

Do You Trust the Process?: Trust-Contingent Reinforcement Learning for Resource Allocation

NAINA BALEPUR, University of Illinois Urbana-Champaign, USA

KESHAV SUBRAMONIAN, University of Illinois Urbana-Champaign, USA

HARI SUNDARAM, University of Illinois Urbana-Champaign, USA

As many governmental bodies increasingly turn to AI to assist with decision-making, it is essential that we explore the impacts of its use beyond the explicit goals of these policies. In this paper we focus on RL for algorithmic governance—the application of computational tools to policy domains. These policies often optimize for social welfare—like taxation to decrease economic inequity or border control to reduce the spread of COVID-19. However, we know that the use of AI in decision-making can directly decrease citizens’ trust in those decisions. Designing trustworthy AI means not only ensuring our tools are competent and fair, but also recognizing that the use of AI as a decision-making proxy can have a strong impact on the trusting relationships built between communities and institutions. In this work, we develop a trust-contingent RL modeling approach for resource allocation to communities by an institution. We use a two-stage Deep Deterministic Policy Gradient approach to learn a resource allocation that fulfills the needs of the community while staying within budget and encouraging trust between the community and institution. We consider three models that align with our approach: trust-optimizing, trust-blind, and trust-sensitive. Each of these models incorporate trust in the reward function and state space to differing degrees. We find that for each model, our proxy for institutional trust sees the largest increase in communities with low initial trust and low resource needs. In such cases, our trust-optimizing model produces an allocation that increases mean trust by 117% more than a uniform allocation, and 89% more than the allocation produced by our trust-blind model. When simulating the policy produced by our trust-optimizing model, we identify a tension between increasing trust and providing resources to meet community members’ needs. This tension, however, does not exist with the trust-sensitive model. We study further how the structure of the community network impacts outcomes, and find that trust is most consistently improved on star graphs. This work underscores the importance of institutional trust in algorithm design and implementation, and we propose a modeling approach that can be used in the future to keep trust at the forefront.

CCS Concepts: • **Computing methodologies** → **Reinforcement learning**; • **Human-centered computing** → *Social networks*; • **Social and professional topics** → **Computing / technology policy**.

Additional Key Words and Phrases: Algorithmic Governance, Resource Allocation, Institutional Trust

ACM Reference Format:

Naina Balepur, Keshav Subramonian, and Hari Sundaram. 2026. Do You Trust the Process?: Trust-Contingent Reinforcement Learning for Resource Allocation. In *The 2026 ACM Conference on Fairness, Accountability, and Transparency (FAccT '26)*, June 25–28, 2026, Montreal, QC, Canada. ACM, New York, NY, USA, 25 pages. <https://doi.org/10.1145/3805689.3812274>

1 Introduction

In recent years, reliance on algorithmic governance has increased dramatically [20, 30, 40, 44]. In particular, RL has been deployed to learn policies in high-stakes public contexts—from resource allocation during COVID-19

Authors’ Contact Information: Naina Balepur, nainab2@illinois.edu, University of Illinois Urbana-Champaign, USA; Keshav Subramonian, keshav7@illinois.edu, University of Illinois Urbana-Champaign, USA; Hari Sundaram, hs1@illinois.edu, University of Illinois Urbana-Champaign, USA.



This work is licensed under a Creative Commons Attribution 4.0 International License.

FAccT '26, Montreal, QC, Canada

© 2026 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-2596-8/2026/06

<https://doi.org/10.1145/3805689.3812274>

[8, 47, 50, 68, 84, 98] to tax policy design [62, 96]—where optimization objectives directly affect citizens’ lives. While these settings vary widely, they share a common structure: an institution allocates limited resources or interventions to a population with the goal of maximizing social welfare. In this work, we study a simplified but fundamental instance of this problem: resource allocation on social networks.

When AI has been utilized for governance in the past, it has caused two major issues (among others). First, algorithmic governance has harmed marginalized groups [5, 17, 90]. Second, and the focus of this work, is that the use of AI in decision-making can *decrease* citizens’ trust in those decisions. Ingrams et al. [39] find that citizens perceive AI decision-makers to be less competent, honest, and benevolent than their human counterparts. Here, we focus on how an RL policy implemented by an institution might impact *institutional trust*, which we define as the belief by citizens that institutions are acting according to the expectations of the public [63].

Understanding how institutional trust of citizens is fostered and dissolved is key in the area of policy design [3, 31, 45]. When this trust is fractured, it can affect citizen willingness to adopt technology, use services, or comply with policies, like social distancing during COVID-19 [16, 46, 53]. If the use of AI policies breeds mistrust [39], and this mistrust can directly affect citizens’ actions, the simulated outcomes of our policies cannot be accurate if we do not take institutional trust into account. This is especially salient as we experience a global decrease in institutional trust in Western governments [58]. It is our belief that designing trustworthy AI means not only ensuring our tools are competent and fair, but also recognizing that the use of AI as a decision-making proxy can have a strong impact on the trusting relationships built between communities and institutions. Much work in RL for governmental policy ignores this tension between the institution and citizen, but understanding and modeling how these algorithms are received is essential as we consider how communities react to AI decision-makers, including refusal to comply with policies.

In this work, we propose a trust-contingent modeling approach for RL which accounts for a proxy of citizens’ institutional trust when learning optimal policies for resource allocation. We consider a network of citizens alongside an institution which provides resources to these citizens. One such example institution is a humanitarian engineering organization (*e.g.*, Engineers Without Borders [11]) which provides engineering solutions to under-resourced communities. In our model, each community member has some prior trust in the organization (due to reputation and individual propensity to trust) [33, 61], as well as some desire for resources. We utilize a two-stage Deep Deterministic Policy Gradient (DDPG) framework in which the RL agent (equivalently, the institution or organization) allocates resources to adequately meet the needs of each community member, while staying within budget. The RL agent is first trained to minimize demand mismatch across all community members, and then refined in a multi-step environment that explicitly models each citizen’s trust evolution as a consequence of neighbor influence and the resources received [28]. We develop three RL models using this approach: trust-optimizing, trust-blind, and trust-sensitive. Our trust-optimizing model explicitly includes trust in the reward function; trust-blind does not observe trust dynamics in any way, as training is halted after the first stage; and trust-sensitive observes how trust dynamics have downstream effects on the reward, without directly optimizing for trust.

We find that both our trust-optimizing and trust-sensitive models succeed in improving our trust metric in communities over the trust-blind and uniform allocation baselines. We do identify a trade-off in the trust-optimizing policy between meeting resource needs and improving trust, though this tension does not exist for trust-sensitive policies. Our trust-sensitive model is a key contribution—rather than presuming that trust is an explicit organizational objective, this model isolates how trust dynamics can shape outcomes through downstream behavior. The network structure of the community also has an effect on outcomes; we find that networks with concentrated degree distributions (like cliques) do not lend themselves as well to an improvement in our trust metric as those with bimodal distributions (like star graphs). Through this work, we make two contributions:

Trust-Contingent RL Modeling Approach: We are the first to propose an RL approach that relies on the trust of the actors involved. In contrast, prior RL literature ignores the tension between AI decision-makers and the institutional trust of those affected by algorithms. For example, an RL solution to resource allocation during COVID-19 [98] does not account for the trust of citizens in their models. And while the public goods allocation problem in [48] is tackled using RL and based in a trust game, this work does not consider the trust of citizens in the institution, but rather citizens between each other. Not only does this disregard the impacts to citizens, but also leads to less informative policy simulations. Through three instantiations of our approach, we find that explicitly optimizing for our trust metric is an effective approach for an institution to preserve trust in a community. We hope that this approach can be used in the future to guide trust-contingent modeling in other areas of algorithmic governance.¹

Preventing Community Harms: We anticipate community harms and simulate an algorithmic solution. While lack of trust is partly fueled by the use of AI itself, RL policies can also cause harm to specific individuals [5, 17, 90]. Humanitarian engineering organizations in particular have caused harm and promoted neocolonialism [10], sometimes unintentionally. In this work we anticipate how failing to optimizing for trust directly might result in harm to the relationships between community members and institutions, through our trust-blind model. Further, we examine the impacts of a counterbalance through our trust-sensitive model that could be implemented with some organizational oversight. Understanding the harms that can come from policies is largely under-explored; to responsibly develop algorithms we must predict future harm and propose solutions.

Our results demonstrate the importance of considering institutional trust when designing RL policies. With sufficient data, these simulations can be used as predictive tools to estimate downstream effects of implementing RL policies in communities, perhaps in conjunction with proposed solutions to decrease inequitable algorithmic outcomes, like participatory design for production of fair and efficient public decision-making algorithms [21, 52, 85].

2 Related Work

RL for Policy: Reinforcement learning has been used in many sectors to design real policies. Yu et al. [95] survey uses of RL in healthcare, and Deliu [23] designs an RL algorithm to deliver personalized wellness-oriented behavioral recommendations. RL algorithms have also informed potential COVID-19 policies including border control, vaccination location setting, and resource allocation [8, 47, 50, 68, 84, 98]. Proposed solutions to issues in climate policy [81] and economic governance [1, 62, 96] also come from RL policies. Some work has considered the ethical implications of using RL to solve high-stakes problems [18, 29, 91], and others have studied the difficulties in implementing real-world solutions [25]. This work is motivated by the abundance of algorithmic governance driven by reinforcement learning.

Trustworthy AI: In RL applications [15, 35], trustworthiness is often used as a proxy for reliability or accuracy, rather than a personal, dynamic relationship. In this work, we model institutional trust being lost or gained by citizens as a direct effect of RL algorithms. Our formulation and attempt to quantify institutional trust in this way is novel.

In other domains, researchers have focused on human relationships with AI, and how trust is built between them. In theoretical work, Manzini et al. [59] investigate trust in AI assistants, which they study as a function of competence and alignment. Most recently, a large body of work has emerged on the trustworthiness of AI systems as they aid in human decision-making, especially through LLM's [19, 27, 42, 69, 80, 88]. Purves and Davis [71] find that AI opacity leads to loss of institutional trust from citizens. There has also been in-depth exploration of trust in the field of human-robot interaction [34, 51, 54, 92]; most relevant to our work, Guo and Yang [32] model and predict human trust in robots after they interact. However, across various domains, very little work

¹Code can be found at https://github.com/CrowdDynamicsLab/RL_Trust.

has modeled how algorithms affect institutional trust [6]. In particular, no work to date has considered the effect of citizen trust on the effectiveness of RL policy.

Modeling Institutional Trust: Institutional trust is the belief by citizens that institutions are acting according to the expectations of the public [63]. Work has been done in sociology and business management to understand which factors can alter citizens’ trust in institutions; in particular, Kaasa and Andriani [43] find that individuals tend to trust institutions less when there is a large power distance, meaning sense of participation and civic responsibility are important factors. PytlikZillig et al. [72] find experimentally that increased knowledge stabilizes institutional trust attitudes of community members. The COVID-19 pandemic spotlighted institutional trust; Jiang et al. [41] categorize citizens using unique trust profiles, and find correlations with compliance with COVID-19 protocols.

Though much empirical work has been done to understand institutional trust, it has never been formally modeled. Trust dynamics (though not institutional) on social networks have been formalized in prior work [49, 66, 77, 97]. These works focus on modeling and predicting trust between members of a community, rather than the trust of community members in an institution, which we explore in this work.

