Universiteit Leiden

# Master Computer Science

**Two-Layer Agent-Based Model for Climate Cooperation:** Coupling National Policy Decisions and Consumer Behavior


| | |
|---|---|
| Name: | Senjuti Bala |
| Student ID: | s4043499 |
| Date: | 27/08/2025 |
| 1st supervisor: | Dr. Rob Saunders |
| 2nd supervisors: | Dr. Zane Kripe |


Master's Thesis in Computer Science


Leiden Institute of Advanced Computer Science
Leiden University
Einsteinweg 55
2333 CC Leiden
The Netherlands

# Contents

**Abstract**

Traditional models addressing global climate cooperation often link emissions directly to GDP which creates an absence of consumer behavioral patterns and their feedback on national policy. This thesis introduces a two-layer agent-based model (ABM) that explicitly couples country-level policy agents with consumer-level adoption agents. Layer-1 models national governments as boundedly rational decision makers and Layer-2 models heterogeneous consumer behaviors, and whether they can afford green technology or not. A feedback loop links the national subsidies to household adoption, aggregated emissions to climate vulnerability, and fairness perceptions feed back into policy choices. The model provides a reproducible and extensible ABM framework that demonstrates how multi-level feedback can shape stability and equity in climate cooperation.



**Keywords:** Agent-Based Modeling, Climate Cooperation, Multi-Agent Systems, Consumer Adoption, Fairness.

# 1 Introduction

The 2015 Paris Agreement is well known for its historic turning point in global climate policy [1]. Yet after 10 years of this movement the uneven outcomes have created a gap in vulnerable nations where they are constantly struggling to translate global commitments into feasible national action. The motivation behind this research is the experience that comes with growing up in a developing country such as Bangladesh, which is one of the most vulnerable countries to climate change due to its population and geographical exposure [14, 15]. Traditional models of climate policy often treat countries as monolithic agents tying carbon emissions directly to GDP or aggregate investment but omit intra-national heterogeneity and transnational feedback. However such view ends up ignoring the reality that emissions are not just outcomes of national policy but also of household-level consumption behaviors. This behavioral dynamics is shaped by affordability, access to infrastructure and technology, and socio-political awareness within the region or the nation. Existing agent-based models (ABM) have made significant progress in modeling either macro level governance or micro-level consumer behavior but they typically treat both system in isolation without accounting the influence of one another [2, 3]. Modeling climate cooperation across heterogeneous and interdependent actors thus requires an architecture that can formally represent both institutional policy and individual-level behavior, along with the dynamic feedback between them. This thesis addresses this gap by introducing a two-layer agent-based modeling framework that couples national policy decision-making with population-level behavioral dynamics. The model consists of:

- **Layer 1**, which simulates national agents i.e., developed, developing, and vulnerable countries. This would essentially let the model make climate policy decisions based on fairness, cost, aid, and risk.
- **Layer 2**, which simulates populations whose emissions depend on consumption and adoption of green technology. This behavioral dynamics is influenced by income, available infrastructure, and awareness.

A feedback loop connects the two layers, (1) national policy affects consumer incentives and resulting emissions, (2) emissions update climate vulnerability and fairness perception, (3) these in turn influence future national policy decisions. The key contributions of this thesis are as follows:

- A formal two-layer ABM architecture coupling country-level and population-level dynamics.
- A utility-based framework for simulating fairness perception, aid, and climate risk in national policy behavior.
- Empirical grounding using real-world economic, environmental, and behavioral parameters for agent design.
- A set of simulation scenarios exploring cooperation emergence under various aid and fairness regimes.

By modeling these interactions over a long time horizon of 100 years, the framework allows for the simulation of long-term cooperation trajectories under conditions of heterogeneity and bounded rationality. It also enables systematic exploration of conditions under which cooperation is fair, effective, and self-reinforcing. The central research question is as follows:

> **How does coupling national climate policy incentives with consumer adoption dynamics in a two-layer agent-based model influence the emergence of fair and sustainable climate cooperation across economically diverse countries?**

## 1.1 Objectives and Contributions

- Model the interplay between national climate policies and household-level consumer behaviors through a two-layer agent-based framework. This includes defining `CountryAgent` (GDP, emissions, aid, climate risk, fairness) and `ConsumerAgent` (income, cost, awareness, access), with adoption decisions linked to policy subsidies.

- Simulate dynamic feedback loops where national policies affect consumer adoption, consumer adoption alters emissions, and emissions update climate vulnerability and fairness. This in turn influence subsequent policy choices.
- Compare outcomes across developed (Type A), developing (Type B), and vulnerable (Type C) countries under baseline and policy-driven scenarios.
- Evaluate how fairness mechanisms and international aid influence policy outcomes and emission pathways across different country types.
- Identify the conditions under which fair and stable cooperation emerges using scenario analysis of adoption rates, emissions trajectories, fairness scores, and policy activity.

# 2 Background

In order to properly address and establish the research question it is important to revisit the policy foundations. The 2015 Paris Agreement acknowledges disparities among countries in capabilities and responsibilities and emphasizes voluntary commitments, equity, and flexibility by shifting from top-down emissions targets to nationally determined contributions. On this event nearly every nation committed to reducing emissions and supporting global mitigation goals. However, nearly a decade later the outcomes fall short of expectations. As apparent as it is to all of us, global emissions have not declined at the pace which is vital for us to avoid catastrophic warming, and many developing countries remain in a bind where they are expected to decarbonize without the financial and technological means to do so. While the Paris Agreement has succeeded in increasing political participation but so far it has not delivered a mechanism for sustained, fair, feasible, and enforceable cooperation, specially for vulnerable countries.

The dynamics of first-hand experience in a developing country like Bangladesh have shaped the perspective through which global climate policy and fairness are approached in this master's thesis. A developing country like Bangladesh with high exposure to climate-related risks exemplifies the challenges that arise at the intersection of industrial dependence and climate vulnerability. While the sea level and climate risks are on the rise, Bangladesh continues to be one of the leading exporters in the garments industry [15, 14]. A considerable share of its emissions is linked to production for international consumption where proper regulatory enforcement and clean technology uptake remain absent. Recent studies have documented the embedded environmental costs of such export flows, including groundwater depletion, wastewater discharge, and chemical pollution [10, 11]. These patterns are not isolated but represent broader systemic dynamics in which low-income economies absorb environmental externalities while high-income countries retain consumption advantages. Institutional assessments further indicate that these structural asymmetries in trade, labor, and climate exposure hinder the equitable distribution of mitigation responsibilities and access to sustainability-enabling infrastructure [15, 14].

