation (when environment dynamics are known) and off-policy estimation (when dynamics are unknown) of interventions on short- and long-term outcomes, for both groups and individuals.

## 1 Introduction

How do we design fair policies for complex, evolving systems? Recently, the literature on fairness in dynamical systems (a.k.a. “feedback loops”) has begun exploring the role of algorithmic systems in shaping their environments over time (Hashimoto et al., 2018; Lum & Isaac, 2016; Ensign et al., 2018; D’Amour et al., 2020). The key insight from these papers is that the repeated application of algorithmic tools in a changing environment can cause fairness impacts in the long-term distinct from those in the short-term.

While the methods in this literature are quite disparate, we note that causal directed acyclic graphs (DAGs) (Pearl, 2009; Richardson & Robins, 2013) serve as a unifying framework for all of these papers (assuming loops are rolled out over finite horizons). While causal DAGs have been used to study one-shot fair decision-making (Kusner et al., 2017; 2019; Kilbertus et al., 2017), they are uncommon in fairness settings involving *sequential* decisions. The mechanism of *intervention* enables causal reasoning that we argue addresses critical problems in the practical deployment of “fair” sequential decision makers, such as off-policy evaluation from *biased* observational data. To illustrate the benefits of causal modeling, we provide a case study showing how causal reasoning can be brought to bear in the lending setting proposed by Liu et al. (2018). We show empirically how causal reasoning improves off-policy evaluation and enables sensitivity analysis.

## 2 The Lending Model: A Causal Re-interpretation

**Notation** There are several ways to encode causal assumptions in DAG form. In this paper, we focus on structural causal models (SCMs) (Pearl, 2009)[1]. Graphically, *endogenous* and *exogenous* nodes depict the structural assumptions of the SCM: each endogenous node is the output of a deterministic structural equation while the exogenous nodes represent stochasitcity in the generative process. For example, $Z = f_Z(\text{Parents}(Z), U_Z)$ might represent some (possibly stochastic) *treatment policy*. We denote by $p$ the joint distribution specified by the SCM. Under the *do*-operation (i.e. graph surgery), various *interventional distributions* can be derived. *Atomic intervention* on the treatment value ("What would happen if all treatments were set to 1?") induces the distribution $p^{do(Z=1)}$, while *policy intervention* on the treatment function ("What would happen if a new non-constant treatment function $\hat{f}_Z$ were applied?") induces the distribution $p^{do(f_Z \to \hat{f}_Z)}$. When computing expected outcomes under intervention, we specify the interventional distribution in the subscript of the expectation.

[1] Overviews of SCMs at various levels of detail can be found elsewhere (Pearl, 2009; Madras et al., 2019; Buesing et al., 2019)

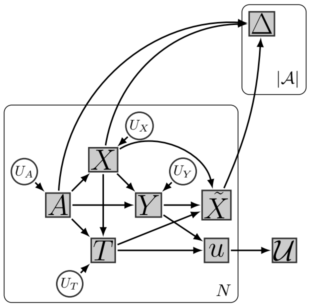


Figure 1: Our structural causal model (SCM) re-interpretation of the one-step model from Liu et al. (2018). See Section 2 for discussion and Appendix B.3 for symbol legend.

**Lending SCM** Liu et al. (2018) studied single-step dynamics of threshold-based classifiers, with a special focus on lending. Our SCM re-interpretation of this model can be seen in Figure 1. In this model, a person with group membership (a.k.a. sensitive attribute) $A$ receives a credit score $X$, and applies to a bank for a loan. The bank makes a binary decision $T$ about whether to award the loan using the policy $f_T$. The binary potential outcome $Y$ is realized, which is converted to institutional profit or loss only if $T = 1$[2]. Finally, the applicant's credit score is modified to $\tilde{X}$ (increased on repayment, decreased on default, static if $T = 0$)[3]. The bank's utility is measured through their profit $\mathcal{U}$ (a sum over the individual profits $u$) as well as the expected score change $\Delta_j$, representing the average change in credit score after one time-step among members of group $A = j$. Varying the loan policy can achieve different values of $\mathcal{U}, \Delta_j$, resulting in outcomes with different fairness properties.

Liu et al. (2018) consider the effect of various threshold policies for loan assignment under this model, namely the expected values of $\mathcal{U}$ and $\Delta$ for some policies with group-specific thresholds $\tau \triangleq (\tau_0, \tau_1)$ that offer loans to applicants of group $j$ with score $X$ if and only if their credit score $X > \tau_j$. They show that different thresholds satisfy different criteria: maximum profit (MAXPROF), demographic parity (DEMPAR), and equal opportunity (EQOPP). In the language of our paper, comparing threshold policies is done through policy evaluation and intervention. Denoting by $\pi_\tau$ a policy with per-group thresholds $\tau$, these results can be phrased with the tool of policy intervention: we evaluate the policy $\pi_\tau$ by estimating the quantities $\mathbb{E}_{p^{do(f_T \rightarrow \pi_\tau)}}[\mathcal{U}]$ and $\mathbb{E}_{p^{do(f_T \rightarrow \pi_\tau)}}[\Delta_j]\ \forall j$, for various $\tau$ computed under different fairness criteria.