Fair Resource Allocations: Deciding how to fairly allocate resources has been thoroughly studied in the game theoretic space [12, 13, 38, 60]. One application area is disaster relief—Dönmez et al. [24] survey the humanitarian logistics space, covering many methods for fair resource allocation. Some work has developed resource allocation policies that can be adjusted to decision-makers needs [37], while other work presents several heuristic solutions for fair distribution of food rescue resources, simultaneously ensuring efficiency [65]. In our work we use RL as a tool to mitigate the challenges of allocating resources within the changing dynamics of the environment, as has been done in prior work in other contexts [86, 87, 93].

3 Problem Statement

Consider a community network $G = (V, E)$ with community members $v \in V$. The set of undirected edges E represents the set of active relationships maintained among individuals, along which information can flow. Now, consider an institution (or organization) \mathcal{I} , which is not part of the network, but is able to observe it. The exact information that can be observed by \mathcal{I} is model dependent, and summarized in Table 1. The goal of the institution is to provide resources to the community, which come from the institution’s budget, ρ . Each community member $v \in V$ has some initial want or need for resources, ω_v , which can change as each citizen’s expectations of the institution change as well. Each individual $v \in V$ is also initialized with a baseline institutional trust value of τ_v , which will be updated as the organization and community interact.

The goal of institution \mathcal{I} is to provide some resources (like solar panels or cook stoves, in the case of a humanitarian engineering organization) to community members V and distribute them according to their needs [74]. We utilize an RL agent to learn the optimal policy to allocate resources to maximize the agent’s reward function, \mathcal{R} , which too depends on the model being used (see Section 4). Reward function \mathcal{R} may depend on the how well each citizen’s needs were met, any additional budget gained by the institution, and the trust of the community members. The learned resource allocation policy results in an allocation $a = (a_1, a_2, \dots, a_{|V|})$ for each timestep, such that $\sum_{i=1}^{|V|} a_i \leq \rho$. Once this policy is learned, we simulate the distribution of these resources to the community. Each timestep, for each individual, trust value τ_v and need ω_v are updated to account for the newest allocation. Further details regarding trust dynamics can be found in Section 4.4.3. In this work, we study institutional trust as a social phenomenon, and investigate how varied resource allocations, especially from AI decision-makers, can foster or dissolve the institutional trust of community members. We seek to answer the following questions:

RQ1: Trust-Optimizing Model Suppose institution \mathcal{I} desires a trusting relationship with community members $v \in V$ (i.e., trust is explicitly optimized for). Can institution \mathcal{I} learn the optimal policy to allocate resources while minimally eroding trust? How does this model compare to one where the organization is trust-blind?

RQ2: Trust-Sensitive Model Suppose institutional trust is not an explicit organizational objective, but rather has downstream consequences for institutional reward. Can \mathcal{I} use a trust-sensitive model to learn a policy that minimally erodes trust?

RQ3: Network Effects How do characteristics of the network $G = (V, E)$ impact final trust values and dynamics?

4 Trust-Contingent Reinforcement Learning

This section presents the models used by institution \mathcal{I} to learn resource allocation policies. We first describe the dataset creation process and an overview of the modeling framework, followed by a detailed description of the two-stage learning procedure. Our primary model is trust-optimizing, and we focus on this model when describing Stages 1 and 2 in this section. The trust-blind model can be viewed as an ablation of the trust-optimizing model that considers only Stage 1. We next detail how the citizen attributes of institutional trust and resource needs are impacted by resource allocation decisions. We conclude by describing the modifications to Stages 1 and 2 that create our trust-sensitive model, in which trust values are not directly observed but cause downstream consequences for behavior and thus reward.

4.1 Dataset Creation

We synthesize 202 networks (160 for training and 42 for testing) with 10 nodes each, using the process described below.

4.1.1 Network Formation Process. We create networks according to various classic network formation models. Our aim in choosing these is to capture a range of structures and degree distributions. We produce a set consisting of the following network classes: Erdős–Rényi (ER) [26], Barabási-Albert (BA) [2], Watts-Strogatz (WS) [89], Stochastic Block Model (SBM) [36], Star Graph, 10-Clique, and two disjoint 5-Cliques.

4.1.2 Citizen Attributes. Each citizen (or node) $v \in V$ is initialized with mutable values for institutional trust and resource need. We describe how these values are assigned in Section 5.

Institutional trust τ_v : Each community member is assigned a scalar [57] institutional trust value $\tau_v \in [0, 1]$. In alignment with the definition, τ_v is the belief that institution \mathcal{I} will provide resources according each citizen v 's needs.

Resource need ω_v : Each community member is assigned a scalar resource need $\omega_v \in [0, 1]$. This need represents the quantity of resources the individual desires in order to sustain themselves each time step. The institution \mathcal{I} observes this need and distributes resources accordingly.

4.2 Overview

We formulate the resource allocation problem as a Markov Decision Process (MDP) defined by the tuple $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$. At each timestep t , the RL agent observes a state $s_t \in \mathcal{S}$ that encapsulates the current needs and network features of the community. Based on this state, the agent selects an allocation action $a_t \in \mathcal{A}$, and then transitions into a new state s_{t+1} according to the transition dynamics $\mathcal{P}(s_{t+1} | s_t, a_t)$, which govern how citizen trust and resource needs evolve in response to the allocation. The agent receives a scalar reward $r_t = \mathcal{R}(s_t, a_t)$ reflecting how well the allocation aligned with the institution's objectives. The goal of the RL agent is to learn a policy $\pi(a_t | s_t)$ that maximizes the expected cumulative discounted reward, $\mathbb{E}[\sum_{t=0}^T \gamma^t r_t]$ where $\gamma \in [0, 1)$ is the discount factor.

DDPG [56] is a model-free, off-policy, actor-critic method that employs two neural networks: the actor, which generates actions, and the critic, which evaluates these actions through a Q-function, $Q(s, a)$ [82]. The Q-function represents the expected cumulative discounted reward of taking action a in state s , and following the actor’s policy thereafter. Due to its ability to learn complex policies in dynamic environments, DDPG is ideal for learning resource allocations on networks [82]. Specifically, our problem requires a deterministic, continuous action space—the agent must output a precise fractional distribution per node with no room for stochastic variance. Additionally, as an off-policy algorithm, DDPG’s replay buffer allows for efficient reuse of past experience. The focus of this paper is not on the detailed description or development of the RL architecture, but rather on leveraging DDPG in a simple problem setting to produce insights on trust. Those interested in technical DDPG details can read further [56, 82, 83].

Because we evaluate three distinct operational paradigms, the formulations of state space \mathcal{S} and reward function \mathcal{R} vary across our models. We detail the foundational state, action, and reward formulations in Section 4.4, and outline the structural modifications for the trust-sensitive model in Section 4.5. We provide a high-level overview of the observations for each model in Table 1.

Table 1. Our model variants and the agent observation for each.

THE MODEL	DESCRIPTION	RL AGENT OBSERVATION
Trust-optimizing	To answer RQ1 , trust is explicitly included in the reward function	Resource needs ω of citizens, social network structure G , and trust values τ
Trust-blind	To answer RQ1 , trust dynamics are not observed or accounted for in any way	Resource needs ω of citizens
Trust-sensitive	To answer RQ2 , trust dynamics have downstream effects on reward without directly optimizing for trust	Resource needs ω of citizens, and social network structure G

4.3 Stage 1: One-Step Fair Allocation (Trust-Blind Model)

Stage 1 of our trust-contingent RL modeling approach serves as a warm-up procedure which minimizes noise in Stage 2. The goal is to learn a static allocator that minimizes shortfall, or unmet need, only being given needs of citizens, ω .

In a single time step, the agent observes a vector of needs, $\omega = (\omega_1, \omega_2, \dots, \omega_{|V|})$ and chooses an allocation $a = (a_1, a_2, \dots, a_{|V|})$ such that $\sum_{i=1}^{|V|} a_i = \rho$, enforcing the budget. We define the *shortfall* at node v as $s_v = \max(\omega_v - a_v, 0)$, and *oversupply*, or excess allocation, as $x_v = \max(a_v - \omega_v, 0)$. Then, we define the Stage 1 loss as the convex penalty $L_{short}(a, \omega) = \sum_{i=1}^{|V|} (\lambda_{short} s_i^2 + \lambda_{over} x_i^2)$. Hyperparameters λ_{short} and λ_{over} weight the influence of shortfall and oversupply. To strongly discourage oversupply and encourage redistribution to other nodes, we set $\lambda_{short} = 2.0$ and $\lambda_{over} = 4.0$.

Recall that this static allocator encompasses our trust-blind model. The agent only attempts to match the needs of each node while remaining within budget. This optimization is in line with need-based justice or fairness [67] though other fairness formulations could be considered in the future, which we discuss in Section 6. For the trust-optimizing model, Stage 1 gives the agent a sensible direction for resource allocation, and a stable initialization for the actor network, which reduces the probability of vanishing or exploding gradients in the actor-critic network.

4.4 Stage 2: Trust-Optimizing Allocation on Networks

Building on pre-trained weights from Stage 1, in Stage 2, the agent learns the final allocator in a trust-aware environment, which is defined on the social network G .

4.4.1 State Space (\mathcal{S}_{TO}). At time step t , the *true environment state* consists of: resource needs ω^t of citizens, trust levels τ^t of citizens, and the social network $G = (V, E)$ which is unchanging. The *state space* for the agent is a tuple consisting of: resource needs ω^t , and normalized degree vector \tilde{d} , such that all $d \in \tilde{d}$ are between 0 and 1.

4.4.2 Action Space (\mathcal{A}_{TO}). As in Stage 1, the agent outputs an allocation a^t , such that for each time step t , $\sum_{i=1}^{|V|} a_i^t = \rho$.

4.4.3 Trust Dynamics (\mathcal{P}_{TO}). Given an allocation a^t at time step t , the satisfaction of node v is given as:

$$\sigma_v^t = \left(\frac{\min(a_v^t, \omega_v^t)}{\omega_v^t} \right)^p.$$

This is the fraction of resource need fulfilled, raised to the p -power. We set $p = 1.3$ to increase the penalty when the served fraction is far from 1. After time step t , both satisfaction and social diffusion affect the trust τ of every node at time step $t + 1$. Trust of node v is updated as an exponential moving average:

$$\tilde{\tau}_v^{t+1} = \tau_v^t + \eta(\sigma_v^t - \tau_v^t).$$

Parameter η determines the update speed. Recall that institutional trust is a proxy for each node's expectation of the institution \mathcal{I} . As each node's resource needs are met or unmet, the expectation will change accordingly. These dynamics capture slow erosion of trust under poor service, and gradual accumulation of trust under consistent satisfaction. Trust also diffuses through the network. Recall that edges E of G represent active relationships between nodes in V along which information may flow. Thus, as nodes interact with their neighbors on network G , trust will diffuse:

$$\bar{\tau}_v^{t+1} = \frac{1}{|N(v)|} \sum_{i \in N(v)} \tilde{\tau}_i^{t+1}.$$

We define $N(v)$ as the set of v 's neighbors. Combining these two values, the final institutional trust value of node v at time step $t + 1$ is given as:

$$\tau_i^{t+1} = (1 - \delta)\tilde{\tau}_v^{t+1} + \delta\bar{\tau}_v^{t+1}.$$

Parameter δ weighs the relative importance of satisfaction and social influence on trust. Our trust dynamics are motivated by prior research. In particular, Cabiddu et al. [14] find: 1) initial trust in an algorithm is a function of a user's propensity to trust, and 2) if a user perceives that an algorithm has positive social influence, there is a higher likelihood of trust building. In addition, Mayer et al. [61] and Hancock et al. [33] both emphasize user propensity to trust, as well as reputation of trustee, as important factors when considering overall trust between two individuals. We incorporate these properties into our model; the original trust of each individual τ_v is assigned at random, we take this to be each individuals' propensity to trust. In addition, we split the "positive social influence" observed by Cabiddu et al. [14] into two components: individual satisfaction and trust (and by proxy, satisfaction) of neighbors. In our models, we choose specific values for our coefficients η and δ , which we discuss in Section 5.