Fairness is a key concept when it comes to climate cooperation and this cooperative behavior is often evaluated not only by its effectiveness in reducing emissions but also by its fairness. Principles of equity is recognized in the climate policy literature include *historical responsibility*, *capability to pay*, and *equal per capita rights*. The perception of fairness focuses on whether countries view agreements as legitimate and whether they are willing to participate. For this thesis, fairness is operationalized in a simplified form as the deviation of a country's emissions intensity from the global average. This approach is adopted from Zhang (2022) [20] that allows fairness perceptions to influence aid redistribution, policy incentives and ultimately stability cooperation.

# 3 Related Work

From the literature review it is noticeable that the research on climate cooperation has drawn from three main areas: (i) agent-based models of consumer and policy behavior, (ii) multi-agent negotiation and cooperation frameworks, and (iii) evolutionary or strategic modeling of incentives. This thesis contributes to the advancement is climate cooperation strategies. And also joins as a participant to the the tradition of simulation modeling that is continuously seeking to enrich the present understanding through exploration. While the work done by researchers is resounding and resonating this thesis aims highlight a gap of a feedback loop between a two-layer policy-behavior model.

One of the core inspiration behind this thesis is the work done by Rai and Henry (2016) which modeled a residential energy technology adoption across U.S. neighborhoods using consumer micro-drivers i.e., affordability, awareness, and peer effects [2]. The model revealed how adoption patterns emerge heterogeneously across space and social networks. Such examples tells us that affordability and awareness alone are insufficient without supportive local structures. Peer effects are one of the essential reason behind why influence works in a network. However, this work highlighted the bottom-up adoption dynamics but it did not connect the adoption patterns back into the national climate policy or any decision making processes. For this thesis their household design was adapted into consumer-level agents of Layer-2. Then the work was extended to aggregate the consumers' adoption into national emissions, reshaping the climate vulnerability, fairness, and ultimately national policies. Another extension of Rai and Henry's work is done by Mishra et al. (2025) provided a comprehensive review of agent-based modeling applications in energy transitions, spanning consumers, cities, microgrids, and market systems [3]. The authors have emphasized the importance of behavioral realism and system feedbacks for understanding transition dynamics. Even though their review highlighted many modeling directions but it does not present a framework that explicitly connects the consumers household level adoption behavior with national policy feedback. This thesis has taken inspiration from their macro-level framing idea to design policy agents in Layer-1 and attempted to connect the consumer-level adoption feedback. Thus the idea of creating a coupled two-layer model was initiated.

Beyond addressing the baselines other important works have shaped specific components of the model as well. An important precedent is the ENGAGE framework introduced by Gerst et al. (2013) [21]. ENGAGE is a multi-level agent-based model designed after Putnam's two-level game. This work explicitly linked three governance scales: international treaty negotiations, national policy formation, and domestic economic-technological dynamics. At the domestic level heterogeneous firms and households act as agents within an evolutionary representation of economic growth, energy technology and climate change. This design allows for bottom-up constraints from domestic actors to influence international policy feasibility, and vice versa. While this work shares the spirit of multi-level coupling it differs from the present work in scope. The scope of this thesis foregrounds consumer adoption and fairness-based feedback into national climate policy. The observation after reviewing different works reveals that recent ABM studies rarely reference ENGAGE despite its early contribution to multi-scale integration. However, this underscores both the novelty and timeliness of reconnecting micro-behavior to policy outcomes. The agent-based simulation implemented by Mutlu and Fescioglu-Unver (2011) demonstrated carbon emissions trading system using the JADE framework. In this work country agents adaptively bid for permits in centralized markets [4]. In their work it was shown how institutional mechanisms can coordinate emissions reduction through adaptive strategies. The model presented in this thesis compliments this by shifting emphasis from centralized market to decentralize cooperation among diverse agents. On the other hand negotiation-based models also provide important context. Perrault et al. (2022) introduced RICE-N which extends the well-known RICE integrated assessment model with multi-agent reinforcement learning to simulate negotiations between regions [5]. RICE-N presented how self-interested agents might converge toward agreements under repeated negotiation. However the emissions in RICE-N are modeled as a direct function of GDP which does not resonate with the original research question of a consumer-driven national policy dynamic. Sterman et al. (2018) approached climate policy from a different angle. They have designed the World Climate Simulation as an interactive role-play coupled to the C-ROADS model [6]. Their system provides real-time feedback to participants so that they can communicate trade-offs and get informed on climate action. While powerful for education it was not intended as a predictive scientific tool. The inspiration was drawn from its feedback-loop design but formalize it computationally to allow systematic scenario analysis. Hu et al. (2023) further extended RICE-N by introducing dynamic grouping protocols, enabling alliances to form, negotiate internally, and shift over time [7]. Their work highlighted how coalition structures affect global agreements and emissions. Finally strategic and incentive-based models enrich the perspective on adaptation. Liu et al. (2023) constructed an evolutionary game involving governments, enterprises, financiers, and the public to study how incentives and pressures affect the sustainability of green technology innovation [8]. Their work on replicator dynamics show how strategies evolve under changing payoffs. Their replicator dynamics show how strategies evolve under changing payoffs. This logic of incentives informs the experimental scenarios of this thesis which test how fairness rules and aid flows reshape long-term cooperation. Tijdeman (2022) built a bilateral negotiation model using utility-based bidding and forecasting [9]. His work illustrates how adaptive agents can negotiate under uncertainty. However it is limited to bilateral cases and does not incorporate fairness. The extension of this thesis was the negotiation perspective to multiparty cooperation, embedding fairness-weighted utilities to reflect differentiated responsibilities.

Prior models have advanced participation in either micro-level adoption, macro-level transitions or negotiation mechanisms. However the attempts to integrate these perspectives are not very common. One reason is methodological, micro-level ABMs of consumer adoption often focus on behavioral diversity and diffusion processes. On the other hand macro-level models emphasize on aggregated dynamics and policy outcomes. This makes it challenging to reconcile their time scales and data requirements. Another potential reason is disciplinary. Studies of negotiation and fairness typical emerge from political science and economics. And consumer adoption models are rooted in behavioral and energy studies. This fragmented research communities can be one of the potential reason. As a result models that combine household adoption, national policy, and international cooperation still remain thin on the ground. This thesis aims to combine these strands by developing a two-layer ABM where national policies influence consumer adoption, adoption shapes emissions, and fairness perceptions feed back into national decision-making. In doing so, it directly builds on Rai and Henry's micro-level adoption dynamics,extends Mishra et al.'s macro-level feedback framing, and incorporates fairness and cooperation mechanisms from related negotiation and strategic models.