## 3 EXPERIMENTS

We begin by relaxing the assumptions of known dynamics and show that causal inference improves off-policy estimation. We then extend the original model to a longer time horizon and provide a sensitivity analysis. In a supplementary experiment (Appendix C), we extend the model to include the credit scoring bureau as a additional agent in the system.

### 3.1 OFF-POLICY EVALUATION AND LEARNING

We consider a *off-policy evaluation* scenario where the bank has historical data from a profit-maximizing policy (MAXPROF) and wishes to learn and estimate the quality of an equal opportunity policy (EQOPP) before deploying it. Assume that *observational data* must be used (i.e. online A/B testing is unsafe or unethical). In the lending SCM (see Figure 1 for depiction and Appendix B for full specification), the key (non-trivial) structural functions of the SCM are:

$$\begin{aligned} X &= f_X(A, U_X) \\ T &= f_T(T, X, A, U_T) \\ Y &= f_Y(X, A, U_Y) \end{aligned} \tag{1}$$

[2] Therefore this model does not capture a notion of opportunity loss for *not* extending a loan to applicants who *are* qualified.

[3] Likewise, the applicant's score does not change in the absence of a loan; this assumption may be inaccurate, since *not* receiving a loan could create additional financial issues for the applicant.

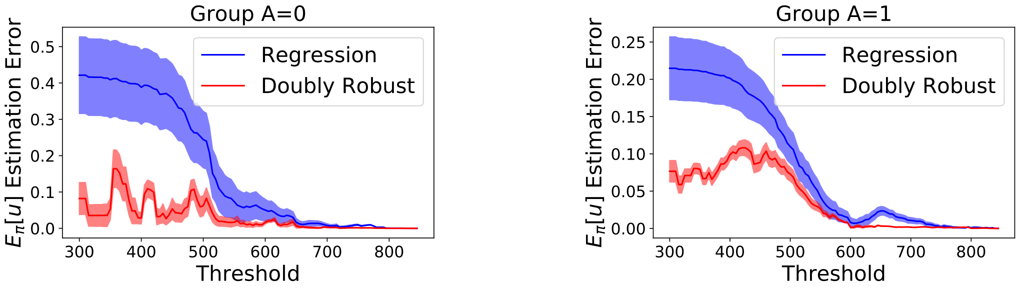


Figure 2: Errors of $\mathcal{E}_{Reg}$ and $\mathcal{E}_{DR}$ (regression and doubly robust) for off-policy estimation of $\mathbb{E}_\pi[u]$ of single threshold policies from observational data. Estimation for $\Delta$ (also linear in $Y$) is similar.

which are the feature distribution, the historical treatment policy, and the outcome distribution, respectively. The change in score $\Delta$, the bank's utility $u$, and the next-step score $\tilde{X}$, are simple functions of the other variables: $(\Delta, u) = (c_+, u_+)$ if $Y = 1$ or $(c_-, u_-)$ if $Y = 0$, and $\tilde{X} = X + \Delta$ (for constants $c_+, u_+ > 0; c_-, u_- < 0$). As in Liu et al. (2018), we focus on threshold policies, which are defined by group-specific thresholds $\tau \triangleq (\tau_0, \tau_1)$ that offer loans to applicants of group $j$ with score $X$ if and only if their credit score $X > \tau_j$.

Liu et al. (2018) make a very strong assumption in their method — that these underlying dynamics parameters $(f_X, f_T, f_Y, c_+, c_-, u_+, u_-)$ of the system are known (this is stronger than just assuming the causal structure, as we do in Fig. 1). While some of these unknown parameters (e.g. $u_+, u_-, f_T$) are easy to estimate from data. One in particular is difficult: the outcome function $Y = f_Y(X, A, U_Y)$. Note that $Y$ is a *causal* quantity representing a *potential outcome* (Rubin, 2005): it is the probability of a person repaying a loan *were they to receive one*[4]. In observational data, $Y$ is *missing-not-at-random*: we only observe $Y$ when a loan was given. Therefore, straightforward estimates may be biased or high variance. This difficulty of estimating $Y$ propagates into the rest of the problem, i.e. $u$ and $\Delta$ also represent potential outcomes that are missing if $T = 0$. Therefore, choosing the policy thresholds—which involves estimating $(u, \Delta)$—is inherently a causal problem.