4.4.4 Resource Need Dynamics (\mathcal{P}_{TO} contd.) Once allocation a^t is provided, and trust values τ^{t+1} are updated, resource needs of node v also updates according to:

$$\omega_v^{t+1} = \max(\omega_v^t - a_v^t, 0) + \alpha\sigma_v^t.$$

Parameter α scales how resource need evolves over time; in our implementation, we set $\alpha = 0.2$. Intuitively, nodes request a bit more if their needs were met entirely, and exactly the same amount if they didn't receive any resources. Otherwise, new resource requests not only encompass any unmet needs, but also grow proportionally with satisfaction from the previous time step. Therefore, unmet resource needs persist, while satisfied needs partially recover this. Given these dynamics, resource need ω_v^t remains between α and 1 for each node v and every time step.

4.4.5 *Reward Function (\mathcal{R}_{TO})*. The reward function is a scalarized objective:

$$\mathcal{R}_{TO}^t = D \sum_{i=1}^{|V|} \tilde{d}_i (\tau_i^{t+1} - \tau_i^t) - \frac{L_{short}(a, \omega)}{|V|}.$$

Recall that \tilde{d}_v is the normalized degree of node v , $\tau_v^{t+1} - \tau_v^t$ is the trust gained by node v in time step t , and $L_{short}(a, \omega)$ is the penalty from Stage 1. Parameter D determines the relative strength of the prior allocator from Stage 1, compared to the new allocator in Stage 2, and we discuss this choice in Section 5. We weight trust gains by the relative degree of the node because hub-targeting behavior does not always emerge naturally. As we know, opinion dynamics on networks can be driven by high degree nodes [75], so we provide this prior to encourage prioritization of these nodes.

4.5 Trust-Sensitive Allocation on Networks

In the previous subsections we focused the trust-optimizing model. Now, we discuss the modifications made to Stages 1 and 2 to produce our trust-sensitive model. Now, rather than institution \mathcal{I} explicitly optimizing for trust, an external mechanism alters the institution's budget ρ as institutional trust levels change. This might reflect waning external funding from an external agency, or lack of donations from the community directly as trust decreases.

The distinction between the trust-sensitive and trust-optimizing models is intended to be structural and conceptual, rather than mathematical, as our changes result in an optimization objective that includes trust by proxy. In this framing, the institution does not intrinsically value trust, as its objective is strictly to satisfy community needs. It is forced to account for trust only because of downstream financial consequences. Here, we contrast an institution actively trying to build trust with one that is externally motivated by its environment. To formalize this conceptual shift within our MDP framework, the specific components for the trust-sensitive model are modified as follows:

4.5.1 *State Space (\mathcal{S}_{TS})*. The state space is identical to \mathcal{S}_{TO} , concatenated with budget ρ^t .

4.5.2 *Action Space (\mathcal{A}_{TS})*. The action space is also identical to \mathcal{A}_{TO} , though ρ varies between time steps.

4.5.3 *Transition Dynamics (\mathcal{P}_{TS})*. The underlying social dynamics governing citizen trust and resource needs remain identical to those described in Section 4.4.3 and Section 4.4.4. However, the environment features an additional transition rule for the budget ρ^t , which evolves based on community trust. As Stage 2 begins, the budget is initialized as ρ , and we index ρ by t to track its evolution. The new budget ρ^{t+1} updates according to:

$$\rho^{t+1} = \rho^t + r + D \sum_{i=1}^{|V|} (\tilde{d}_i) (\tau_i^{t+1} - \tau_i^t).$$

We enforce that the budget must remain between 0 and 5. We use $r = 0.05$ as a small boost to the institution's budget each time step to balance strongly negative trust gains from the limited budget in initial time steps. The third term is identical to the weighted reward from trust gain in the trust-optimal model.

4.5.4 *Reward* (\mathcal{R}_{TS}). In Stage 1, since we are operating on a single time step, the budget is a constant ρ , and the reward solely targets the shortfall penalty. In Stage 2, the reward function explicitly incorporates the new budget per node as reward to the institution \mathcal{I} , alongside the penalty from Stage 1:

$$\mathcal{R}_{TS}^t = \frac{\rho^{t+1}}{|V|} - \frac{L_{short}(a, \omega)}{|V|}.$$

5 Simulation Results

In this section we present the results from simulating resource allocations as learned by our RL models. As a reminder, our dataset consists of social networks $G = (V, E)$, and for each $v \in V$ a trust value τ_v and resource need ω_v are both drawn from independent distributions. With this data, the institution \mathcal{I} learns a policy that dictates how resources should be distributed in the community. Then, we simulate the distribution of resources, and analyze several metrics to determine the organization’s success, as well as the impact on the community members.

Before presenting results, we describe our simulation parameters. We consider networks of 10 nodes; we keep our numbers small for simplicity and interpretability. The size of the network could be extended with some hyperparameter tuning. We simulate 7 time steps, *i.e.*, the number of rounds the institution provides resources to the community. We set the institution \mathcal{I} ’s budget to $\rho = 1$; this can be easily extended as the sum of the allocation will scale to the total resource constraint. We fix other parameters for our simulation as well: η , which determines the update speed of trust in the community is set to 0.5; δ , which weights the relative importance of social influence and satisfaction on trust is set to 0.35; and D , which determines the relative strength Stage 2 compared to the prior allocator of Stage 1 is set to 8. Because no prior research provides ecologically valid guidance for these parameter values, we chose them following extensive exploratory analysis, guided by our modeling priorities. As for DDPG hyperparameter tuning: we have discount factor $\gamma = 0.9$, actor learning rate $a_\pi = 10^{-4}$, critic learning rate $c_\pi = 10^{-3}$, and batch size = 64. A comprehensive list of hyperparameters, full model architecture details, and reward convergence curves during training are provided in Appendix A.

In the following subsections, we vary two values: the distributions of initial institutional trust values and resource needs across community members. We draw values for both of these citizen attributes from the following distributions: {Beta(2, 6), Beta(2, 4), Beta(2, 2), Beta(4, 2), Beta(6, 2)}. We train each model on a training set of 160 networks. We then run each testing simulation twenty times for each of the seven graph families we consider (see Section 4.1.1). We present the average of those simulations for all research questions.

5.1 RQ1: Trust-Optimizing Model

In **RQ1** we suppose institution \mathcal{I} desires a trusting relationship with community members $v \in V$. We ask whether the institution can learn a policy to allocate resources while minimally eroding trust, and compare this model to one that is trust-blind. We use a uniform allocation of resources to all community members as a baseline. For a visual depiction of the resource allocation process, refer to Figure 1.

We first explore the raw difference in mean institutional trust values after the resource allocation has been completed according to each policy. In Figure 2 we show these results as heatmaps, showcasing how results change as initial trust and resource needs vary. As expected, in each type of community, the trust-optimizing policy is most successful in increasing institutional trust across community members (Figure 2a). When initial trust is low (drawn from distributions Beta(2, 6) and Beta(2, 4)), the trust-optimizing policy is able to make the biggest difference in institutional trust, especially when resource needs of community members are low as well. In this case, our trust-optimizing model produces an allocation that increases mean trust by 117% more than a uniform allocation, and 89% more than the allocation produced by our trust-blind model. Returning to our definition of institutional trust, citizens have high trust in institution \mathcal{I} when it is acting according to

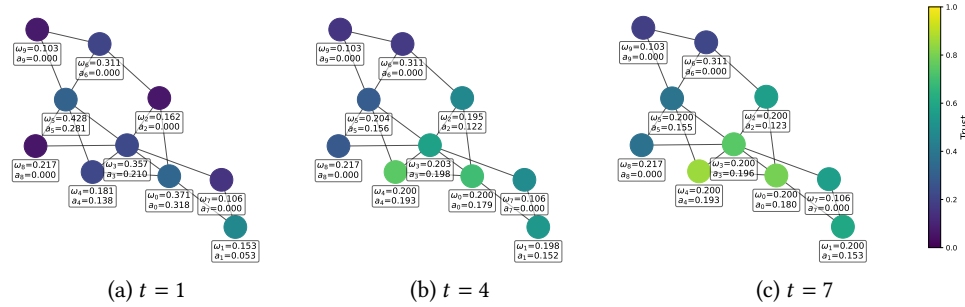


Fig. 1. Depiction of the trust-optimizing resource allocation process by institution \mathcal{I} to community G across 7 time steps. Each node is labeled; ω_v gives the resource need for node v , and a_v represents the allocation given by institution \mathcal{I} at the given time step. We draw τ_v and ω_v for each node v independently from $\text{Beta}(2, 6)$. Node color depicts trust at each time step, a legend is shown to the right. We show only time steps 1, 4, and 7 for space.

their expectations. When resource needs are low, it is more likely that the institution will be able to meet those expectations, thus increasing trust. When trust is already high, the institution must work very hard to allocate resources according to the expectations of the public, which may not be possible with budget ρ . However, we show here that our trust-optimizing resource allocation policy prevents major loss in trust that occurs in high need and trust communities with a trust-blind or uniform resource allocation policy (Figure 2b and c).

Next, let’s examine how well each of these policies are able to meet the needs of community members—we want to improve trust in communities without losing the ability of the institution \mathcal{I} to provide resources to the community. In Figure 3, we compute, accumulated over all time steps, $\|a - \omega\|_1^2$. Essentially, the square of the sum of under-supplied and over-supplied resources across all nodes. Here, we see that using the trust-optimizing policy for resource allocation (Figure 3a) does have consequences when it comes to meeting the needs of the community, though minor. As expected, the trust-blind policy (Figure 3b) is most effective here, as the only factor in the optimization is the resource needs of the community. The loss differential between the trust-optimizing and trust-blind policy is at worst, 10.1%. Through this metric we uncover a tension between increasing trust and providing adequate services to the community when considering our trust-optimizing model. We do not

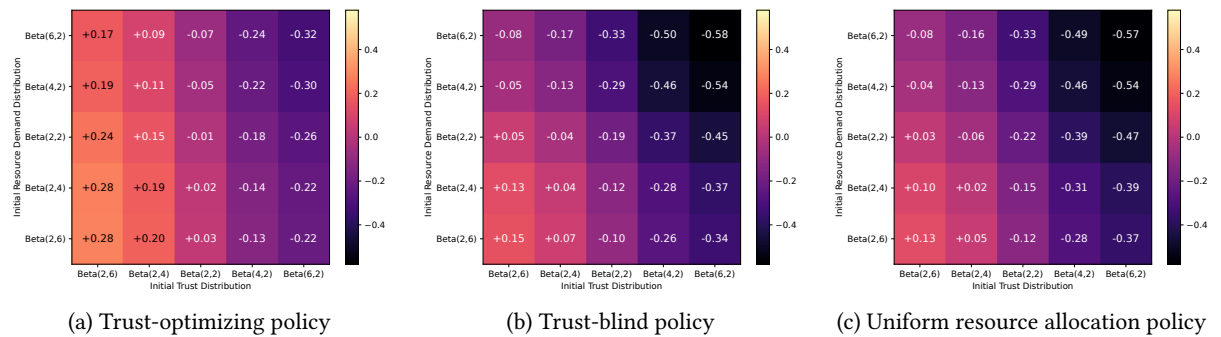


Fig. 2. Raw difference in trust after resource allocation, for trust-optimizing, trust-blind, and uniform resource allocation policies. Low trust and low need communities stand to benefit most from trust-optimizing policies, as they present the largest opportunities for building trusting relationships.