# 4 Methodology

## 4.1 Modeling Choice

The methodological foundation of this thesis is agent-based modeling (ABM). Climate cooperation is very much a real-world issue with diverse multi-level interactions, characterized by heterogeneity, bounded rationality, and emergent dynamics. This particular issue brings a multitude of complications which cannot be fully captured by equilibrium-based Integrated Assessment Models (IAMs) such DICE or RICE[16, 17]. Now IAMs (Integrated Assessment Model) still remain useful as it aim to provide a quantitative description of policy-relevant insights into environmental change globally, sustainable development issues, and the interaction between human and earth. But while it offers an aggregated rationality but it also takes us away from heterogeneity and constant adaptation. As discussed before more recent work on IAMs is RICE-N [5] which incorporate multi-agent reinforcement learning to simulate negotiations between a list of countries to focus on global emissions. However, this work still treats global carbon emissions primarily as a direct function of GDP neglecting the consumer-level adoption behaviors and barriers.

To address this particular context agent-based modeling offers several advantages [18, 19]. Firstly, it offers the representation of heterogeneity which in this case, countries differ in economic capacity and vulnerability both geographically and financially [7], and the population also as known as the "consumers" differ in income, awareness, and access [2]. Secondly, agent-based modeling offers the most interesting feature, bounded rationality, the idea that directly points to limitations and decisions based on those limitations. Agent-based modeling focuses on the limitation where the agents adapt incrementally rather than globally optimizing [9, 8]. And, lastly, ABM helps in enabling the study of cooperation dynamics, including tipping points, club formation, and collapse of agreements. These features cannot be imposed analytically [6]. Sometimes simulating different dynamics can be important for one to grasp insights that are not immediately apparent from numerical results alone.

After deciding on the model, the design choice is a two-layer architecture. The Layer-1 models country-level agents making utility-based policy decisions which is inspired by macro-level ABM frameworks for energy transition [3] and emission trading [4]. The Layer-2 models consumer adoption decisions which extends the micro-level energy adoption studies to a cross country context. This design choice ideally helps addressing the research question whether carbon emissions of an individual country can be addressed by its population, also in this case known as consumer agents. By coupling these layers through a feedback loop, the model aims to capture how national policies designated by a government body shape adoption within the region and how this reshapes the vulnerability and fairness feeding back into policy. This feedback loop is an extension of earlier work that has typically treated policy and behavior in isolation [2, 3]. The model is designed for reproducibility and extension. All rules, parameters, and update functions are explicitly documented. The simulations are repeated across random seeds to capture a stochastic variation, and results are reported using median and inter-quartile range statistics. This model design enables replication and comparison with alternative approaches such as MARL-based negotiation, [5] or evolutionary game-theoretic simulations of technology adoption [8].

## 4.2 Assumptions

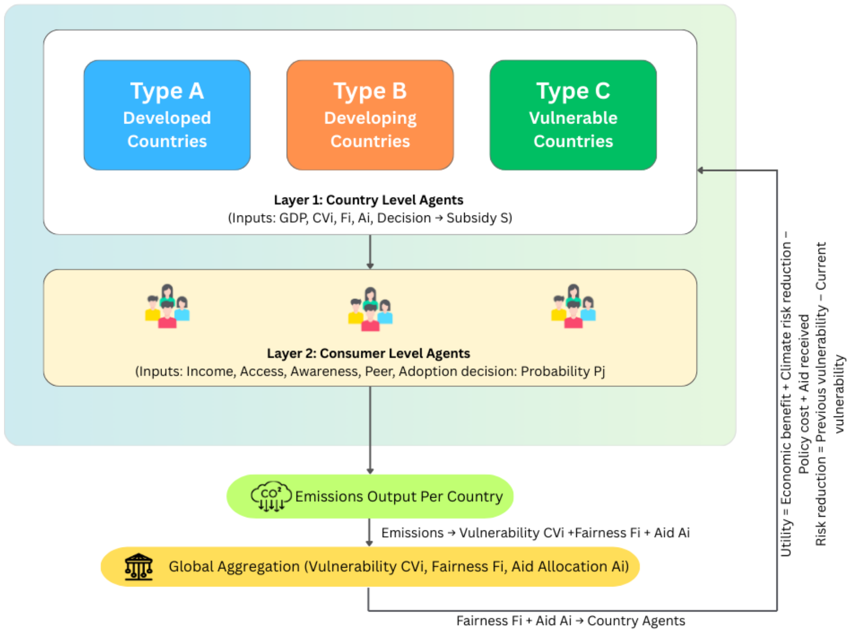


Figure 1: Conceptual Architecture of the Two-layer agent-based model. National policies influence consumer adoption, which drives emissions. And the emissions update climate vulnerability and fairness. Which is feeding back into national policy decisions. Core loop, *policy* ↔ *consumers* ↔ *emissions* ↔ *vulnerability/fairness/aid* ↔ *policy.*

Several assumptions have been taken to create the foundation of the model. The assumption is consistent with conventions in computational climate policy modeling [6, 7].

- **Country Types**: For this thesis three representative types are introduced and modeled instead of simulating all nations. They are, *Type A* (developed), *Type B*(developing), and *Type C* (vulnerable). This abstraction particularly mirrors the grouping approaches in climate negotiation models [7] and educational simulations such as C-ROADS [6].
- **Normalization of variables**: The economic capacity, vulnerability, and consumer attributes of both layer 1 and 2 are normalized to the interval [0,1]. This is similar to other agent-based climate models [3, 8]. This method have reduced calibration dependence and permitted comparability across different agents types.
- **Bounded Rationality:** The three types of country agents update subsidy intensity incrementally by evaluating candidate policies rather than solving a global optimization. This logic is consistent with bounded rationality assumptions in negotiation and learning models [9, 8].
- **Consumer Agents**: The consumer agent attributes (income, awareness, access) are designed from uniform distributions within each country agent type. This abstraction reflects the heterogeneity observed in real populations while remaining analytically tractable [2].
- **Simplified Aid and Fairness Rules**: The methods for redistribution and collaboration are implemented as functions i.e., fairness redistribution, threshold club, climate club (later discussed in depth). This is similar to previous experimental designs in fairness-oriented climate ABMs [20].
- **Temporal Abstraction**: One simulation step corresponds to approximately one year, with a 100-step horizon representing a century of cooperation. This idea resonates with the time frames used in IAMs [17, 5].
- **No Endogenous Negotiation**: The negotiation between the country agent types is not explicitly modeled. Instead strategic rules (e.g., tit-for-tat retaliation) are introduced as scenarios. This

isolates the effect of institutional logics on outcomes while avoiding confounds of complex bargaining protocols [7].