Given a policy $\pi$, we focus on computing an off-policy estimator $\mathcal{E}(\pi) \approx \mathbb{E}_{p^{do(f_T \rightarrow \pi)}}[u]$. We consider two estimators. The naive baseline is derived from logistic regression on the observational data: first learn a function to approximate $f_{\text{Reg}}(X, A) \approx \mathbb{E}_{p^{\text{obs}}}[u|X, A]$ in the observational data; then apply this regression for every individual where $\pi$ suggests giving the treatment: $\mathcal{E}_{Reg}(\pi) = \mathbb{E}_{p^{\text{obs}}(X,A)}[f_{\text{Reg}}(X, A)|\pi(X, A) = 1]$. The causally-inspired estimator treats estimating $u$ as a missing data problem. Noting that $(X, A)$ satisfy the backdoor criterion w.r.t. $u$ justifies the use of a Doubly Robust estimator (Zhang et al., 2012), a variant of inverse probability weighting that exhibits low variance (Bang & Robins, 2005). With $C_i = \mathbb{1}[\pi(X_i, A_i) = T]$, we have

$$\mathcal{E}_{DR} = \frac{1}{n} \sum_{i=1}^{n} \Big[ \frac{C_i(\pi) u_i}{P(C_i(\pi) = 1|X_i, A_i)} \\ - \frac{C_i(\pi) - P(C_i(\pi) = 1|X_i, A_i)}{P(C_i(\pi) = 1|X_i, A_i)} f_{Reg}(X_i, A_i) \Big].$$

We can use an analogous estimator for $\Delta$, where the same backdoor criterion holds.

We generate observational data from the SCM in Figure 1, under a MaxProf threshold policy. We then consider a new policy $\pi_\tau$ with per-group thresholds $\{\tau_j\}$. We compute the estimators $\mathcal{E}_{Reg}(\pi_\tau)$ and $\mathcal{E}_{DR}(\pi_\tau)$ for varying values of these thresholds. Figure 2 shows that the causally motivated estimator $\mathcal{E}_{DR}$ achieves lower off-policy estimation error on both sensitive groups, across the threshold range. Note the high estimation error of the baseline $\mathcal{E}_{Reg}$ for low values of $\tau$. This is because the historical policy typically does not award loans to applicants with low scores, meaning there are fewer data available for the regression.

Ultimately, the goal of estimating these quantities is to improve policy learning. We can formulate an objective that trades off between utility and an equal opportunity objective $\delta_{EqOpp} = |P(T = 1|Y =$

[4] Using the notation of Rubin (2005), we could denote it as $Y_1$.

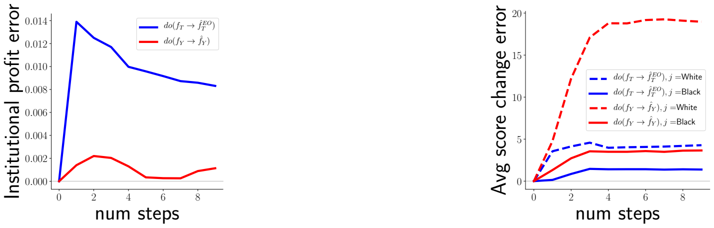


Figure 4: Evaluating multi-step policy robustness to distribution shift for various choice of intervention distribution $q$. Sensitivity of institutional utility—formally $|\mathbb{E}_q[\mathcal{U}] - \mathbb{E}[\mathcal{U}]|$—and sensitivity of group avg. score change—formally $|\mathbb{E}_q[\Delta_j] - \mathbb{E}[\Delta_j]|$—are shown as a function of steps. Expected profit is relatively robust to both interventions, whereas the expected per-group score changes are relatively more sensitive to these interventions.

$1, A = 0) - P(T = 1|Y = 1, A = 1)|$. The objective is $\mathcal{V}_\pi = \mathcal{U} - \lambda\delta_{EqOpp}$. We hope to maximize this, with some hyperparameter $\lambda \in \mathbb{R}$ governing the tradeoff. Estimating $\delta_{EqOpp}$ itself presents a challenging causal problem, since $Y$ is frequently unobserved (see Appendix A for details).

Using the estimators presented above, we can construct an off-policy estimate of $\mathcal{V}_\pi$ We search over the space of two-threshold policies (one threshold per group) to find the policy with the highest off-policy estimate of the objective on a validation set. We then calculate the true value of $\mathcal{V}_\pi$ on a held-out test set, using our simulator to generate the true potential outcomes. The estimator $\mathcal{E}_{DR}$ that more fully incorporates causal reasoning in the parameter estimation finds a better objective value, ultimately yielding an improved policy (see Fig. 3). We emphasize that this improvement requires assumptions about causal structures, but not precise knowledge of the system dynamics.

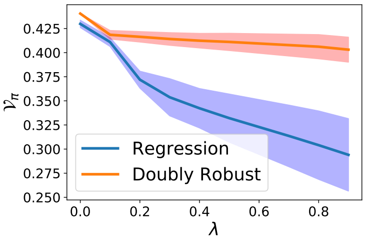


Figure 3: Test set value of a fairness-utility objective using the two off-policy estimators. Hyperparameter $\lambda$ governs the tradeoff. Higher $\mathcal{V}_\pi$ is better.