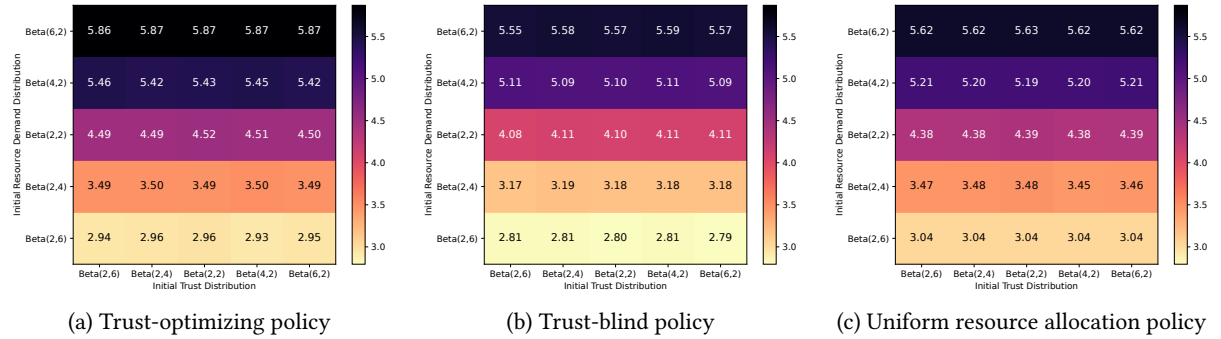


Fig. 3. Loss from misalignment with citizen needs, for trust-optimizing, trust-blind, and uniform resource allocation policies. The trust-blind policy is most effective at accurately meeting the needs of citizens. At worst, the trust-optimizing policy’s loss is 10.1% higher than trust-blind.

make any claims regarding which should be prioritized, but model designers can take note of this tradeoff as they consider their preferred outcomes—neither metric alone tells the entire story regarding community outcomes. Thus, in response to RQ1, we answer: yes, institution \mathcal{I} is able to learn a policy which prioritizes trust, with a slight decrease in effectiveness of resource allocation.

5.2 RQ2: Trust-Sensitive Model

In RQ2 we suppose institution \mathcal{I} does not explicitly desire a trusting relationship with community members $v \in V$. Rather, institutional trust has downstream consequences put in place either by the community or an overseeing organization, and thus affects institutional reward, the dynamics of which we discussed in Section 4.5. We compare the policy produced by our trust-sensitive model to a uniform allocation of resources to all community members as a baseline. We also present the results of our trust-sensitive model after only training Stage 1—this is similar to the trust-blind model, but with a dynamic budget to allow for direct comparison to our trust-sensitive model.

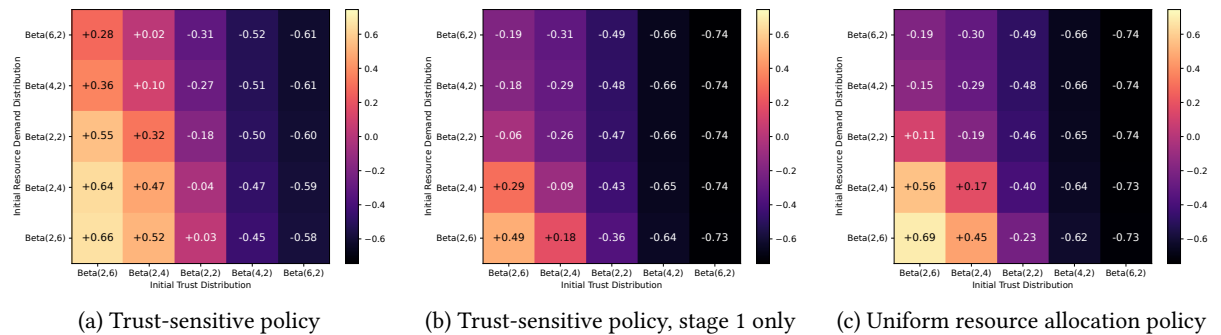


Fig. 4. Raw difference in trust after resource allocation, for trust-sensitive, trust-sensitive (stage 1 only) and uniform resource allocation policies. In this case, low trust and high resource need communities benefit most from the allocation learned by trust-sensitive policy. The dynamic budget implemented in this model allows for institutions to fulfill larger needs when trust is improving.

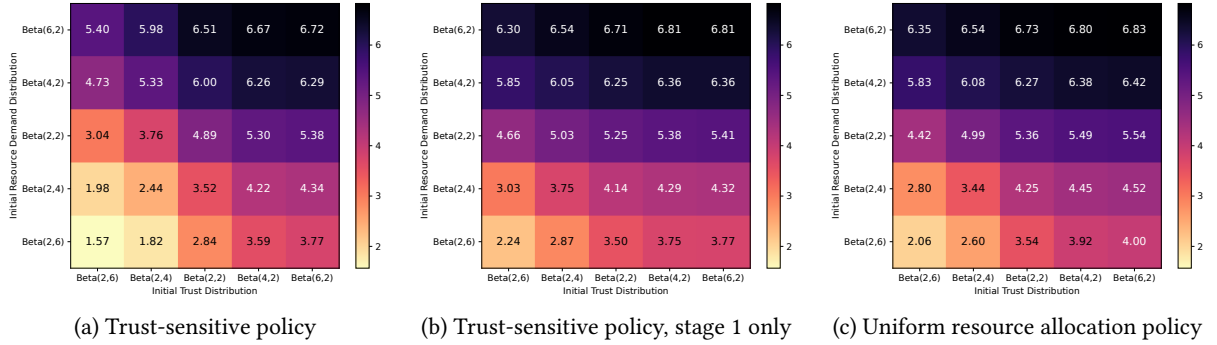


Fig. 5. Loss from misalignment with citizen needs, for trust-sensitive, trust-sensitive (stage 1 only) and uniform resource allocation policies. In almost every case, the trust-sensitive allocation best matches the resource needs of the community members.

As with RQ1, we first review the raw difference in mean institutional trust values after the resource allocation has been completed according to each policy. In Figure 4 we show these results as heatmaps. The trust gains and losses under these policies span a wider range of values than those examined in RQ1, driven by the dynamic budget ρ , which expands the range of the RL agent’s possible allocations. Looking at our primary policy of interest, trust-sensitive (Figure 4a), we can see that the allocation produced by this policy is the most successful at increasing trust values in the community in most cases, except in the lowest resource need and initial trust communities. For low trust and resource need (both drawn from Beta(2, 6)), the uniform resource allocation (Figure 4c) outperforms trust-sensitive. We hypothesize that this is the case because the low-want communities will easily be satisfied (or even over-supplied) by a uniform resource allocation policy when ρ is high. In every other case, however, the two-stage trust-sensitive policy outperforms the single-stage trust-sensitive and uniform policy—especially in low trust, high resource-need communities. According to these results, the dynamics in this model could realistically be used to provide a counterbalance to institutions that don’t wish to promote institutional trust in the communities they serve. We found that these consequences are sufficient for an institution to learn an allocation that promotes trust.

Next, we once more examine the ability of these policies to meet community members’ needs (Figure 5). In this case, it is clear that our trust-sensitive policy (Figure 5a) is more effective at meeting the needs of community members, when compared to both the single-stage trust-sensitive and uniform resource allocation policies (Figure 5b and c). We find these results to be especially promising, as the tension between trust and need-based resource allocation we noted with the trust-optimizing model, is no longer present. Thus, to answer RQ2: yes, this trust-sensitive model can learn a policy that minimally erodes trust, and also produced a needs-based allocation when compared to baseline models.

5.3 RQ3: Network Effects

In RQ3 we ask how the characteristics of community network $G = (V, E)$ impact trust outcomes. We use our trust-sensitive model to present these results, and consider at each of the seven graph families we described in Section 4.1.1.

Looking at Figure 6, our trust-sensitive policy performs similarly on most graphs, with the exception of the star graph, a network where one central node serves as the hub, with degree $|V| - 1$, and surrounding nodes each have degree 1, connected only to the hub. Barabási-Albert, Watts-Strogatz and the Stochastic Block Model all produce graphs that are meant to capture some aspects of real social networks, Erdős–Rényi is often used as a

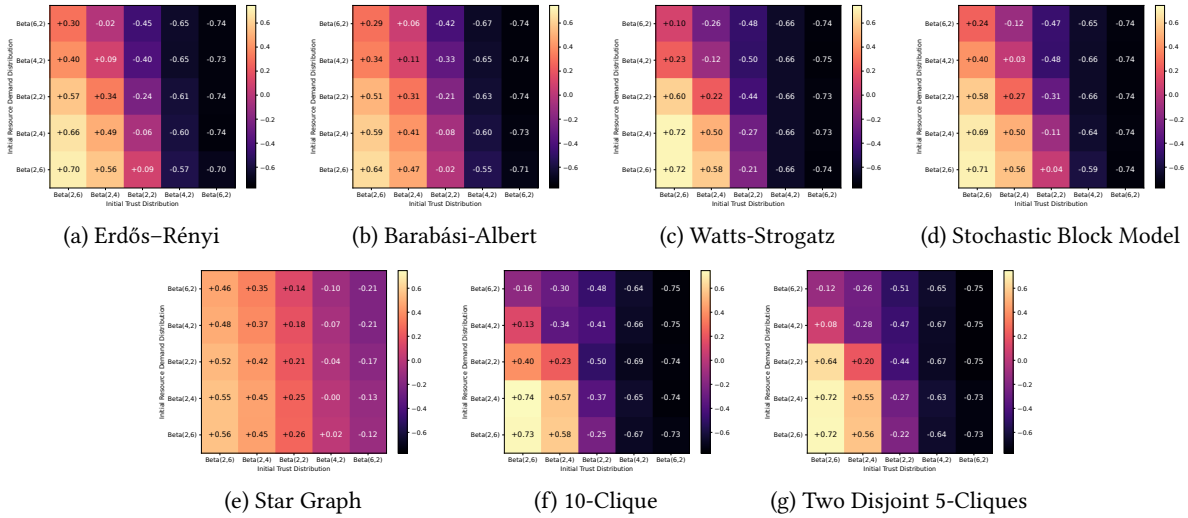


Fig. 6. Raw difference in trust after resource allocation, for trust-sensitive policy, on each class of community network. Institutional trust is most consistently improved on star graphs, where one node has degree $|V| - 1$, and all other nodes have degree 1. The trust-sensitive policy performs well on all networks with low trust and low resource needs, but the clique structures experience the steepest drops in trust improvements as we move toward higher trust or resource needs.

random baseline, and the cliques are meant to represent very densely connected communities. We hypothesize that trust-sensitive resource allocation is most effective on star graphs (Figure 6e) because the single central node can be targeted by the institution \mathcal{I} , even when resource needs are very high. This then leads to higher trust in the whole community, since the central node communicates with every other node. In contrast, the clique graphs (Figure 6f and g) see a steep drop-off in trust improvements as we move away from low trust and resource need communities. Again, we hypothesize that the institution \mathcal{I} is unable to provide resources to every node in high need scenarios, and this leads to massive decreases in trust because most (if not all) nodes are in communication with each other. The other graph families (Figure 6a-d) produce very similar results, though Barabási-Albert (Figure 6b) comes close to the results of the star graph due to the degree inequality between nodes. Thus, we conclude that network structure has an effect on trust outcomes, and that this is largely due to degree distributions. We observe that our trust-sensitive resource allocation policy performs worse on communities with very narrow degree distributions, and better on those that are bimodal, or have a larger spread.