Now these assumptions created for this particular thesis have both strengths and limitations. They allow the model to isolate the interaction of fairness, aid, and adoption of green technology while maintaining interpretability. This determines the scope of the thesis: this model is *not* intended as a *forecasting tool* for generating precise national emissions, but this model addresses a *heuristic framework* for analyzing the emergence and stability of cooperation between countries and its consumers under asymmetric conditions.

## 4.3 Layer 1: country-Level Agents

The three types of country level agents represent the national government bodies making climate policy decisions under both economic and environmental constraints. Each of the country agent type $i$ is modeled as a boundedly rational decision maker which updates its subsidy intensity $S_i$ based on the trade-off between economic benefit, climate risk, policy cost, and aid received. This design of Layer-1 is directly inspired from the macro-level ABMs of energy transition [3], emissions trading simulations [4], and frainess-based cooperation models [20].

### 4.3.1 Agent Variables:

Each country agent maintains the following state variables:

| Variable | Symbol | Description |
|---|---|---|
| Economic capacity | $GDP_i$ | Normalized GDP proxy representing economic strength and ability to finance climate policy. |
| Climate vulnerability | $CV_i$ | Dynamic vulnerability index (0–1) representing exposure to climate damages. Updated each timestep based on global emissions. |
| Emissions | $E_i$ | National emissions derived from adoption outcomes in Layer 2. |
| Subsidy intensity | $S_i$ | Policy effort to subsidize green technology adoption ($0 \leq S_i \leq 1$). |
| Aid received | $A_i$ | Transfers allocated from redistribution or cooperation mechanisms. |
| Fairness index | $F_i$ | Perception of equity, based on relative emissions intensity. |
| Utility | $U_i$ | Current welfare, optimized through incremental policy updates. |

Table 1: Core variables for country-level agents.

### 4.3.2 Utility Function

Each country's utility function is defined as:

$$U_i = \alpha \cdot GDP_i + \beta \cdot (CV_{i,t-1} - CV_{i,t}) - \gamma \cdot S_i + A_i$$

where $\alpha, \beta, \gamma$ are weight parameters. The terms correspond to:

- **Economic benefit:** A function of $GDP_i$, rewarding stronger economies.
- **Risk reduction:** The difference between previous and current vulnerability ($CV_{i,t-1} - CV_{i,t}$), following Mishra et al. [3].
- **Policy cost:** A penalty proportional to subsidy intensity $S_i$.
- **Aid:** Additional utility from transfers $A_i$.

Instead of globally optimizing the utility function, here the countries use incremental bounded rationality. This evaluates the candidate's subsidy levels $S_i - \epsilon, S_i, S_i + \epsilon$ and updates to the option with the highest predicted utility. This approach ensures that the decisions are evolving gradually than solving a full optimization problem.

### 4.3.3 Climate Vulnerability Update Rule

Climate vulnerability evolves as a function of global emissions and country-specific sensitivity $R_i$:

$$CV_{i,t} = CV_{i,t-1} + \delta \cdot \left( \frac{E_{\text{global}}}{E_{\max}} \right) \cdot R_i - \kappa \cdot \frac{\max(E_i^{ref} - E_i, 0)}{E_i^{ref}}$$

where $\delta$ and $\kappa$ are parameters for risk pressure and relief, $E_{\text{global}}$ is global emissions, $E_{\max}$ is the historical maximum, and $E_i^{ref}$ is baseline emissions. This follows Mishra et al. [3], adapted to include relief when national emissions fall below baseline.

### 4.3.4 Fairness Index

Fairness perception is modeled as simplified version of fairness metric proposed by Zhang [20]:

$$F_i \;=\; 1 - \left| \frac{E_i}{GDP_i} - \frac{1}{N} \sum_{k=1}^{N} \frac{E_k}{GDP_k} \right|$$

Even though Zhang et al. used a ratio based index in their research the simplified version is adapted to an absolute deviation form. Here the, $F_i = 1$ indicates perfect fairness, and $F_i \to 0$ indicates strong deviation from the global average emissions intensity.

### 4.3.5 Aid Mechanisms

Three formalized aid mechanisms are implemented, each corresponding to an experimental scenario:

- **None:** $A_i = 0$.
- **Fairness redistribution:** $A_i = \lambda \cdot (1 - F_i)$, following redistribution rules in Zhang [20].
- **Threshold club:** $A_i = \lambda$ if $\frac{E_i}{GDP_i} \leq \tau$, else 0, modeling conditional aid membership.

### 4.3.6 Policy Update Logic

At each step, countries evaluate candidate subsidy levels $\{S_i - \epsilon, S_i, S_i + \epsilon\}$, where $\epsilon$ is a small increment. For each candidate, the utility $U_i$ is predicted using expected adoption outcomes from Layer 2. The subsidy intensity is updated to the value yielding the highest predicted utility, consistent with bounded rationality and satisfying behavior [9].

## 4.4 Layer 2: Consumer-Level Agents

The blueprint of Layer-2 is directly inspired and informed by two prior works at both the micro and macro scales. From the work of Rai and Henry [2] it was demonstrated that the diversity in household-level income, awareness, and infrastructural access offers a non-linear adoption dynamics. His work highlights the need for micro-level diversity in energy adoption models. On the other hand, Mishra et al. [3] showcased a two-layer architecture where national policies interact with sectoral adoption. This offers the reader a communication between environmental entities at a macro level. Building on these insights from the two works: the Layer-2 in this thesis models individual consumer agents nested within the country contexts. This ensures that adoption decisions reflect both individual-level variation as well as macro-level constraints.

The Layer-2 introducing consumer-level agents is constitute of individual households making adoption decisions regarding green technology. Each consumer agent $j$ is embedded within a country agent $i$. The consumer agent $j$ inherits macro-level constraints i.e., subsidy intensity $S_i$ and national vulnerability $CV_i$. The purpose behind this consumer driven Layer-2 is to target the diversity in adoption behavior and to use this pattern to translate the policy into actual emission outcomes.

#### 4.4.1 Agents Attributes

Each consumer agent is assigned the following normalized attributes:

- **Income ($I_j \in [0, 1]$):** This is the consumer's purchasing power. Higher income increases adoption probability, but interacts with policy subsidies.
- **Awareness ($A_j \in [0, 1]$):** This is the consumer's social and informational exposure to green technology which is initialized from a uniform distribution. (*Inspired by Rai and Henry [2].*
- **Access ($W_j \in [0, 1]$):** This is the consumer's infrastructural availability (e.g., charging stations, grid access). Higher access means adoption barriers in decreasing.
- **Adoption state ($Y_j \in \{0, 1\}$):** This is a binary variable where $Y_j = 1$ indicating adoption and $Y_j = 0$ otherwise. This means once adopted, the consumer agent remains adopted.