## 3.2 Sensitivity Analysis of Long-term Outcomes

Sensitivity analysis (Rosenbaum, 2014; Saltelli et al., 2008) measures how a model responds to changes in its underlying assumptions. It is especially relevant to long-term fairness, where errors compound over time. Causal DAGs are a natural match for sensitivity analysis since they make structural assumptions explicit. We show how to conduct a long-term sensitivity analysis in the lending SCM by casting sensitivity analysis as on-policy evaluation under an intervention that accounts for model mismatch[5].

To handle multiple steps we alter the structural equation on scores to the recursive update $X^{t+1} = f_{X^{t>0}}(X^t, Y^t, T^t)$ (See Figure 5 in Appendix B.2 for depiction). $A$ affects treatments and outcomes at every step. We analyze the sensitivity of the EqOpp policy to two forms of model mismatch. In the first, do$(f_T \to \hat{f}_T^{EO})$ recomputes the per-group thresholds under the EqOpp constraint, but using *incorrect* statistics from the credit bureau. In particular, the marginal $p(Y|X)$ was is for both group's repayment probabilities rather than the correct $p(Y|X, A)$. The second intervention do$(f_Y \to \hat{f}_Y)$ is more severe, as $p(Y|X)$ is used to *sample* potential outcomes $Y$ rather than just set the thresholds within $f_T$. We measure error under each intervention relative to the "ground truth" baseline where the correct potential outcome distributions are used to set thresholds and sample data. We measure how these errors compound over time (Figure 4). We observe the institutional profits are surprisingly robust to both forms of intervention, while the per-group outcomes are more sensitive to these interventions, especially to do$(f_Y \to \hat{f}_Y)$. This suggests a policy fairness sensitivity to assumptions around the distribution over potential outcomes across sensitive groups, which underscores the importance of accurately estimating this distribution from observational data as in Section 3.1.

[5] "Mismatch" refers here to structural equations with misspecified functional forms, not incorrect assumptions of causal structure.

## References

Bang, H. and Robins, J. M. Doubly robust estimation in missing data and causal inference models. *Biometrics*, 61(4):962–973, 2005.

Buesing, L., Weber, T., Zwols, Y., Racaniere, S., Guez, A., Lespiau, J.-B., and Heess, N. Woulda, coulda, shoulda: Counterfactually-guided policy search. In *International Conference on Representation Learning*, 2019.

D'Amour, A., Srinivasan, H., Atwood, J., Baljekar, P., Sculley, D., and Halpern, Y. Fairness is not static: deeper understanding of long term fairness via simulation studies. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pp. 525–534, 2020.

Ensign, D., Friedler, S. A., Neville, S., Scheidegger, C., and Venkatasubramanian, S. Runaway feedback loops in predictive policing. In *Conference on Fairness, Accountability and Transparency*, pp. 160–171, 2018.

Gumbel, E. J. and Lieblein, J. Some applications of extreme-value methods. *The American Statistician*, 8(5):14–17, 1954.

Hashimoto, T. B., Srivastava, M., Namkoong, H., and Liang, P. Fairness without demographics in repeated loss minimization. In *International Conference on Machine Learning*, 2018.

Kilbertus, N., Carulla, M. R., Parascandolo, G., Hardt, M., Janzing, D., and Schölkopf, B. Avoiding discrimination through causal reasoning. In *Advances in Neural Information Processing Systems*, pp. 656–666, 2017.

Kusner, M., Russell, C., Loftus, J., and Silva, R. Making decisions that reduce discriminatory impacts. In *International Conference on Machine Learning*, pp. 3591–3600, 2019.

Kusner, M. J., Loftus, J., Russell, C., and Silva, R. Counterfactual fairness. In *Advances in Neural Information Processing Systems*, pp. 4066–4076, 2017.

Liu, L., Dean, S., Rolf, E., Simchowitz, M., and Hardt, M. Delayed impact of fair machine learning. In *International Conference on Machine Learning*, pp. 3156–3164, 2018.

Lum, K. and Isaac, W. To predict and serve? *Significance*, 13(5):14–19, 2016.

Maddison, C. J., Tarlow, D., and Minka, T. A* sampling. In *Advances in Neural Information Processing Systems*, pp. 3086–3094, 2014.

Madras, D., Creager, E., Pitassi, T., and Zemel, R. Fairness through causal awareness: Learning causal latent-variable models for biased data. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, pp. 349–358. ACM, 2019.

Oberst, M. and Sontag, D. Counterfactual off-policy evaluation with gumbel-max structural causal models. In *International Conference on Machine Learning*, pp. 4881–4890, 2019.

Pearl, J. Causal inference in statistics: An overview. *Statistics surveys*, 3:96–146, 2009.

Reserve, U. F. Report to the congress on credit scoring and its effects on the availability and affordability of credit. *Washington, DC: Board of Governors of the Federal Reserve System*, 2007.

Richardson, T. S. and Robins, J. M. Single world intervention graphs (swigs): A unification of the counterfactual and graphical approaches to causality. *Center for the Statistics and the Social Sciences, University of Washington Series. Working Paper*, 128(30):2013, 2013.