6 Discussion

6.1 Formalization of Institutional Trust

In this work, we explored the incorporation of institutional trust into a well-studied modeling paradigm. As a complex human concept, trust is impossible to model entirely, and we acknowledge the “formalism trap” whereby our formalization of institutional trust inevitably fails to account for the full meaning of the term [79]. Rather than claim a full formalization of the term, we limit the scope of our inquiry, formalizing institutional trust as a proxy for communities’ expectations of institutions, impacted by prior interactions with the institution and discussion with neighbors.

We allow for flexibility in the modeling approach according to the modeler’s own preferences and priorities. In particular, we make several assumptions and parameters choices that could be changed in future work. First,

we chose to model institutional trust as a scalar [57], though a vectorized representation could allow for more specificity on different dimensions. And though the modeling choices we made were grounded in prior research [9, 14, 28, 33, 55, 61, 73, 76, 94], modifications to our proposed trust dynamics could be made to fit different context or priorities.

Understanding human decision-making is difficult, especially when it comes to complex constructs like trust. Though they are often used, surveys may not be accurate tools for capturing attitudes regarding trust [64], and recent work on human-AI trust [78] has found current surveys inadequate for capturing the full range of trust and distrust. To design such an instrument is beyond the scope of this work, but a qualitative complement to this work is a necessary next step for model validation and deployment. In the future, we plan to conduct interviews with human subjects to better understand which values impact trust, and to what extent—informing our trust update rule.

6.2 Misabstraction Risks

An issue that follows closely from the formalism trap is the risk of misabstraction. De Troya et al. [22] define misabstraction as the representation of complex, contextual phenomena in ways that cause harm upon reintegration into the sociotechnical contexts they were designed for. Misabstraction is a risk in much of computational social science research; in work on decision-making algorithms for policy, this often comes in the form of misinterpretation of results by policymakers, who may take abstract concepts represented in research and apply them directly to policy settings. We take the risk of misabstraction seriously; we engage with this risk directly and emphasize that though we colloquially name “trust” in this paper, we are referring only to a proxy for trust, which as we described in the previous subsection, is only one formalization of many possibilities. While we report that our trust-optimizing model outperforms baselines on increasing institutional trust, we remind the reader that this value is a proxy. We make this difference clear, while emphasizing that such models for algorithmic decision-making are inherently reliant on abstraction of human concepts. Our formalization is only one choice of many, which should be interrogated and modified for every new context.

Misabstraction in this context may cause negative consequences to communities, especially when policies are enacted by institutions who do not prioritize their well-being. We recognize that if an RL agent directly optimizes for network-wide trust, it could learn manipulative behaviors, such as ignoring isolated nodes in favor of influential ones. We address this directly in the model by jointly optimizing for trust and node shortfall, which is our proxy for fairness (see Section 4.4.5). However, negative consequences, such as under-resourcing certain individuals, could still arise. Here, we note that improving trust is not the only metric by which we should be evaluating our policies. We integrate fairness into our work, but looking further at accountability and transparency on the part of the institution is an important direction for future work and for any real-world deployment of this model.

6.3 More Trustworthy Algorithmic Governance

This work is predicated on the knowledge that algorithmic governance exists, and the belief that it will only become more prevalent as AI systems take on greater decision-making roles. Our aim is not to argue for or against algorithmic governance, but to identify one underexplored gap. We demonstrate, through the comparison of policies produced by our trust-blind and trust-optimizing models, that human decision-makers who design AI algorithms can control trustworthy outcomes. We acknowledge that exclusion of trust dynamics is not the most pressing issue with AI governance as it stands, but we hope that through this work we are able to make a small improvement.

7 Future Work and Limitations

Alternate trust-sensitive models: In this work, we examined how trust could have downstream consequences for institution \mathcal{I} through our trust-sensitive model. We chose to implement these consequences in the form of a dynamic budget ρ , which might represent changing funding from an overseeing organization, or even support from a local donor base. However, other consequences of lost trust could be modeled and explored in the future, such as citizens who no longer interact with the institution, or even leave the community. This would make for interesting future work, as the tradeoff between fostering trust and meeting resource needs may differ depending on the implementation.

Need-based fairness: Our modeling approach is predicated on a fair resource allocation minimizing the loss between resource needs and allocations. However, there are other normative allocation principles that could be considered here to motivate the resource distribution process, such as group [7], merit [4], and sufficiency-based [70] fairness framings.

Generalizability of findings: Further exploration on this topic, in different contexts, would be required to generalize the results in this paper to contexts beyond resource allocation. Our aim was to choose a simple problem setting to clearly identify the impacts of trust on RL, which is a deliberate scoping choice on our part. However, we do believe this issue of trustworthy governance could be extended to other areas. One possible extension is to revisit pandemic response policy, while accounting for and modeling untrusting community members who may not follow mandates.

Simulated data: Our work relies on simulated data and random assignment of citizen attributes, such as trust and resource needs. This is a limitation of our work, though our simulated trust dynamics are designed carefully and grounded in prior research [9, 14, 28, 33, 55, 61, 73, 76, 94]. In an ideal setting, there would exist network datasets that describe the institutional trust of each node, with an institution that provides the community with some service. Unfortunately, as far as we are aware, one does not exist. We plan to collect such a dataset ourselves in the future, working with local organizations in order to interrogate the assumptions made in this paper. Since the current work is an exploration of trust and how it affects RL algorithms, we find that aspect outside the scope of this work.

8 Conclusion

In this work we proposed a trust-contingent reinforcement learning modeling approach for resource allocation in communities by an institution. Using a two-stage DDPG framework for three model variants, we learned policies for resource allocation that fulfilled the needs of the community while staying within budget. Our trust-optimizing model explicitly considered trust in the reward function, and succeeded over the trust-blind and uniform resource allocation policies in increasing trust. However, we identified a tradeoff here between improving trust and allocating resources according to the needs of the community. Our trust sensitive model is a key contribution of this work—not only does it model realistic institutional checks, but also produces allocations that improve trust while meeting citizens' needs. We found that network structure also impacted the effectiveness of our models, and networks with highly concentrated degree distributions, like cliques, saw smaller trust improvements than star graphs which contain one hub. In future work, we plan to conduct interviews to interrogate assumptions we made in this model, especially with respect to data synthesis and trust dynamics. In this work we explored one aspect of designing trustworthy AI. We sought to understand, through modeling and simulation, how using AI for decision-making can have direct consequences on the trust of communities. Our work here underscored the importance of institutional trust in algorithm design and implementation, and we hope that the proposed modeling approach can be extended by others.

Generative AI Usage Statement

ChatGPT (GPT 5.2) and Claude (Sonnet 4.6) were used to provide stylistic editing of select sentences, which was verified and/or paraphrased by the authors.

Author Contribution Statement

Naina Balepur: Conceptualization, Methodology, Writing. **Keshav Subramonian:** Methodology, Software. **Hari Sundaram:** Supervision.

Acknowledgments

Thank you to Aditya Karan and the reviewers who provided feedback on early drafts of this work. We additionally thank undergraduate students Xingrui Pei and Krish Mewade for their work on a prior direction of this project, which helped shape the approach we ultimately pursued.

References

- [1] Naoki Abe, Prem Melville, Cezar Pendus, Chandan K Reddy, David L Jensen, Vince P Thomas, James J Bennett, Gary F Anderson, Brent R Cooley, Melissa Kowalczyk, et al. 2010. Optimizing debt collections using constrained reinforcement learning. In Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining. 75–84.
- [2] Réka Albert and Albert-László Barabási. 2002. Statistical mechanics of complex networks. Reviews of modern physics 74, 1 (2002), 47.
- [3] Veronika Alexander, Collin Blinder, and P. Zak. 2018. Why trust an algorithm? Performance, cognition, and neurophysiology. Comput. Hum. Behav. 89 (2018), 279–288. doi:10.1016/j.chb.2018.07.026
- [4] Juha M Alho and Mikko A Salo. 2000. Merit rating and formula-based resource allocation. International Journal of Educational Management 14, 3 (2000), 95–100.
- [5] Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. 2022. Machine bias. In Ethics of data and analytics. Auerbach Publications, 254–264.
- [6] Naina Balepur and Hari Sundaram. 2024. Intervening to Increase Community Trust for Fair Network Outcomes. In The 2024 ACM Conference on Fairness, Accountability, and Transparency. 1827–1837.
- [7] Solon Barocas, Moritz Hardt, and Arvind Narayanan. 2023. Fairness and machine learning: Limitations and opportunities. MIT press.
- [8] Hamsa Bastani, Kimon Drakopoulos, Vishal Gupta, Ioannis Vlachogiannis, Christos Hadjichristodoulou, Pagona Lagiou, Gkikas Magiorkinis, Dimitrios Paraskevis, and Sotirios Tsiodras. 2021. Efficient and targeted COVID-19 border testing via reinforcement learning. Nature 599, 7883 (2021), 108–113.
- [9] Rajeev Bhattacharya, Timothy M Devinney, and Madan M Pillutla. 1998. A formal model of trust based on outcomes. Academy of management review 23, 3 (1998), 459–472.
- [10] Cristian H Birzer and Jaimee Hamilton. 2019. Humanitarian engineering education fieldwork and the risk of doing more harm than good. Australasian Journal of Engineering Education 24, 2 (2019), 51–60.
- [11] Engineers Without Borders. [n. d.]. Engineers Without Borders. <https://ewb-usa.org/>
- [12] Joseph L Bower and Clark G Gilbert. 2005. From resource allocation to strategy. Oxford University Press, USA.
- [13] Thomas C Brown. 1984. The concept of value in resource allocation. Land economics 60, 3 (1984), 231–246.
- [14] Francesca Cabiddu, Ludovica Moi, Gerardo Patriotta, and David G Allen. 2022. Why do users trust algorithms? A review and conceptualization of initial trust and trust over time. European management journal 40, 5 (2022), 685–706.
- [15] Zhong Cao, Shaobing Xu, Xinyu Jiao, Huei Peng, and Diange Yang. 2022. Trustworthy safety improvement for autonomous driving using reinforcement learning. Transportation research part C: emerging technologies 138 (2022), 103656.
- [16] Anetta Caplanova, Rudolf Sivak, and Estera Szakadatova. 2021. Institutional trust and compliance with measures to fight COVID-19. International Advances in Economic Research 27, 1 (2021), 47–60.
- [17] Anupam Chander. 2016. The racist algorithm. Mich. L. Rev. 115 (2016), 1023.
- [18] Melissa Chapman, Caleb Scoville, Marcus Lapeyrolerie, and Carl Boettiger. 2022. Power and accountability in reinforcement learning applications to environmental policy. arXiv preprint arXiv:2205.10911 (2022).
- [19] Rex Chen, Ruiyi Wang, Norman Sadeh, and Fei Fang. 2025. Missing Pieces: How Do Designs that Expose Uncertainty Longitudinally Impact Trust in AI Decision Aids? An In Situ Study of Gig Drivers. In Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency. 790–816.
- [20] John Danaher. 2016. The threat of algocracy: Reality, resistance and accommodation. Philosophy & technology 29, 3 (2016), 245–268.