#### 4.4.2 Affordibility and Adoption Probability

Following the affordability framing in [2], the agent-level affordability is

$$\text{Affordability}_j \;=\; \frac{I_j + S_i}{\max(C_j, \varepsilon)},$$

with a small $\varepsilon > 0$ to ensure numerical stability. Here the $\text{Affordability}_j$ is clipped to $[0, 1]$ to respect normalization.

The $\phi \geq 0$ denote the peer-effects strength and $\text{peer}_j \in [0, 1]$ the local adoption share relevant to agent $j$ (by default, the fraction of adopters within the same country, though a network-based neighborhood can be substituted). The per-step adoption probability is then

$$p_j \;=\; \text{clip}\left( \underbrace{\text{Affordability}_j}_{\text{ability to pay}} \times \underbrace{A_j}_{\text{awareness}} \times \underbrace{W_j}_{\text{access}} \times \underbrace{\left(1 + \phi \cdot \text{peer}_j\right)}_{\text{social influence}},\; 0,\; 1 \right).$$

Stochastic adoption draws $Y_j \sim \text{Bernoulli}(p_j)$. If $Y_j = 1$ at any time. This means the adoption remains persistent once adopted. Importantly here the peer-effect term is multiplicative. This feature essentially means that the social influence amplifies existing adoption capacity but cannot substitute for extremely low income, awareness, or access. Such structure ideally puts an strong emphasis on the presence of structural barriers before social diffusion can accelerate adoption.[2]

#### 4.4.3 Emission Aggregation and Feedback to Layer 1

Country emissions $E_i$ aggregate realized adoption:

$$E_i \;=\; E_i^{ref} \cdot \left(1 - \eta \cdot \frac{\sum_j Y_j}{N_i}\right),$$

where $N_i$ is the number of consumers in country $i$ and $\eta \in [0, 1]$ is the per-adoption emissions reduction factor. The aggregated $E_i$ feeds into Layer 1 to update vulnerability $CV_i$ and fairness $F_i$, closing the loop.

# 5 Implementation Details

The model for this thesis was implemented in Python using the `Mesa` agent-based modeling framework for simulation scheduling and state updates. The random number generation was controlled using Python's `random` module and NumPy's `random.seed`. Seeding was passed explicitly to each simulation run ensuring reproducibility. `Pandas` DataFrames is used for statistical analysis and visualization.

All simulations were run under a time horizon of 100 years. Each country agent is initialized with a fixed population of 300 consumer agents. The number was chosen after preliminary tests to ensure stable aggregate patterns without excessive runtime. For robustness each scenario was run across 100 seeds.[6]. At the start of simulation, each country agent is initialized with normalized economic capacity ($GDP_i$), climate vulnerability ($CV_i$), and baseline emissions ($E_i^{ref}$). Consumer agents are instantiated within each country agent and assigned attributes $I_j$, $A_j$, and $W_j$. These are taken from uniform distributions and assumptions described in Section 2.2. All variables have been scaled to the interval $[0, 1]$. This made it simpler to compare across countries and different experimental scenarios.

## 5.1 Experimental Scenarios

The scenarios for this thesis are defined through a configuration dictionary specifying aid mechanisms, policy update rules, and shock interventions. Following are the experimental scenarios implemented in this thesis:

- **S1 (Baseline):** No aid, no exogenous shocks, bounded subsidy updates available. The baseline serves as a reference point isolating the endogenous dynamics of bounded rationality and consumer adoption without any external support or feedback loop.
- **S2 (Coupled Distribution):** Here the aid is redistributed based on the fairness deficits ($A_i \propto 1 - F_i$). This scenario was implemented to test whether redistributing resources according to fairness perceptions improves the adoption modality or not. If it actually improves then it would stabilize the cooperation and highlight the equity debates in climate negotiation.
- **S3 (Threshold Club):** This scenario lets the countries receive a fixed aid $\lambda$ if their emissions intensity $\frac{E_i}{GDP_i}$ is below a threshold $\tau = 2.4$. This experimental design represents the conditional *Climate Clubs* where membership requires meeting standards. The $\tau = 0.5$ cut-off reflects a formalized benchmark which essentially forces the countries to reduce emissions intensity to at least half of the baseline in order to qualify. This threshold calls out the notion of "good faith" participation in the cooperation between three different types of countries.
- **S4 (Club Incentive):** Same threshold as S3 ($\tau = 0.5$) but the club members receive an additional utility bonus. This is modeled as a fixed payoff boost. This scenario extends the club members' participation by including pros like trade benefits, technology transfers or prestige even. The bonus operationalizes the idea that such agreements amidst different type of nations sometimes has the capacity to provide benefits beyond direct aid: *friends cooperating with each other can be attractive too.*
- **S5 (Tit for Tat Retaliation):** For this scenario the countries reduce their policy rigidity if the cooperation is seen as unfair. If a country's fairness index falls below a threshold ($\theta = 0.6$) or if it fails to provide aid, other countries penalize it by lowering their own policy intensity. This mechanism implements the idea of punishment-based climate clubs. Parameters were set to $\theta = 0.6$, penalty = 0.05, and `punish_no_aid` = True.
- **S6 (Awareness Shock):** This scenario offers a sudden mid-simulation increase at year 50 in consumer awareness ($A_j$). Ideally this simulates the exogenous events such climate disasters, global or local campaigns or effect of behavioral awareness. This is designed to test the effect of bottom-up behavioral change.
- **S7 (Access Expansion):** This scenario speaks about the structural improvement in access ($W_j$), e.g., government offering new improved solar panels, EV charging networks or, renewable infrastructures. The presence of Access Expansion scenario allows to compare it against S6 to test whether shocks (awareness) or access matter more or not.

The parameters were selected to balance tractability with interpretability. The risk pressure $\delta = 0.006$ and relief $\kappa = 0.60$ were calibrated so that vulnerability evolves gradually but responds visibly to emissions changes over a 100-year range of time. (following Mishra et al. [3]) The aid scale $\lambda = 0.08$ was chosen to produce non-trivial transfers while avoiding complete equalization. (following Zhang's redistribution framework [20]) The club threshold $\tau = 2.4$ was set to represent a realistic emissions intensity benchmark. This ensures that only significant mitigation qualifies a country for membership.

## 5.2 Data Collection and Output

The following metrics are recorded at each simulation step:

- Country-level: Adoption rate, emissions $E_i$, vulnerability $CV_i$, fairness $F_i$, subsidy $S_i$, aid $A_i$, and utility $U_i$.
- Consumer-level: Individual adoption decisions $Y_j$ and attributes $I_j$, $A_j$, $W_j$.