Rosenbaum, P. R. Sensitivity analysis in observational studies. *Wiley StatsRef: Statistics Reference Online*, 2014.

Rubin, D. B. Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association*, 100(469):322–331, 2005.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S. *Global sensitivity analysis: the primer*. John Wiley & Sons, 2008.

Schulman, J., Heess, N., Weber, T., and Abbeel, P. Gradient estimation using stochastic computation graphs. In *Advances in Neural Information Processing Systems*, pp. 3528–3536, 2015.

Zhang, B., Tsiatis, A. A., Laber, E. B., and Davidian, M. A robust method for estimating optimal treatment regimes. *Biometrics*, 68(4):1010–1018, 2012.

# A EXPERIMENTAL DETAILS FOR OFF-POLICY EVALUATION AND LEARNING

Here, we discuss details on the setup for the off-policy evaluation experiment in Sec. 3.1.

## A.1 DATA GENERATION

We generate data from the Liu et al. (2018) model, described in full in Appendix B. We use $(c_+, c_-) = (75, -150)$ and $(u_+, u_-) = (1, -4)$. We use a single threshold policy of $\tau_j = 620 \forall\, j$. We generate 13 data sets of 10000 examples each, using 11 for training (to get confidence intervals), 1 for validation, and 1 for test.

In order to use re-weighting estimators, we must have *overlap* i.e. each point $(X, A)$ must have a non-zero probability of receiving each treatment in the observational data. Since a threshold policy does not satisfy this, we flipped the treatment chosen by the threshold policy with a probability of 0.1.

## A.2 TREATMENT AND OUTCOME MODELS

We use L2-regularized logistic regression for both the treatment and the outcome model using the "liblinear" default solver in sklearn. We train a treatment and outcome model on each of the 11 training sets, and use these to construct our confidence intervals.

## A.3 ESTIMATION OF EQUAL OPPORTUNITY DISTANCE

We define the equal opportunity metric $\delta_{EqOpp}$ as

$$\delta_{EqOpp} = |P(T = 1|Y = 1, A = 0) - P(T = 1|Y = 1, A = 1)|. \tag{2}$$

The key unit in this expression is $P(T = 1|Y = 1)$ (removing $A = a$ from the right side for clarity). This is non-trivial to estimate, since $Y$ is unobserved for many cases.

We take the following approach. First, using Bayes rule, we have

$$P(T = 1|Y = 1) = \frac{P(Y = 1|T = 1)P(T = 1)}{P(Y = 1)}. \tag{3}$$

$P(T = 1)$ is easy to estimate from observational data. $P(Y = 1|T = 1)$ is the off-policy estimation question — we use either $\mathcal{E}_{Reg}$ or $\mathcal{E}_{DR}$ to estimate this. We estimate $P(Y = 1)$ using off-policy estimation as well, noting that $P(Y = 1) = P(Y = 1|\tilde{T} = 1)$, if $\tilde{T} \perp Y$. Therefore, we can obtain an estimate for the marginal distribution of $Y$ by doing off-policy estimation for random policies $\tilde{T}$ (again, using either $\mathcal{E}_{Reg}$ or $\mathcal{E}_{DR}$). We choose 10 random Bernoulli policies to obtain 10 estimates of $P(Y = 1)$ and average them.

## A.4 THRESHOLD SEARCH

In both the estimation (Fig. 2) and selection (Fig. 3) experiments, we consider all thresholds $\tau \in [300, 850)$ such that $\tau \text{mod} 5 = 0$ (where 300 and 850 are the minimum and maximum credit scores in the dataset). To choose our best thresholds in the selection experiment, we consider all pairs of group-specific thresholds $(\tau_0, \tau_1)$, and estimate the value of $\mathcal{V}_\pi$ for the policy associated with those thresholds. We find the optimal value on the validation set, and test them to obtain a final value on the test set Since we do not require overlap to hold in the target policy, we consider hard threshold policies (we do not flip any predictions post-hoc, as we do in the observational data). In the selection experiment, we test $\lambda$ in increments of 0.1 from 0 to 0.9.

# B SCM Details

## B.1 Parameterization for Single-Step Model

As briefly discussed above, Liu et al. (2018) propose a one-step feedback model for a decision-making setting then analyze several candidate policies—denoted by the structural equation $f_T$ in our analysis—by simulating one step of dynamics to compute the institution's profit and group outcomes for each policy. Figure 1 shows our SCM formulation of this dynamics model. Here we provide expressions for the specific structural equations used.

To sample over $p(X, A)$ we start with Bernoulli sampling of $A$, parameterized SCM-style like

$$U_{A_i} \sim \text{Bernoulli}(U_{A_i}|\theta); \quad A_i = f_A(U_{A_i}) \triangleq U_{A_i} \tag{4}$$

where $\theta \in [0, 1]$ is the proportion of the $A = 1$ group.

