Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 4, April, 2026
Website: usajournals.org

AGENT-BASED MODELING WITH REINFORCEMENT LEARNING FOR CARBON MARKET DESIGN IN UZBEKISTAN'S INDUSTRIAL SECTOR: A THEORETICAL FRAMEWORK

Abbos Saydullaev[1]
[1] Department of Green Economy and Susutainable Business, Samarkand Branch of Tashkent State University of Economics, Samarkand 140104, Uzbekistan, 0000-0002-6156-0963
abbos.saydullayev@sbtsue.uz

Abstract
Uzbekistan's updated Nationally Determined Contribution commits to a 35% reduction in greenhouse gas intensity per unit of GDP by 2030 relative to 2010, yet the country lacks a carbon pricing mechanism to drive this transition. This paper proposes a theoretical framework integrating agent-based modeling (ABM) with reinforcement learning (RL) to simulate and optimize the design of a cap-and-trade carbon market for Uzbekistan's emissions-intensive industrial sector. The framework addresses a critical research gap: the absence of computational models capturing strategic behavior of few, large, state-affiliated industrial emitters under carbon pricing in developing and transition economies. We define two agent classes-a Government Regulator and heterogeneous Industrial Enterprises-and formulate their decision problems as Markov Decision Processes with mixed continuous-discrete action spaces, where firms simultaneously choose green technology investment proportions (continuous) and net quota trading volumes (discrete). Using mechanism design principles, we derive a payoff matrix for regulator–industry interactions under strict versus loose cap scenarios, mathematically derive the equilibrium carbon price threshold at which firms transition from quota trading to green technology investment, and analyze two initial allocation regimes: grandfathering and

auctioning. Distributional analysis highlights acute risks for emissions-intensive trade-exposed sectors and monotown labor markets in cities such as Almalyk, Navoi, and Angren. The paper concludes that Uzbekistan's oligopolistic industrial structure necessitates hybrid allocation mechanisms with output-based benchmarking for trade-exposed sectors, and proposes computational implementation of the ABM-RL model as a critical next step for empirical parameterization.

Keywords: Agent-based modeling, reinforcement learning, carbon market design, emissions trading, Uzbekistan, mechanism design, Markov Decision Process, green economy transition.

Introduction

Uzbekistan's Green Economy Transition and Industrial Carbon Footprint

Uzbekistan, Central Asia's most populous nation with approximately 35 million inhabitants, stands at a pivotal juncture in its environmental policy trajectory. The country ratified the Paris Agreement in 2018 and submitted an updated Nationally Determined Contribution (NDC) in October 2021, strengthening its pledge from a 10% to a 35% reduction in specific GHG emissions per unit of GDP by 2030 relative to 2010 levels (UNFCCC, 2021). Presidential Decree PP-4477 (2019) formalized the Strategy for the Transition of the Republic of Uzbekistan to a Green Economy for 2019–2030, setting targets to double energy efficiency and raise the renewable energy share to 25% of electricity generation by 2030.

Despite these commitments, Uzbekistan's total greenhouse gas emissions remain approximately 170–200 $MtCO_2e$ per year, with the energy sector accounting for roughly 75–80% of total emissions (World Resources Institute, 2023). The country's primary energy supply is dominated by natural gas at approximately 85%, and its per capita emissions of 5–6 tCO_2e place it among the more emissions-intensive economies in the region. The industrial sector broadly-including energy generation, mining, metallurgy, chemicals, and cement-contributes approximately 28–33% of GDP while concentrating the

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 4, April, 2026
Website: usajournals.org

majority of point-source emissions within a handful of large state-affiliated enterprises (World Bank, 2022).

Four conglomerates dominate the emissions landscape: Uzbekenergo (restructured into Thermal Power Plants JSC), the state power utility operating most thermal generation capacity; Uzbekneftegaz, the state oil and gas holding; Navoi Mining and Metallurgical Combine (NMMC), operating the Muruntau gold mine and uranium extraction facilities; and Almalyk Mining and Metallurgical Combine (AMMC), the primary copper and zinc producer. Together with state-controlled chemical enterprises (Uzkimyosanoat) and expanding cement production, these entities likely account for 60–80% of industrial and energy-sector emissions, creating an exceptionally concentrated emissions profile.

The Problem of Carbon Market Design with Few Large State-Affiliated Players

The structural characteristics of Uzbekistan's industrial sector pose fundamental challenges for emissions trading system (ETS) design that standard economic theory does not adequately address. First, the oligopolistic market structure-potentially 50–100 major installations covering the bulk of national emissions-creates thin market conditions where strategic behavior and price manipulation become rational strategies for dominant firms (Hahn, 1984). Second, the state-affiliated ownership of major emitters introduces soft budget constraints and political objectives that deviate from profit-maximizing assumptions underlying competitive permit market theory (Kornai, 1986). Third, persistent energy subsidies, with domestic natural gas prices estimated at \$50–80 per 1,000 m^3 versus export parity of \$150–250+ per 1,000 m^3, create conflicting price signals that may undermine carbon price effectiveness. Fourth, the emergence of the EU Carbon Border Adjustment Mechanism (CBAM) creates external pressure on Uzbekistan's metal, cement, and fertilizer exports, generating an urgent but complex incentive for domestic carbon pricing.