Outputs are stored in time-series tables (DataFrames) indexed by year and country type. Visualization is done through `matplotlib` plots comparing scenarios across countries and metrics. The results are reported both as full time-series and as selected snapshots of years $(0, 249, 499)$.

# 6 Results

## 6.1 Adoption Trajectories (S1 vs S2 vs S3)

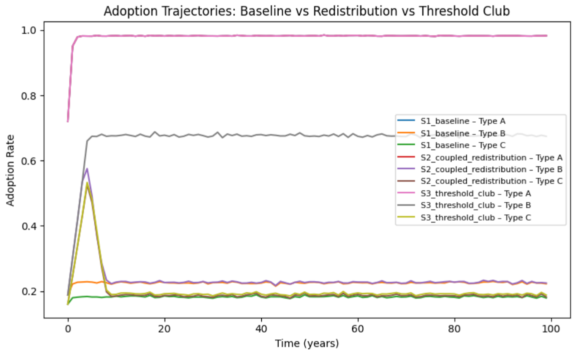


Figure 2: Figure 1. Adoption trajectories: Baseline vs Redistribution vs Threshold Club.

In the baseline (S1) adoption remains high for Type A but stagnates for Type B and C. Coupled redistribution (S2) improves adoption moderately for Type B. Threshold-based clubs (S3) accelerate adoption further in Type B, yet structural barriers leave Type C behind. This highlights that redistribution and clubs mobilize middle economies effectively but do not resolve persistent vulnerabilities for weaker countries.

## 6.2 Emissions Trajectories (S1 vs S2 vs S3)

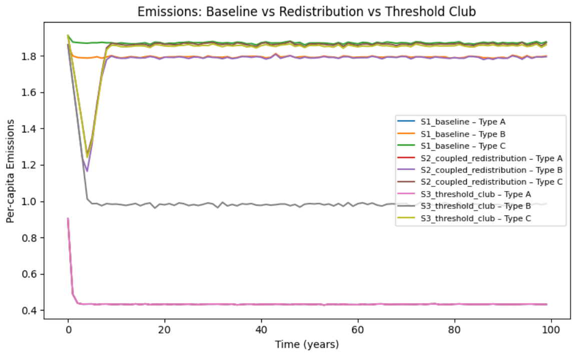


Figure 3: Figure 2. Emissions: Baseline vs Redistribution vs Threshold Club.

In the baseline (S1) emissions fall slightly for Type A but rise for Type B and C. Redistribution (S2) lowers emissions for Type B but has limited effect on Type C. Threshold clubs (S3) reinforce reductions in Type B, but again Type C shows little change. This pattern reflects how redistributive and conditional aid support convergence among middle economies but structural disadvantages persist for vulnerable countries.

## 6.3 Policy Stringency (S1 vs S2 vs S3)

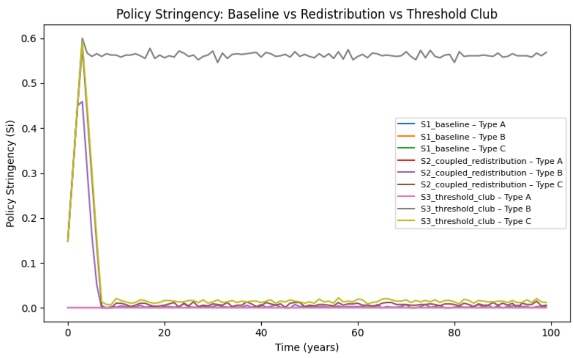


Figure 4: Figure 3. Policy stringency: Baseline vs Redistribution vs Threshold Club.

Under baseline (S1) policy stringency remains minimal. Redistribution (S2) motivates Type B to increase policy activity slightly. Threshold clubs (S3) push Type B toward stronger policies to qualify for aid. Type C remains unable to sustain higher policy levels, reflecting how structural vulnerability limits effective policy action.

## 6.4 Adoption: Incentive vs Punishment Clubs (S4 vs S5)

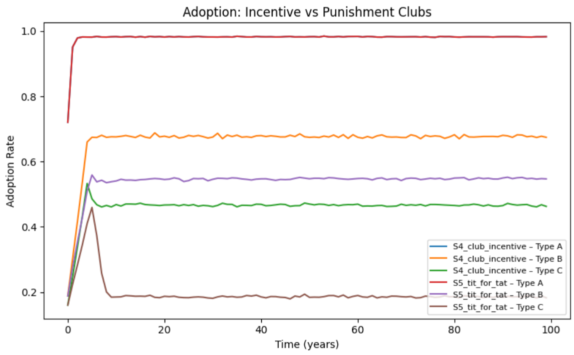


Figure 5: Figure 4. Adoption: Incentive vs Punishment Clubs.

Climate club incentives (S4) accelerate adoption in Type B and modestly improve outcomes for Type C. Punishment-based tit-for-tat retaliation (S5) enforces stricter fairness but suppresses adoption among vulnerable countries. This comparison highlights how positive incentives mobilize cooperation while punishment mechanisms risk discouraging the weakest actors.

## 6.5 Fairness: Incentive vs Punishment Clubs (S4 vs S5)

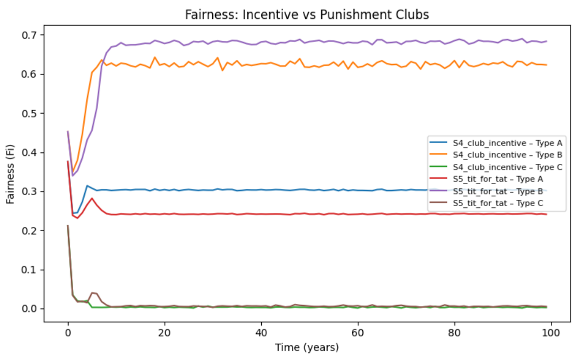


Figure 6: Figure 5. Fairness: Incentive vs Punishment Clubs.

Incentive clubs (S4) provide moderate improvements in fairness by rewarding ambitious countries. Tit-for-tat punishment (S5) raises fairness indices for Type B but leaves Type C near zero, excluded from reciprocity-based cooperation. The results underline a trade-off: incentives mobilize adoption, while punishment enforces fairness at the cost of vulnerable inclusion.

## 6.6 Adoption: Awareness Shock vs Access Expansion (S6 vs S7)

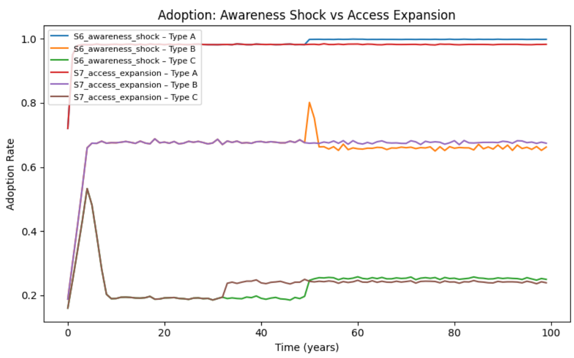


Figure 7: Figure 6. Adoption: Awareness Shock vs Access Expansion.