- [21] John Danaher, Michael J Hogan, Chris Noone, Rónán Kennedy, Anthony Behan, Aisling De Paor, Heike Felzmann, Muki Haklay, Su-Ming Khoo, John Morison, et al. 2017. Algorithmic governance: Developing a research agenda through the power of collective intelligence. *Big data & society* 4, 2 (2017), 2053951717726554.
- [22] Íñigo De Troya, Jacqueline Kernahan, Neelke Doorn, Virginia Dignum, and Roel Dobbe. 2025. Misabstraction in Sociotechnical Systems. In *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency*. 1829–1842.
- [23] Nina Deliu. 2023. Reinforcement learning for sequential decision making in population research. *Quality & Quantity* (2023), 1–24.
- [24] Zehranaz Dönmez, Merve Ayyıldız, Benay Uslu, Ozlem Karsu, and Bahar Y Kara. 2022. Fairness in Humanitarian Logistics: State of the Art and Future Directions. (2022).
- [25] Gabriel Dulac-Arnold, Nir Levine, Daniel J Mankowitz, Jerry Li, Cosmin Paduraru, Sven Gowal, and Todd Hester. 2021. Challenges of real-world reinforcement learning: definitions, benchmarks and analysis. *Machine Learning* 110, 9 (2021), 2419–2468.
- [26] Paul Erdős, Alfréd Rényi, et al. 1960. On the evolution of random graphs. *Publications of the* (1960).
- [27] Alicia Freil, Sabid Bin Habib Pias, Selma Šabanović, and Apu Kapadia. 2025. How Misclassification Severity and Timing Influence User Trust in AI Image Classification: User Perceptions of High-and Low-Stakes Contexts. In *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency*. 2906–2923.
- [28] Jones Granatyr, Vanderson Botelho, Otto Robert Lessing, Edson Emilio Scalabrin, Jean-Paul Barthès, and Fabrício Enembreck. 2015. Trust and reputation models for multiagent systems. *ACM Computing Surveys (CSUR)* 48, 2 (2015), 1–42.
- [29] Travis Greene, Galit Shmueli, and Soumya Ray. 2023. Taking the person seriously: Ethically aware IS research in the era of reinforcement learning-based personalization. *Journal of the Association for Information Systems* 24, 6 (2023), 1527–1561.
- [30] Daria Gritsenko and Matthew Wood. 2022. Algorithmic governance: A modes of governance approach. *Regulation & Governance* 16, 1 (2022), 45–62.
- [31] Siddharth Gulati, S. Sousa, and D. Lamas. 2019. Design, development and evaluation of a human-computer trust scale. *Behaviour Information Technology* 38 (2019), 1004 – 1015. doi:10.1080/0144929X.2019.1656779
- [32] Yaohui Guo and X Jessie Yang. 2021. Modeling and predicting trust dynamics in human–robot teaming: A Bayesian inference approach. *International Journal of Social Robotics* 13, 8 (2021), 1899–1909.
- [33] PA Hancock, Theresa T Kessler, Alexandra D Kaplan, Kimberly Stowers, J Christopher Brill, Deborah R Billings, Kristin E Schaefer, and James L Szalma. 2023. How and why humans trust: A meta-analysis and elaborated model. *Frontiers in psychology* 14 (2023), 1081086.
- [34] Peter A Hancock, Deborah R Billings, Kristin E Schaefer, Jessie YC Chen, Ewart J De Visser, and Raja Parasuraman. 2011. A meta-analysis of factors affecting trust in human-robot interaction. *Human factors* 53, 5 (2011), 517–527.
- [35] Xiangkun He, Wenhui Huang, and Chen Lv. 2024. Toward trustworthy decision-making for autonomous vehicles: A robust reinforcement learning approach with safety guarantees. *Engineering* 33 (2024), 77–89.
- [36] Paul W Holland, Kathryn Blackmond Laskey, and Samuel Leinhardt. 1983. Stochastic blockmodels: First steps. *Social networks* 5, 2 (1983), 109–137.
- [37] CL Hu, X Liu, and YK Hua. 2016. A bi-objective robust model for emergency resource allocation under uncertainty. *International Journal of Production Research* 54, 24 (2016), 7421–7438.
- [38] Leonid Hurwicz. 1973. The design of mechanisms for resource allocation. *The American Economic Review* 63, 2 (1973), 1–30.
- [39] Alex Ingrams, Wesley Kaufmann, and Daan Jacobs. 2022. In AI we trust? Citizen perceptions of AI in government decision making. *Policy & Internet* 14, 2 (2022), 390–409.
- [40] Shiv Issar and Aneesh Aneesh. 2022. What is algorithmic governance? *Sociology Compass* 16, 1 (2022), e12955.
- [41] Lixin Jiang, E. Bettac, Hyun Jung Lee, and T. Probst. 2022. In Whom Do We Trust? A Multifoci Person-Centered Perspective on Institutional Trust during COVID-19. *International Journal of Environmental Research and Public Health* 19 (2022). doi:10.3390/ijerph19031815
- [42] Jeongwon Jo, He Zhang, Jie Cai, and Nitesh Goyal. 2025. AI Trust Reshaping Administrative Burdens: Understanding Trust-Burden Dynamics in LLM-Assisted Benefits Systems. In *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency*. 1172–1183.
- [43] Anneli Kaasa and Luca Andriani. 2022. Determinants of institutional trust: the role of cultural context. *Journal of Institutional Economics* 18, 1 (2022), 45–65.
- [44] Christian Katzenbach and Lena Ulbricht. 2019. Algorithmic governance. *Internet Policy Review* 8, 4 (2019), 1–18.
- [45] Ryan Kennedy, Philip D. Waggoner, and M. Ward. 2021. Trust in Public Policy Algorithms. *The Journal of Politics* 84 (2021), 1132 – 1148. doi:10.1086/716283
- [46] Sherrie YX Komiak and Izak Benbasat. 2006. The effects of personalization and familiarity on trust and adoption of recommendation agents. *MIS quarterly* (2006), 941–960.
- [47] Varun Kompella, Roberto Capobianco, Stacy Jong, Jonathan Browne, Spencer Fox, Lauren Meyers, Peter Wurman, and Peter Stone. 2020. Reinforcement learning for optimization of COVID-19 mitigation policies. *arXiv preprint arXiv:2010.10560* (2020).
- [48] Raphael Koster, Miruna Pislari, Andrea Tacchetti, Jan Balaguer, Leqi Liu, Romuald Elie, Oliver P Hauser, Karl Tuyls, Matt Botvinick, and Christopher Summerfield. 2025. Deep reinforcement learning can promote sustainable human behaviour in a common-pool resource problem. *Nature Communications* 16, 1 (2025), 2824.

- [49] Stavroula Kridera and Andreas Kanavos. 2024. Exploring trust dynamics in online social networks: A social network analysis perspective. *Mathematical and Computational Applications* 29, 3 (2024), 37.
- [50] Gloria Hyunjung Kwak, Lowell Ling, and Pan Hui. 2021. Deep reinforcement learning approaches for global public health strategies for COVID-19 pandemic. *PLoS one* 16, 5 (2021), e0251550.
- [51] Theresa Law and Matthias Scheutz. 2021. Trust: Recent concepts and evaluations in human-robot interaction. *Trust in human-robot interaction* (2021), 27–57.
- [52] Min Kyung Lee, Daniel Kusbit, Anson Kahng, Ji Tae Kim, Xinran Yuan, Allissa Chan, Daniel See, Ritesh Noothigattu, Siheon Lee, Alexandros Psomas, et al. 2019. WeBuildAI: Participatory framework for algorithmic governance. *Proceedings of the ACM on human-computer interaction* 3, CSCW (2019), 1–35.
- [53] Yin-Yang Lee and Julia L Lin. 2011. How much does trust really matter? A study of the longitudinal effects of trust and decision-making preferences on diabetic patient outcomes. *Patient education and counseling* 85, 3 (2011), 406–412.
- [54] Michael Lewis, Katia Sycara, and Phillip Walker. 2018. The role of trust in human-robot interaction. *Foundations of trusted autonomy* (2018), 135–159.
- [55] Che-hui Lien, Jyh-Jeng Wu, Ying-Hueih Chen, and Chang-Jhan Wang. 2014. Trust transfer and the effect of service quality on trust in the healthcare industry. *Managing Service Quality* 24, 4 (2014), 399–416.
- [56] TP Lillicrap. 2015. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971* (2015).
- [57] Michele Loi, Andrea Ferrario, and Eleonora Viganò. 2023. How much do you trust me? A logico-mathematical analysis of the concept of the intensity of trust. *Synthese* 201, 6 (2023), 186.
- [58] M. Lounsbury. 2023. The Problem of Institutional Trust. *Organization Studies* 44 (2023), 308 – 310. doi:10.1177/01708406221131415
- [59] Arianna Manzini, Geoff Keeling, Nahema Marchal, Kevin R McKee, Verena Rieser, and Iason Gabriel. 2024. Should Users Trust Advanced AI Assistants? Justified Trust As a Function of Competence and Alignment. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency*. 1174–1186.
- [60] Catherine A Maritan and Gwendolyn K Lee. 2017. Resource allocation and strategy. 2411–2420 pages.
- [61] Roger C Mayer, James H Davis, and F David Schoorman. 1995. An integrative model of organizational trust. *Academy of management review* 20, 3 (1995), 709–734.
- [62] Qirui Mi, Siyu Xia, Yan Song, Haifeng Zhang, Shenghao Zhu, and Jun Wang. 2023. Taxai: A dynamic economic simulator and benchmark for multi-agent reinforcement learning. *arXiv preprint arXiv:2309.16307* (2023).
- [63] Arthur H Miller and Ola Listhaug. 1990. Political parties and confidence in government: A comparison of Norway, Sweden and the United States. *British journal of political science* 20, 3 (1990), 357–386.
- [64] Alan S Miller and Tomoko Mitamura. 2003. Are surveys on trust trustworthy? *Social Psychology Quarterly* (2003), 62–70.
- [65] Divya J Nair, David Rey, and Vinayak V Dixit. 2017. Fair allocation and cost-effective routing models for food rescue and redistribution. *IJSE Transactions* 49, 12 (2017), 1172–1188.
- [66] Surya Nepal, Wanita Sherchan, and Cecile Paris. 2011. Strust: A trust model for social networks. In *2011 IEEE 10th International Conference on Trust, Security and Privacy in Communications*. IEEE, 841–846.
- [67] Andreas Nicklisch and Fabian Paetzel. 2020. Need-based justice and distribution procedures: The perspective of economics. In *Need-Based Distributive Justice: An Interdisciplinary Perspective*. Springer, 161–189.
- [68] Regina Padmanabhan, Nader Meskin, Tamer Khattab, Mujahed Shraim, and Mohammed Al-Hitmi. 2021. Reinforcement learning-based decision support system for COVID-19. *Biomedical Signal Processing and Control* 68 (2021), 102676.
- [69] Eva Paraschou, Maria Michali, Sofia Yfantidou, Stelios Karamanidis, Stefanos Rafail Kalogeros, and Athena Vakali. 2025. Ties of Trust: a bowtie model to uncover trustor-trustee relationships in LLMs. In *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency*. 1715–1728.
- [70] Thomas Princen. 2005. *The logic of sufficiency*. Mit Press.
- [71] Duncan Purves and Jeremy Davis. 2022. Public trust, institutional legitimacy, and the use of algorithms in criminal justice. *Public Affairs Quarterly* 36, 2 (2022), 136–162.
- [72] L. PytlikZillig, Christopher D. Kimbrough, Ellie Shockley, T. Neal, M. Herian, Joseph A. Hamm, B. Bornstein, and A. Tomkins. 2017. A longitudinal and experimental study of the impact of knowledge on the bases of institutional trust. *PLoS ONE* 12 (2017). doi:10.1371/journal.pone.0175387
- [73] Ling Qi, Yongjie Zhou, Ruoxi Wang, Yang Wang, Yifeng Liu, and Lingyun Zeng. 2021. Perceived quality of primary healthcare services and its association with institutional trust among caregivers of persons diagnosed with a severe mental illness in China. *Journal of Psychiatric and Mental Health Nursing* 28, 3 (2021), 394–408.
- [74] M. C. Ramírez, Irene Bengo, R. Mereu, Astrid Ximena Bejarano R., and Juan C. Silva. 2011. Participative Methodology for Local Development: The Contribution of Engineers Without Borders from Italy and Colombia: Towards the Improvement of Water Quality in Vulnerable Communities. *Systemic Practice and Action Research* 24 (2011), 45–66. doi:10.1007/S11213-010-9175-3
- [75] Ateeq Ur Rehman, Aimin Jiang, Abdul Rehman, Anand Paul, Sadia Din, and Muhammad Tariq Sadiq. 2023. Identification and role of opinion leaders in information diffusion for online discussion network. *Journal of Ambient Intelligence and Humanized Computing* 14,