Kazakhstan's experience with its ETS-launched in 2013 as the first in Central Asia, suspended in 2016 due to overallocation and near-zero prices, and restarted in 2018 with persistent thin trading-serves as a cautionary precedent (Narassimhan et al., 2018). The Kazakh experience demonstrates that transplanting ETS designs from liquid, competitive markets (e.g., the EU ETS)

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 4, April, 2026
Website: usajournals.org

into concentrated, state-dominated industrial sectors without accounting for strategic behavior produces dysfunctional markets.

Research Gap and Contribution

Existing literature on carbon market design has produced sophisticated theoretical frameworks for competitive settings (Montgomery, 1972; Stavins, 1995) and computational models for large, liquid markets (Tang et al., 2015; Zhang & Wei, 2010). Agent-based models have been applied to simulate carbon market dynamics in China and the EU (Han et al., 2025; Chappin & Dijkema, 2010), and reinforcement learning has been used for optimal taxation policy (Zheng et al., 2022). However, no published work combines ABM and RL to model strategic behavior of few large emitters under carbon pricing in a developing or transition economy context. This gap is critical because the very conditions that make carbon market design most challenging-thin markets, state ownership, political economy constraints, monotown dependencies-are precisely the conditions where standard equilibrium analysis breaks down and where computational approaches offer the greatest value.

This paper makes three contributions. First, it proposes a theoretical ABM-RL framework that integrates neoclassical cost minimization with bounded rationality to simulate industrial agent responses to carbon market design parameters. Second, it provides a formal Markov Decision Process formulation with a novel mixed continuous-discrete action space capturing simultaneous green investment and quota trading decisions. Third, it derives analytical equilibrium conditions for the carbon price threshold at which firms transition from permit trading to green technology investment, and constructs a payoff matrix for regulator–industry strategic interaction under alternative allocation regimes.

Methodology

The framework rests on a deliberate synthesis of two modeling traditions. The neoclassical foundation provides the equilibrium benchmark: under perfect competition, rational firms equalize marginal abatement costs with the permit price, achieving least-cost pollution control (Montgomery, 1972). This

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 4, April, 2026
Website: usajournals.org

benchmark defines the socially optimal allocation against which ABM outcomes are evaluated.

The agent-based modeling layer relaxes three key neoclassical assumptions. Following Simon (1955) and Arthur (2006), agents exhibit bounded rationality-they learn from experience rather than solving complete optimization problems. Following Farmer and Foley (2009), the model does not impose market clearing or equilibrium as an initial condition but allows these properties to emerge (or fail to emerge) from agent interactions. Following Tesfatsion (2006), agents are heterogeneous in their cost structures, ownership types, and behavioral rules.

The reinforcement learning component operationalizes bounded rationality through a principled learning framework. Rather than assuming agents follow fixed behavioral rules (as in traditional ABM) or solve complete optimization problems (as in neoclassical theory), RL agents discover strategies through trial-and-error interaction with the market environment (Sutton & Barto, 2018). This approach has two theoretical advantages: it naturally captures learning dynamics in nascent carbon markets where agents lack experience, and it allows agents to discover strategic behaviors (including collusion and manipulation) that may not be anticipated by the modeler (Calvano et al., 2020).

Agent Definitions

The model defines two classes of agents operating within a cap-and-trade institutional environment.

Agent Class 1: Government Regulator (GR). The regulator is a single agent that sets market design parameters and seeks to achieve dual objectives: (a) an aggregate emissions target aligned with Uzbekistan's NDC commitment (35% intensity reduction by 2030), and (b) market stability, defined as bounded permit price volatility within a policy-acceptable corridor. The regulator's decision variables are the aggregate emissions cap $\bar{E}_t$, the allocation method (grandfathering share $\alpha \in [0,1]$ versus auction share $1 - \alpha$), and the non-compliance penalty rate f per tonne of excess emissions.

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 4, April, 2026
Website: usajournals.org

Agent Class 2: Industrial Enterprises (IE). The model includes heterogeneous industrial agents indexed $i \in \{1, \dots, N\}$, representing major emitting installations. Each agent is characterized by:

- Baseline emissions $\bar{e}_i$ (tonnes CO_2e per compliance period)
- A convex abatement cost function $C_i(a_i)$ with $C_i'(a_i) > 0$, $C_i''(a_i) > 0$
- Output level q_i and output price p^q
- Ownership type $\omega_i \in \{\text{SOE,private}\}$, where SOEs face modified objective functions incorporating political constraints and soft budget effects
- Technology state $\tau_i \in [0,1]$, representing the proportion of production capacity using green technology, which shifts the abatement cost curve

Each enterprise seeks to minimize total compliance costs (abatement costs plus net permit trading costs plus any non-compliance penalties) while maintaining production targets. The critical departure from standard models is that SOE agents may weight employment preservation and production targets alongside cost minimization, reflecting the political economy constraints documented in Uzbekistan's monotown-dependent industrial structure.