Awareness shocks (S6) increase adoption significantly in Type B but only marginally in Type C. Access expansion (S7) directly lowers structural barriers, producing stronger adoption gains for Type C. This highlights how behavioral interventions affect capable economies, while structural support is critical for vulnerable settings.

## 6.7 Emissions: Awareness Shock vs Access Expansion (S6 vs S7)

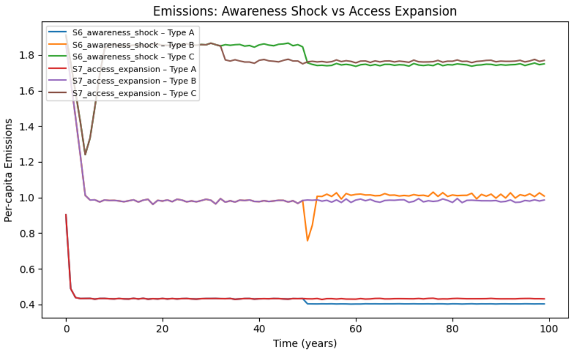


Figure 8: Figure 7. Emissions: Awareness Shock vs Access Expansion.

Awareness shocks (S6) reduce emissions modestly, mainly through improved adoption in Type B. Access expansion (S7) produces greater reductions for Type C, showing that infrastructure investments have stronger long-term climate impact in vulnerable contexts. Together these scenarios illustrate how both behavioral and structural interventions shape emission pathways differently across economies.

# 7 Discussion

The experiments with the time range of 100 years provide insights into systematic differences in early dynamics of cooperation and inequity between the three different country types. In the baseline (S1) adoption progresses steadily only for country type A with an advanced economy. But the developing (Type B) and vulnerable (Type C) economies remain stagnant (see Fig.(Fig. 2) Emissions fall modestly in Type A but it shows a rise for the rest reflecting that unilateral action is insufficient (Fig. 3). The policy stringency remains minimal across all groups (Fig. 4). Fairness Index remains low which indicates persistent inequitable burden sharing while policy activity is minimal. Introducing redistribution Scenario 2 (coupled dynamics) shows a slight shift as the pattern starts evolving. Overall the performance improves moderately for Type B with higher adoption and lower emissions compared to the baseline. Interestingly the Threshold-based clubs (Scenario 3) amplifies this effect by motivating Type B to strengthen its policies to qualify for aid in Layer 1. This is reflecting in the increase in adoption and stringency (Figs. 2 and 4). Yet in both scenarios Type C show very little change. This persistence indicates that redistribution and clubs mobilize middle economies but do not resolve the structural disadvantages of the vulnerable ones.

The contrast between incentive and punishment clubs (Scenario 4 and 5) further illustrates such dynamics. Incentive clubs (Scenario 4) is designed to promote broader adoption which produced modest improvements even for Type C (Fig. 5).The rewared mechanisms helps increasing the fairness indices slightly. But on the other hand punishment based reciprocity (Scenario 5) raises the adoption more sharply for Type B (Fig. 6). But the ultimate cost was faced by the most vulnerable since it reduces their adoption pattern. This points to a trade-off, that incentives encourage exclusivity whereas punishment enforces fairness but risks excluding the weaker actors. This echoes the real-world tensions in climate negotiations. Where cooperative exclusivity often competes against the desire for equitable burden sharing. The risk remains in the sign of dependency as middle income countries benefit disproportionately while vulnerable countries remain excluded. Such tension among different country types highlights a central policy dilemma, *should cooperation prioritize rapid adoption among capable agents or long term equity?*

Behavioral and structural interventions (Scenario 6 and Scenario 7) reveal another layer of differentiation. In the graph it is noticeable that awareness shocks(S6) drive significant adoption in Type B but provide only marginal benefits to Type C (Fig. 7). The result still shows the limited emissions reductions. In contrast access expansion (S7) lowers structural barriers directly which shows a spike in its early stage enabling stronger adoption and similarly reduces emission at the early stage for Type C (Fig. 8). This highlights how behavioral nudges can be effective in capable contexts. However structural support and access is critical in vulnerable settings for achieving long-term climate impact.

Overall the scenarios and its comparisons underscore a consistent theme across all experiments. Cooperation mechanisms that assume equal capacity has a tendency to risk amplifying equity; on the other hand structural interventions aimed at reducing barriers are essential for inclusivity. Simply it tells us that *access is key*. Redistribution and clubs help the middle economies (Type B) to converge. The incentives widen the participation of the consumers. And, access expansion strengthens the position of the most vulnerable. Within the addressed Two-Layer Feedback Loop, the government policy $S_i$ modifies household adoption probabilities $p_j$. Which then gets aggregated into national emissions $E_i$. These emissions later on feed into fairness indices $F_i$ and climate vulnerability $CV_i$. Which in turn update the utility $U_i$ aiming to guide subsequent policy choices by the government of individual country types.The adoption probability is itself driven by the micro-drivers introduced in Layer 2. Where affordability $(I_j + S_i)/C_j$, awareness $W_j$, and peer influence $(1 + \phi \cdot \text{neighbor adoption})$ jointly determine how effectively policy incentives diffuse through populations. This structure highlighted that equitable climate cooperation requires a mechanism design that is sensitive to both capacity and heterogeneous consumer-level drivers rather than assuming uniform response functions.

# 8 Conclusion

This thesis proposed and investigated a two-layer agent-based model of climate cooperation that included both national policy agents and diverse consumer agents. The model showed how policy, adoption, emissions, and climate vulnerability form a feedback loop that can either sustain cycles of cooperation or collapse under inequitable and fragile conditions. Overall this thesis is set out to answer the research question: *How does coupling national climate policy incentives with consumer adoption dynamics in a two-layer agent-based model influence the emergence of fair and sustainable climate cooperation across economically diverse countries?* The results demonstrated that linking top-down policy mechanisms with bottom-up adoption behavior *can* provide valuable insights into cooperation dynamics between different governments. National subsidies shape household adoption, adoption influences emissions, and emissions feed back into vulnerability and fairness perceptions which complete the loop between policy and behavior. In this sense the thesis succeeds in showcasing the mix of bottom-up and top-down structure which has the ability to generate self-reinforcing cycles in developed settings, conditional convergence developing economies and address the stagnation in vulnerable countries.