- 11 (2023), 15301–15313.
- [76] Sanjit Kumar Roy, James F Devlin, and Harjit Sekhon. 2015. The impact of fairness on trustworthiness and trust in banking. *Journal of Marketing Management* 31, 9-10 (2015), 996–1017.
- [77] Noel Sardana, Robin Cohen, Jie Zhang, and Shuo Chen. 2018. A Bayesian multiagent trust model for social networks. *IEEE Transactions on Computational Social Systems* 5, 4 (2018), 995–1008.
- [78] Nicolas Scharowski, Sebastian AC Perrig, Nick von Felten, Lena Fanya Aeschbach, Klaus Opwis, Philipp Wintersberger, and Florian Brühlmann. 2025. To Trust or Distrust AI: A Questionnaire Validation Study. In *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency*. 361–374.
- [79] Andrew D Selbst, Danah Boyd, Sorelle A Friedler, Suresh Venkatasubramanian, and Janet Vertesi. 2019. Fairness and abstraction in sociotechnical systems. In *Proceedings of the conference on fairness, accountability, and transparency*. 59–68.
- [80] Krithi Shailya, Shreya Rajpal, Gokul S Krishnan, and Balaraman Ravindran. 2025. LExT: Towards Evaluating Trustworthiness of Natural Language Explanations. In *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency*. 1565–1587.
- [81] Salman S Shuvo, Yasin Yilmaz, Alan Bush, and Mark Hafen. 2019. Scenario Planning for Sea Level Rise via Reinforcement Learning. In *2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. IEEE, 1–5.
- [82] Ebrahim Hamid Sumiea, Said Jadid Abdulkadir, Hitham Seddig Alhussian, Safwan Mahmood Al-Selwi, Alawi Alqushaibi, Mohammed Gamal Ragab, and Suliman Mohamed Fati. 2024. Deep deterministic policy gradient algorithm: A systematic review. *Heliyon* (2024).
- [83] Haining Tan. 2021. Reinforcement learning with deep deterministic policy gradient. In *2021 International Conference on Artificial Intelligence, Big Data and Algorithms (CAIBDA)*. IEEE, 82–85.
- [84] M Irfan Uddin, Syed Atif Ali Shah, Mahmoud Ahmad Al-Khasawneh, Ala Abdulsalam Alarood, and Eesa Alsolami. 2022. Optimal policy learning for COVID-19 prevention using reinforcement learning. *Journal of Information Science* 48, 3 (2022), 336–348.
- [85] Michael Veale, Max Van Kleek, and Reuben Binns. 2018. Fairness and accountability design needs for algorithmic support in high-stakes public sector decision-making. In *Proceedings of the 2018 chi conference on human factors in computing systems*. 1–14.
- [86] David Vengerov. 2007. A reinforcement learning approach to dynamic resource allocation. *Engineering Applications of Artificial Intelligence* 20, 3 (2007), 383–390.
- [87] Jiadai Wang, Lei Zhao, Jiajia Liu, and Nei Kato. 2019. Smart resource allocation for mobile edge computing: A deep reinforcement learning approach. *IEEE Transactions on emerging topics in computing* 9, 3 (2019), 1529–1541.
- [88] Ruotong Wang, Ruijia Cheng, Denae Ford, and Thomas Zimmermann. 2024. Investigating and designing for trust in ai-powered code generation tools. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency*. 1475–1493.
- [89] Duncan J Watts and Steven H Strogatz. 1998. Collective dynamics of ‘small-world’ networks. *nature* 393, 6684 (1998), 440–442.
- [90] Meredith Whittaker, Kate Crawford, Roel Dobbe, Genevieve Fried, Elizabeth Kaziunas, Varoon Mathur, Sarah Mysers West, Rashida Richardson, Jason Schultz, Oscar Schwartz, et al. 2018. *AI now report 2018*. AI Now Institute at New York University New York.
- [91] Jess Whittlestone, Kai Arulkumaran, and Matthew Crosby. 2021. The societal implications of deep reinforcement learning. *Journal of Artificial Intelligence Research* 70 (2021), 1003–1030.
- [92] Rosemarie E Yagoda and Douglas J Gillan. 2012. You want me to trust a ROBOT? The development of a human–robot interaction trust scale. *International Journal of Social Robotics* 4 (2012), 235–248.
- [93] Hao Ye, Geoffrey Ye Li, and Bing-Hwang Fred Juang. 2019. Deep reinforcement learning based resource allocation for V2V communications. *IEEE Transactions on Vehicular Technology* 68, 4 (2019), 3163–3173.
- [94] David S Yeager, Valerie Purdie-Vaughns, Sophia Yang Hooper, and Geoffrey L Cohen. 2017. Loss of institutional trust among racial and ethnic minority adolescents: A consequence of procedural injustice and a cause of life-span outcomes. *Child development* 88, 2 (2017), 658–676.
- [95] Chao Yu, Jiming Liu, Shamim Nemati, and Guosheng Yin. 2021. Reinforcement learning in healthcare: A survey. *ACM Computing Surveys (CSUR)* 55, 1 (2021), 1–36.
- [96] Stephan Zheng, Alexander Trott, Sunil Srinivasa, David C Parkes, and Richard Socher. 2022. The AI Economist: Taxation policy design via two-level deep multiagent reinforcement learning. *Science advances* 8, 18 (2022), eabk2607.
- [97] Cai-Nicolas Ziegler and Georg Lausen. 2005. Propagation models for trust and distrust in social networks. *Information Systems Frontiers* 7, 4 (2005), 337–358.
- [98] Kai Zong and Cuicui Luo. 2022. Reinforcement learning based framework for COVID-19 resource allocation. *Computers & Industrial Engineering* 167 (2022), 107960.

A Reproducibility and Training Details

To facilitate reproducibility and provide insight into our Deep Deterministic Policy Gradient (DDPG) implementation, we detail the environment constants, network architectures, and hyperparameters used to train our models.

A.1 Environment Parameters

Table 2 outlines the foundational constants that govern the resource allocation simulation, including the network size, underlying social dynamics, and the reward function weights.

Table 2. Simulation and Environment Constants

Parameter Name	Symbol	TO Value*	TS Value*
<i>Network & Simulation</i>			
Number of Nodes	$ V $	10	10
Episode Length (Stage 2)	T	7	7
<i>Initial Distributions</i>			
Initial Trust	τ_0	$\sim \text{Beta}(\cdot, \cdot)^{**}$	$\sim \text{Beta}(\cdot, \cdot)^{**}$
Initial Resource Need	ω_0	$\sim \text{Beta}(\cdot, \cdot)^{**}$	$\sim \text{Beta}(\cdot, \cdot)^{**}$
<i>Social Dynamics</i>			
Need Update Scaling	α	0.2	0.2
Satisfaction Exponent	p	1.3	1.3
Trust Update Speed	η	0.5	0.5
Social Influence Weight	δ	0.35	0.35
<i>Reward & Budget</i>			
Stage 1 Penalty Weight	λ_{short}	3.0	3.0
Stage 2 Penalty Weight	λ_{short}	2.0	2.0
Fixed Budget (Per Time Step)	ρ	1.0	N/A
Initial Budget	ρ^0	N/A	1.0
Maximum Budget Cap	ρ_{max}	N/A	5.0
Budget Boost Per Step	r	N/A	0.05
Trust Gain Multiplier	D	8.0	8.0

*TO = Trust-Optimizing model, TS = Trust-Sensitive model.

**Drawn from $\{\text{Beta}(2, 6), \text{Beta}(2, 4), \text{Beta}(2, 2), \text{Beta}(4, 2), \text{Beta}(6, 2)\}$.

A.2 Model Architecture

For the resource allocation model, the Deep Deterministic Policy Gradient (DDPG) implementation utilizes two distinct neural network architectures: the Actor and the Critic.

The Actor network parametrizes the policy, mapping the state space to specific resource allocation actions.

- **Hidden Layers:** Two fully connected layers containing 128 units each.
- **Hidden Activations:** Rectified Linear Unit (ReLU) activation functions are applied after each hidden layer.

- **Output Layer:** The final output layer employs a Temperature-Scaled Softmax activation function with a temperature parameter set to $T = 1.5$. This ensures the network outputs valid fractional allocations.

The Critic network approximates the action-value function, estimating the expected return $Q(s, a)$ of taking a specific action in a given state.

- **Input:** The state and action vectors are concatenated before being passed into the network.
- **Hidden Layers:** Two fully connected layers containing 128 units each.
- **Hidden Activations:** ReLU activation functions are applied after each hidden layer.
- **Output Layer:** A single linear unit outputs the scalar Q-value.

A.3 Hyperparameters and Training Details

The training procedure for the Deep Deterministic Policy Gradient (DDPG) resource allocation model relies on several key hyperparameters, detailed in Table 3. The training occurs in a multi-stage process, optimizing first for shortfall penalty reduction and subsequently for trust dynamics.

A.3.1 Replay Buffer and Sampling. To ensure stable learning and break the correlation between consecutive samples, we utilize an experience replay buffer. The buffer stores transition tuples (s_t, a_t, r_t, s_{t+1}) , up to a maximum capacity of 10^5 transitions. During each training iteration, a uniform random mini-batch of size 64 is sampled from the buffer to update the Actor and Critic networks.

A.3.2 Exploration and Target Networks. Exploration is encouraged by adding Gaussian noise to the Actor’s selected actions during the training phase. Rather than a fixed variance, the noise scale is linearly decayed over the course of training, starting from an initial scale (0.3 for Stage 1; 0.2 for Stage 2) down to a minimum threshold of 0.01. Additionally, to stabilize the learning of the Q-function, we employ target networks for both the Actor and Critic, which are updated via soft updates using the parameter $\tau = 0.001$.

Table 3. Key hyperparameters used for training the DDPG resource allocation model.

Hyperparameter	Value
Actor Learning Rate	1×10^{-4}
Critic Learning Rate	1×10^{-3}
Discount Factor (γ)	0.9
Soft Update Rate (τ)	0.001
Replay Buffer Capacity	10^5
Mini-Batch Size	64
Exploration Noise Initial Scale	0.2 – 0.3 (decays to 0.01)

A.4 Reward Convergence

The total number of training iterations allocated to each stage depends on the specific variant of the model being trained: the Trust-Optimizing model versus the Trust-Sensitive model.

A.4.1 Trust-Optimizing Model. The Trust-Optimizing model focuses primarily on maximizing trust over the long term and requires a less exhaustive pre-training phase. For this variant, Stage 1 (which optimizes strictly for the shortfall penalty) is trained for 50,000 episodes to establish a functional baseline allocation policy. The model then advances to Stage 2 (which incorporates the trust dynamics and network topology), where it is trained for 1,000,000 episodes.