We then sample scores by the inverse CDF trick[6]. Given an inverse cumulative distribution function $\text{CDF}_j^{-1}$ for each group $j \in \{0, 1\}$, we can write

$$U_{X_i} \sim \text{Uniform}(U_{X_i}|[0, 1]) \tag{5}$$

$$X_i = f_X(U_{X_i}, A_i) \triangleq \text{CDF}_{A_j}^{-1}(U_{X_i}) \tag{6}$$

Liu et al. (2018) discuss implementing threshold policies for each group $j \in \{0, 1\}$, which are parameterized by thresholds $c_j$ and tie-breaking Bernoulli probabilities $\gamma$ (for simplicity of exposition we assume the tie-breaking probability is shared across groups). The original expression was

$$\mathbb{P}(T = 1|X, A = j) = \begin{cases} 1 & X > c_j \\ \gamma & X = c_j \\ 0 & X < c_j. \end{cases} \tag{7}$$

Then, after denoting by $\mathbb{1}(\cdot)$ the indicator function, we can rephrase this distribution in terms of a structural equation governing treatment:

$$U_{T_i} \sim \text{Bernoulli}(U_{T_i}|\gamma) \tag{8}$$

$$\begin{aligned} T_i &= f_T(U_{T_i}, X_i, A_i) \\ &\triangleq 1^{\mathbb{1}(X_i > c_{A_i})} \cdot U_{T_i}^{\mathbb{1}(X_i = c_{A_i})} \cdot 0^{\mathbb{1}(X_i < c_{A_i})}. \end{aligned} \tag{9}$$

A policy $f_T$ (which itself may or may not satisfy some fairness criteria) is evaluated in terms of whether loans were given to creditworthy individuals, and in terms of whether each demographic group successfully repaid any allocated loans on average. To capture the notion of *creditworthiness*, we introduce a *potential outcome* $Y$ (repayment if the loan were given) for each individual, which is drawn[7] from $p(Y|X, A)$[8]. By convention $T = 1$ as the "positive" treatment (e.g., got loan) and $Y = 1$ as the "positive" outcome (e.g., would have repaid loan if given) Note that $Y$ is independent of $T$ given $X$, meaning $Y$ is really an indicator of *potential* success. Formally, the potential outcome $Y$ is distributed as $Y_i \sim \text{Bernoulli}(Y_i|\boldsymbol{\rho}(X_i, A_i))$ for some function $\boldsymbol{\rho} : X \times A \to [0, 1]$. We reparameterize this as a structural equation using the Gumbel-max trick[9] (Gumbel & Lieblein, 1954;

[6] This standard trick is used for sampling from distributions with know densities. Recalling that $\text{CDF}_p : \mathcal{X} \to [0, 1]$ is a monotonic (invertible) function representing $\text{CDF}_p(X') = \int_{-\infty}^{X'} dX p(X < X')$. Then to sample $X' \sim p$ we first sample $U \sim \text{Uniform}(U|[0, 1])$ then compute $X' = \text{CDF}_p^{-1}(U)$.

[7] The authors denoted by $\boldsymbol{\rho}(x)$ the probability of potential success at score $X$. Various quantities were then computed, e.g., $\boldsymbol{u}(x) = u_+\boldsymbol{\rho}(x) + u_-(1 - \boldsymbol{\rho}(x))$. We observe that this is equivalent to marginalizing over potential outcomes $\boldsymbol{u}(x) = \mathbb{E}_{p(Y|X)}[u_+Y + u_-(1 - Y)]$; in our simulations we compute such expectations via Monte Carlo sampling with values of $Y$ explicitly sampled.

[8] The authors use $\boldsymbol{\rho}(X) = p(Y|X)$ in their analysis (suggesting that potential outcome is independent of group membership conditioned on score) but $\boldsymbol{\rho}(X, A) = p(Y|X, A)$ in the code, i.e. the potential outcome depends differently on score for each group. The SCM as expressed in Figure 1 represents the codebase version.

[9] This trick reparameterizes a Categorical or Bernoulli sample as a deterministic transformation of a Uniform sample. See Oberst & Sontag (2019) for discussion of how to perform counterfactual inference for SCMs with Categorical random variables.

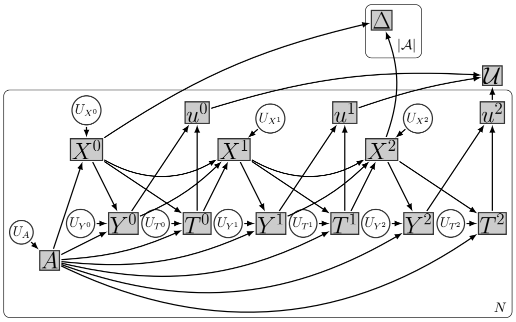


Figure 5: Phrasing the model from Liu et al. (2018) as an SCM enables a multi-step extension for measuring long-term impacts, e.g., in the two-step version shown here.