While the results are quite interesting but it also has several limitations, especially when it comes to real-world situations. The model is *intentionally* simplified. Agents act with bounded rationality but do not learn across time. Country agents do not adapt strategies beyond incremental subsidy updates. And the consumer agents do not evolve preferences or networks. The absence of intertemporal learning means that no agent remembers or adapts to past cooperation failures. This limitation underestimates both instability and resilience. And unfortunately this normalization and generalization approach fail to respect countries at an individual level. Similarly the model abstracts away geographical context, interpersonal relationships, and differences in energy systems. In practice whether a country is relying on coal or hydro or renewable energies is a matter of their situation and geographical context.Data quality and availability also matter. In this case the model assumes normalized inputs. Where the actual calibration would require socio-economic and energy datasets. The largest impact of these simplifications is seen in Type C countries in every scenarios. They are *vulnerable for a reason.*

Despite these limitations the framework highlights several important lessons. Redistribution alone is insufficient.Club incentives can accelerate adoption but risk leaving vulnerable countries behind. Behavioral shocks are effective primarily in middle-income settings. These findings suggest that top-down mechanisms nor bottom-up interventions alone are enough. Cooperation requires their interplay, and also the model needs to learn from its mistakes and successes. Future work should extend this model by incorporating intertemporal learning at both the consumer and country level. The emissions to specific energy mixes must be linked with country types. A network structure that takes an account for geography and trade would be very interesting to simulate but equally hard to accomplish. In this thesis the fairness is formulaic but in reality the perception differs considering historical responsibility, per capita equity or development rights. A future work could model diverse fairness heuristics across countries. The micro-drivers introduced in Layer-2 in this thesis vary in range but do not evolve. Extensions could study behavioral learning, media influence or income growth, even the effect of education awareness within social networks. Such extensions would not only reduce the oversimplification but also allow for more policy-relevant exploration of fairness, resilience, and stability in global climate cooperation.

In conclusion this thesis provides a partial but meaningful answer to its central research question. It shows that coupling national policies with consumer adoption dynamics changes the trajectory of cooperation. However the equity and convergence remain fragile under simplified assumptions. Rather than a cheer to claim it as a success story, this thesis should be read as a starting point: how demonstration of multi-level feedback can be modeled, why taking the consumer as a stakeholder matters, and of course, why representation is important. This thesis is an invitation to build richer simulations that take strong account for learning, geography, energy systems, and real-world data complexity.

## References


[1] UNFCCC. (2015). *Adoption of the Paris Agreement*. United Nations Framework Convention on Climate Change, FCCC/CP/2015/L.9/Rev.1.

[2] Rai, V., & Henry, A. D. (2016). Agent-based modelling of consumer energy choices. *Nature Climate Change*, 6(6), 556–562.

[3] Mishra, A., Rai, V., & Henry, A. D. (2022). *Agent-based modeling of energy transitions: A review and outlook*. Renewable and Sustainable Energy Reviews, 162, 112403. https://doi.org/10.1016/j.rser.2022.112403

[4] Mutlu, F., & Fescioglu-Unver, N. (2011). Agent based simulation of carbon emissions trading market. *IEEE Proceedings*, DOI: 10.1109/SIMUL.2011.17

[5] Perrault, A., et al. (2022). AI for global climate cooperation: Modeling negotiation strategies in climate policy. *arXiv preprint arXiv:2208.07004*.

[6] Sterman, J. D., et al. (2018). Combining role-play with interactive simulation to motivate informed climate action: Evidence from the World Climate simulation. *PLOS ONE*, 13(8), e0202877.

[7] Hu, Y., et al. (2023). Dynamic grouping for climate change negotiation: Facilitating cooperation and balancing interests through effective strategies. *arXiv preprint arXiv:2307.13893*.

[8] Liu, Y., et al. (2023). Evolutionary game analysis on behavioral strategies of green technology innovation. *Journal of Cleaner Production*, 406, 137017.

[9] Tijdeman, M. (2022). Predicting the outcomes of automated negotiations. Master's Thesis, Utrecht University.

[10] Rasmussen, K. H., & Setiawati, M. D. (2025). Water Unequal Exchange: Embedded Groundwater, Chemicals, and Wastewater in Textile Trade from Bangladesh to the EU and the USA. *Sustainability*, 17(4818). https://doi.org/10.3390/su17114818

[11] Shahriar, S., Smallbone, K., & Wyche, K. (2024). An Effective Solution: Water Pollution by the Textile Industry in Bangladesh. *Int. J. Environ. Pollut. Remed.*, 12. https://doi.org/10.11159/ijepr24.004

[12] International Labour Organization. (2022). *Decent Work in the Textiles and Garment Sector*. ILO Sectoral Policies Department. https://www.ilo.org/wcms_838007

[13] International Labour Organization. (2019). *Sectoral Brief: Textiles, Clothing, Leather and Footwear*. Geneva: ILO. Retrieved from https://www.ilo.org/sector/Resources/publications/WCMS_669355

[14] International Labour Organization. (n.d.). *Climate vulnerability and export dependence in Bangladesh*. Internal research report. (Y0672.pdf).

[15] International Labour Organization. (2022). *Decent Work in the Textiles and Garment Sector*. ILO Sectoral Policies Department. https://www.ilo.org/wcms_838007

[16] Nordhaus, W. D. (1993). *Optimal Greenhouse Gas Reductions and Tax Policy in the "DICE" Model*. American Economic Review, 83(2), 313–317.

[17] Nordhaus, W. D., & Yang, Z. (1996). *A Regional Dynamic General-Equilibrium Model of Alternative Climate-Change Strategies*. American Economic Review, 86(4), 741–765.

[18] Bonabeau, E. (2002). *Agent-based modeling: Methods and techniques for simulating human systems*. Proceedings of the National Academy of Sciences, 99(suppl 3), 7280–7287. https://doi.org/10.1073/pnas.082080899

[19] Epstein, J. M. (1999). *Agent-based computational models and generative social science*. Complexity, 4(5), 41–60. https://doi.org/10.1002/(SICI)1099-0526(199905/06)4:5¡41::AID-CPLX9¿3.0.CO;2-F

[20] Zhang, W. (2022). *Modeling fairness in climate cooperation: An agent-based approach.* Journal of Environmental Modelling and Software, 150, 105336. https://doi.org/10.1016/j.envsoft.2022.105336

[21] M.D. Gerst, P. Wang, A. Roventini, G. Fagiolo, G. Dosi, R.B. Howarth, M.E. Borsuk, "Agent-based modeling of climate policy: An introduction to the ENGAGE multi-level model framework," *Environmental Modelling & Software*, vol. 44, pp. 62–75, 2013.