A.4.2 Trust-Sensitive Model. The Trust-Sensitive model requires a highly robust initial allocation strategy before introducing the complexities of trust dynamics. To achieve this, Stage 1 is trained for a full 1,000,000 episodes. Once the shortfall penalty is thoroughly minimized, Stage 2 is subsequently trained for another 1,000,000 episodes to fine-tune the trust-aware behavior.

The learning curves for these training phases are shown in Figures 7, 9, 8, and 10. The top panel of each figure illustrates the moving average of the episodic reward over time, while the bottom panel displays the corresponding Actor and Critic network losses.

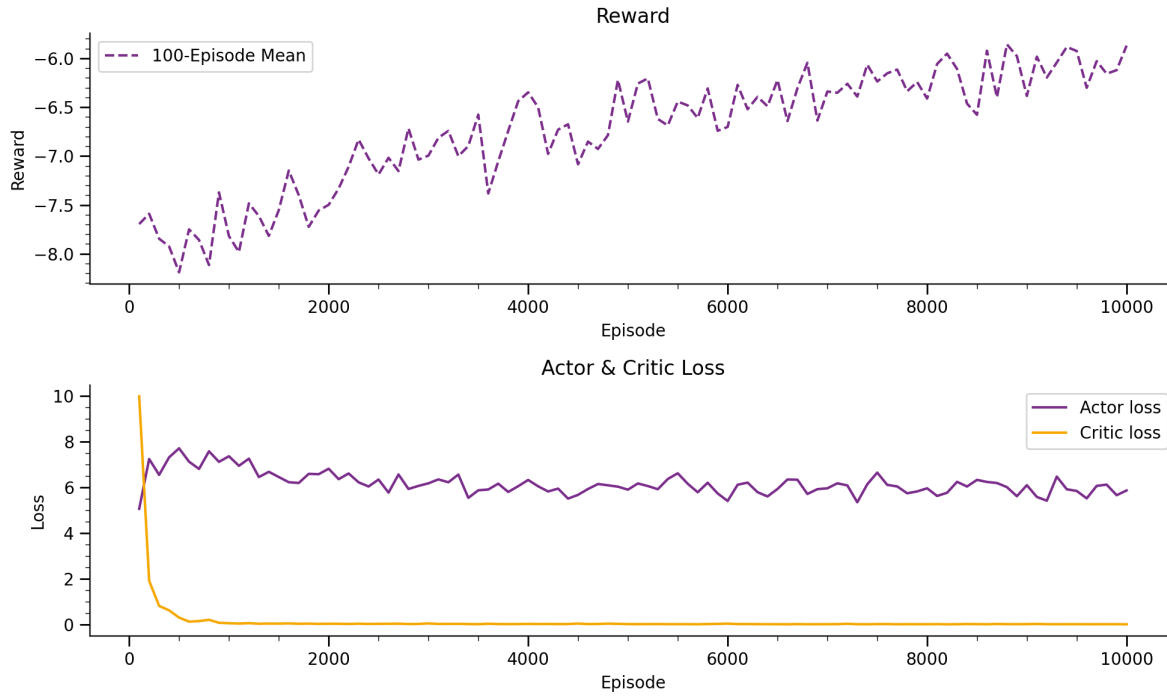


Fig. 7. Stage 1 training metrics for the Trust-Optimizing model. This shorter pre-training phase establishes a baseline allocation policy before introducing trust dynamics.

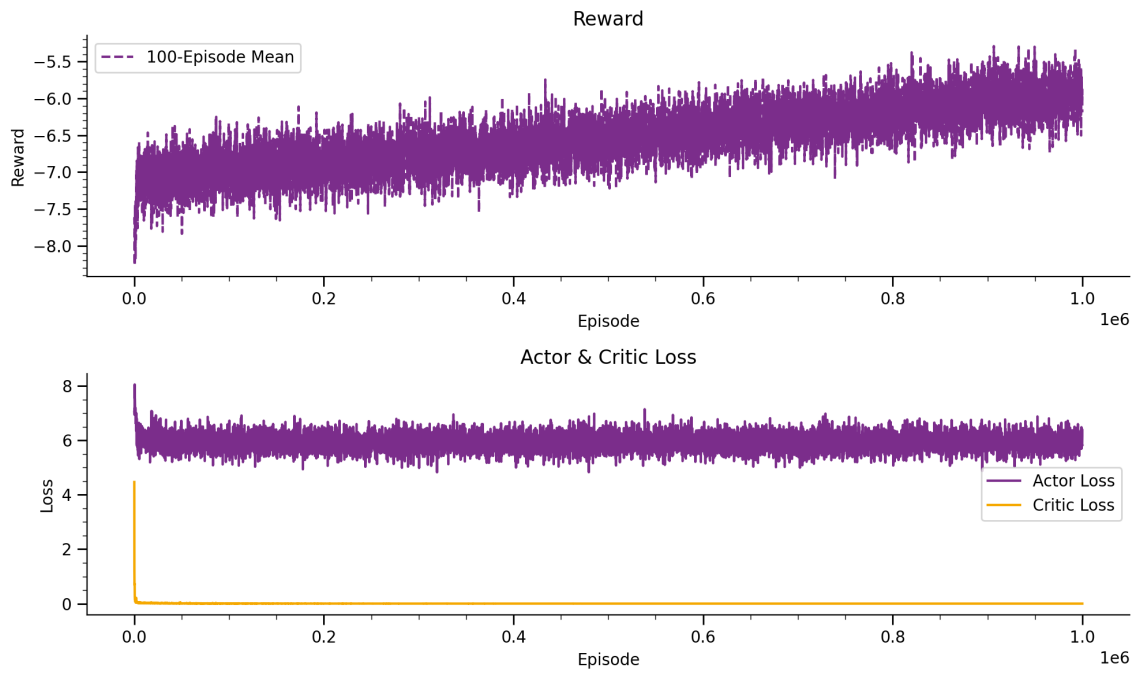


Fig. 8. Stage 1 training metrics for the Trust-Sensitive model over 10^6 episodes. This extended pre-training heavily penalizes shortfall to build a robust foundational policy.

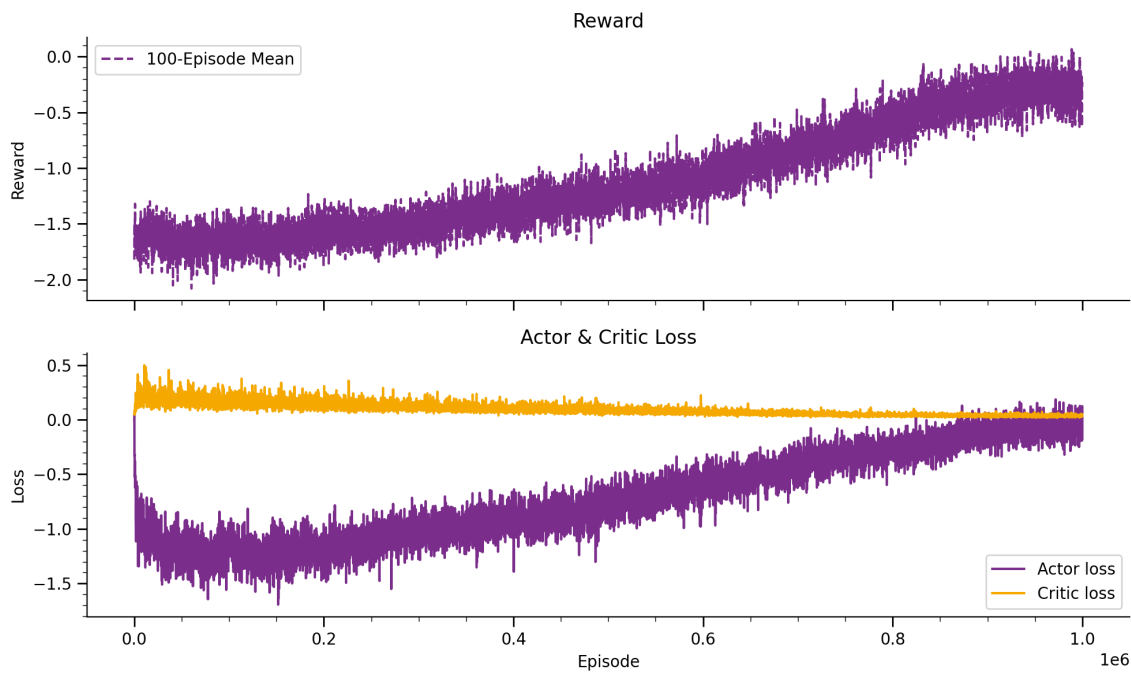


Fig. 9. Stage 2 training metrics for the Trust-Optimizing model over 10^6 episodes, optimizing for long-term trust maximization.

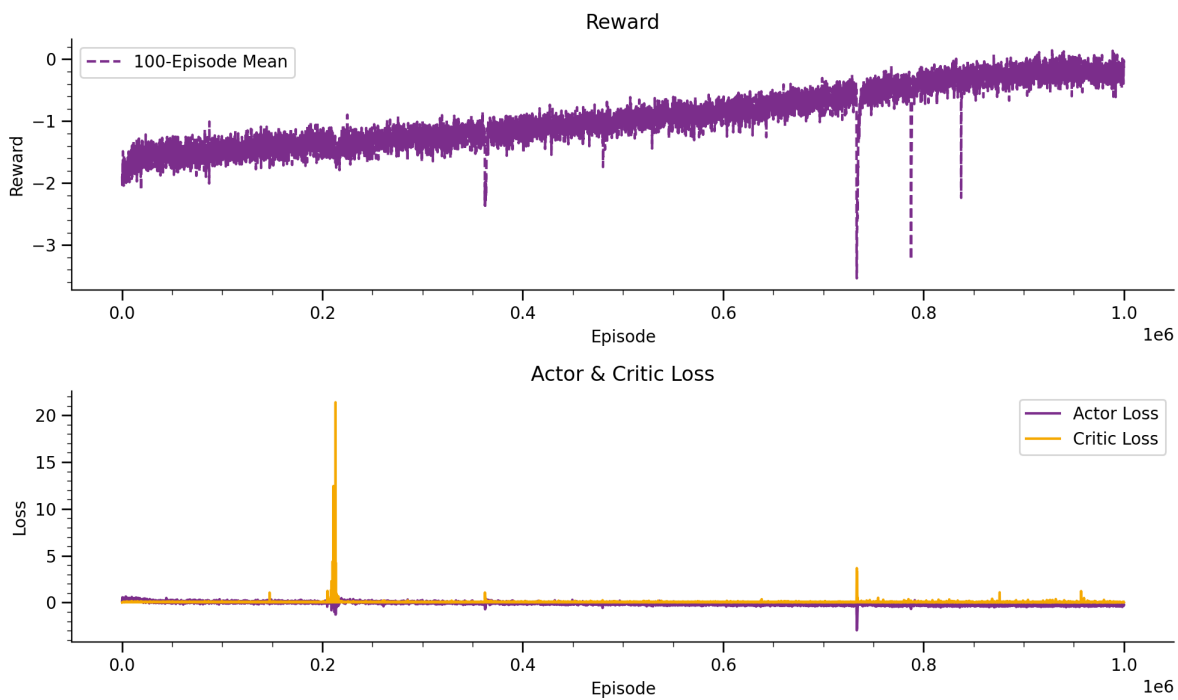


Fig. 10. Stage 2 training metrics for the Trust-Sensitive model over 10^6 episodes, fine-tuning the trust-aware behavior based on the Stage 1 baseline.