Maddison et al., 2014):

$$
\begin{aligned}
U_{Y_i} &\sim \text{Uniform}(U_{Y_i}|[0,1]) \qquad (10)\\
Y_i &= f_Y(U_Y, X_i, A_i)\\
&\triangleq \mathbb{1}\left(\log \frac{\boldsymbol{\rho}(X_i, A_i)}{1-\boldsymbol{\rho}(X_i, A_i)} + \log \frac{U_Y}{1-U_Y} > 0\right). \qquad (11)
\end{aligned}
$$

The institutional utility $u_i$ and the updated individual score $\tilde{X}_i$ are deterministic functions of the outcome $Y_i$ and the treatment $T_i$, and the original score $X_i$:

$$
u_i = f_u(Y_i, T_i) \triangleq \begin{cases} u_+^{\mathbb{1}(Y_i)=1} \cdot u_-^{\mathbb{1}(Y_i)=0} & \text{if } T_i = 1 \\ 0 & \text{else} \end{cases}, \qquad (12)
$$

$$
\tilde{X}_i = f_{\tilde{X}}(X_i, Y_i, T_i) \triangleq \begin{cases} X_i + c_+^{\mathbb{1}(Y_i)=1} \cdot c_-^{\mathbb{1}(Y_i)=0} & \text{if } T_i = 1 \\ X_i & \text{else} \end{cases}. \qquad (13)
$$

As mentioned in Section 2, $\{u_+, u_-, c_+, c_-\}$ are fixed parameters that encode expected gain/loss in utility/score based on payment/default of loan.

There are two *global* quantities of interest. Firstly, the institution cares about its overall utility at the current step (ignoring all aspects of the future), expressed as

$$
\mathcal{U} = f_{\mathcal{U}}(u_{1\ldots N}) \triangleq \frac{1}{N} \sum_{i=1}^{N} u_i. \qquad (14)
$$

Secondly, society (and possibly the institution) might care the average per-group score change induced by the policy, expressed for group $A = j$ as

$$
\Delta_j = f_{\Delta_j}(X_{1\ldots N}, \tilde{X}_{1\ldots N}, A_{1\ldots N}) \triangleq \frac{1}{N_{A_j}} \sum_{i=1}^{N} (\tilde{X}_i - X_i)^{\mathbb{1}(A_i=j)}, \qquad (15)
$$

with $N_{A_j} \triangleq \sum_{i'} \mathbb{1}(A_{i'} = j)$ is the size of the $A_j = 1$ group.

### B.2 Multi-step Extension

Figure 5 depects the SCM for the multi-step extension to the one-step lending model proposed by Liu et al. (2018).

### B.3 Symbol Legends

Table 1 decodes the symbols used in the various SCMs (e.g., Figure 1).

| **Symbol** | **Meaning** |
|---|---|
| $N$ | Number of individuals |
| $\lvert\mathcal{A}\rvert$ | Number of demographic groups |
| $A_i$ | Sensitive attribute for individual $i$ |
| $U_{A_i}$ | Exogenous noise on sensitive attribute for individual $i$ |
| $X_i$ | Score for individual $i$ |
| $U_{X_i}$ | Exogenous noise on score for individual $i$ |
| $Y_i$ | Potential outcome (loan repayment/default) for individual $i$ |
| $U_{Y_i}$ | Exogenous noise on potential outcome for individual $i$ |
| $T_i$ | Treatment (institution gives/withholds loan) for individual $i$ |
| $U_{T_i}$ | Exogenous noise on treatment for individual $i$ |
| $u_i$ | Utility of individual $i$ (from the institution's perspective) |
| $\Delta_i$ | Expected improvement of score for individual $i$ |
| $\tilde{X}_i$ | Score for individual $i$ after one time step |
| $\mathcal{U}$ | Global utility (from institution's perspective) |
| $\Delta_j$ | Expected change in score for group $j$ |

Table 1: Symbol legend for Figure 1

## C Multi-Actor Experiment

In this section we discuss an additional experiment that demonstrates how causal DAGs can be used as *expressive simulators* when environment dynamics are known.

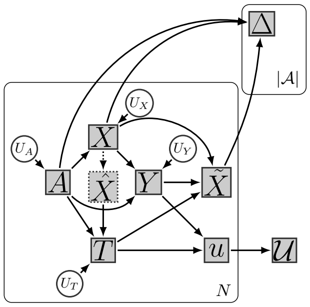
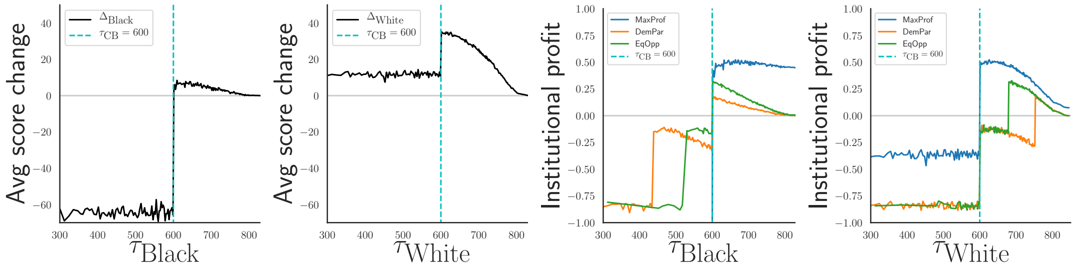


Figure 6: An extension of the lending (SCM) that emphasizes the role of the credit scoring bureau.

**Intervention by credit bureau** Liu et al. (2018) conduct experiments based on statistics of FICO credit scores assigned by the credit bureau TransUnion (Reserve, 2007). We note that these credit score decisions themselves constitute a policy; and moreover, the language of interventions in the SCM framework allows us to characterize decisions made by the *credit bureau* (rather than the bank) using the same fairness and profit metrics as before.[10] The credit bureau enters the SCM by reinterpreting $X_i$ as *features* related to creditworthiness of an individual, then introducing $\hat{X}_i = f_{\hat{X}}(X_i)$ as a *score* that is deterministically computed by the agency from the features (See Fig. 6). When $f_{\hat{X}}$ is the identity function, we recover the original model. Policy evaluation under double intervention $\mathcal{M}^{\mathrm{do}(f_T \to \hat{f}_T, f_{\hat{X}} \to \hat{f}_{\hat{X}})}$ captures the sensitivity of the bank's decisions to the decisions of the credit bureau (and vice versa).

Figure 7 shows the effect on the average utility $\mathbb{E}[\mathcal{U}]$ and average per-group score change $\mathbb{E}[\Delta_j]$ of a simple policy intervention by the credit bureau. The intervention involves the bureau setting the minimum score to 600 for all applicants via the structural equation $\hat{f}_{\hat{X}}(X) = \min(X, 600)$. This intervention is unlikely in the real world because it contradicts the profit incentives of the bureau, which encourage well-calibrated scores. Nevertheless, it coarsely captures a potential scenario where

[10] Note that recent changes by the credit scoring bureau Fair Isaac Corp. (`https://www.wsj.com/articles/fico-changes-could-lower-your-credit-score-11579780800`) can be characterized as such an intervention.



(a) Score change, min. group. (b) Score change, maj. group. (c) Profit as fn. of min. thresh. (d) Profit as fn. of maj. thresh.

Figure 7: Policy evaluation under credit bureau intervention $\hat{f}_{\hat{X}}(X) = \min(X, \tau_{CB})$ with $\tau_{CB} = 600$. Group score change—formally $\mathbb{E}_{p^{do(f_{\hat{X}} \to \hat{f}_{\hat{X}}, f_T \to \hat{f}_T)}}[\Delta_j]$—and institutional profits—$\mathbb{E}_{p^{do(f_{\hat{X}} \to \hat{f}_{\hat{X}}, f_T \to \hat{f}_T)}}[\mathcal{U}]$—are shown as functions of the two group thresholds $\{\tau_j\}$ under several fairness constraints.

an actor besides the bank seeks to encourage fair outcomes in a group-blind way, since under the new scoring policy minority applicants are more likely to receive loans. However, we see in Figure 7a that the average group outcome for Black applicants is negative when the bank's group threshold $\tau_{Black}$ is below 600, since in this case its policy offers loans to individuals who have good scores on paper but are unlikely to repay the loans. Interestingly, the expected profit (Figure 7b) under credit bureau intervention differs depending on the fairness criteria of the bank. This is because each fairness criteria differently constrains the relationship between the two thresholds $\{\tau_{Black}, \tau_{White}\}$, so the choice of fairness criteria implicitly sets how many applicants with boosted scores ($X < 600$, thus $\hat{X} = 600$) are selected for loans. DEMPAR is more sensitive to the credit bureau intervention than EQOPP; it obeys a stricter fairness constraint and offers more loans to applicants with boosted scores (who are are unlikely to repay, and disproportionately belong to the minority group).

**Causal model as computation graph** A causal DAG can be thought of an *expressive* simulator, whose capabilities extend the classical graphical model via the mechanism of *intervention*. Accordingly, the standard tools for optimization/learning in computation graphs (Schulman et al., 2015) can be brought to bear in order to *learn* policies that capture optimal rewards *across many interventional settings*. While there are several obstacles in applying standard gradient based learning to the lending setting, we found in unreported experiments that biased gradient estimators are capable of learning bureau policy that improves expected score change outcomes for the disadvantaged group by marginalizing out uncertainty about the bank's choice of fairness criterion. Such an approach holds promise in scaling to high dimensional datasets, which we leave for future work.

When both causal structure (which edges are present in graph) and dynamics (functional form of structural equations) are known, then can be jointly leveraged to produce *expressive simulators* capable of generating trajectories under many interventional distributions. This implies that a parameterized machine learning system interacting with a dynamic environment can adjust its parameters to optimize for short- or long-term fairness outcomes, even when there is uncertainty about which interventions will occur (for example those induced by other actors) at test time.