Agent characteristics and decision variables

Parameter	Government Regulator	Industrial Enterprise
Objective	Minimize emissions subject to economic stability	Minimize compliance cost subject to production targets
Decision variables	Cap $\bar{E}_t$, allocation share α, penalty f	Green investment ϕ_i, net trading volume Δ_i
Constraints	NDC target, political feasibility	Budget, technology availability, production requirements
Information	Aggregate emissions, market price	Own emissions, permit holdings, market price
Ownership	Government	SOE or private, with heterogeneous cost structures

Reinforcement Learning Formulation as Markov Decision Process

Each industrial enterprise's decision problem is formulated as a Markov Decision Process (MDP) defined by the tuple $\langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma \rangle$ (Puterman, 1994; Sutton & Barto, 2018).

State space. The state vector for agent i at decision epoch t is:

$$s_t^i = \left(p_t^c,\ \tau_t^i,\ A_t^i,\ e_t^i,\ \bar{E}_t,\ q_t^i,\ p_t^q,\ t\right) \in \mathcal{S}$$

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 4, April, 2026
Website: usajournals.org

where $p_t^c \in \mathbb{R}_+$ is the current carbon permit price, $\tau_t^i \in [0,1]$ is the firm's green technology state (proportion of capacity using clean technology), $A_t^i \in \mathbb{R}_+$ is the firm's current permit holdings (initial allocation plus banked permits minus surrendered), $e_t^i \in \mathbb{R}_+$ is cumulative emissions in the current compliance period, $\bar{E}_t$ is the announced aggregate cap, q_t^i is the firm's output level, p_t^q is the output market price, and t indexes the time step within the compliance period.

Action space. Each agent selects an action from a mixed continuous-discrete (parameterized) action space:

$$a_t^i = \left(\phi_t^i,\ \Delta_t^i\right) \in \mathcal{A} = [0, \bar{\phi}] \times \{-\bar{\Delta}, \dots, -1,0,1, \dots, \bar{\Delta}\}$$

The action vector comprises two components:

Continuous component-Green technology investment proportion $\phi_t^i \in [0, \bar{\phi}]$: the fraction of available capital allocated to abatement technology investment in period t. This investment reduces the firm's marginal abatement cost curve according to:

$$MAC_i\left(a_i; \tau_t^i\right) = c_i \cdot \left(1 - \tau_t^i\right) \cdot e^{-\eta a_i}$$

where the technology state evolves as $\tau_{t+1}^i = \tau_t^i + \delta\left(\phi_t^i\right)$ with a concave function $\delta(\cdot)$ reflecting diminishing returns to investment. The parameter $\eta > 0$ governs how strongly green technology reduces abatement costs.

Discrete component-Net quota trading volume $\Delta_t^i \in \{-\bar{\Delta}, \dots, \bar{\Delta}\}$: a signed integer representing the number of permit units the firm seeks to trade, where $\Delta_t^i > 0$ denotes buying and $\Delta_t^i < 0$ denotes selling.

This parameterized action space structure follows the hybrid actor-critic approach of Fan et al. (2019) and Masson et al. (2016), requiring decomposition of the policy into a continuous head $\pi_c(\phi \mid s; \theta_c)$ and a discrete head $\pi_d(\Delta \mid s, \phi; \theta_d)$.

Non-compliance modeling. Non-compliance is not modeled as a deliberate agent choice but as an outcome of the emissions penalty function. If a firm's actual emissions exceed its permit holdings at the end of the compliance period, a penalty is automatically applied:

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 4, April, 2026
Website: usajournals.org

$$P_i^{NC} = f \cdot max(0,\ e_i^T - A_i^T)$$

where e_i^T is total emissions at compliance deadline T, A_i^T is total permit holdings, and f is the penalty rate (typically set at 2–3× the prevailing permit price to ensure compliance incentives). This formulation follows the enforcement theory of Stranlund and Dhanda (1999), where non-compliance emerges endogenously from the agent's cost-minimizing behavior rather than being an explicit strategic choice.

Reward function. The period reward for industrial agent i is:

$$r_t^i = \underbrace{p_t^q \cdot q_t^i}_{\text{revenue}} - \underbrace{c_i(q_t^i, \tau_t^i)}_{\text{production cost}} - \underbrace{C_i(a_t^i)}_{\text{abatement cost}} - \underbrace{p_t^c \cdot |\Delta_t^i|}_{\text{trading cost}} - \underbrace{\kappa \cdot \phi_t^i \cdot K_i}_{\text{investment cost}} - \underbrace{P_i^{NC}}_{\text{penalty}}$$

where κ is a capital cost parameter and K_i is the firm's capital stock. The agent maximizes the expected discounted cumulative reward:

$$V^{\pi_i}(s_0^i) = \mathbb{E}^{\pi_i}\left[\sum_{t=0}^{T} \gamma^t \cdot r_t^i \mid s_0^i\right]$$

with discount factor $\gamma \in [0.95, 0.99]$, reflecting the firm's time preference over the planning horizon.

Regulator reward. In the two-level formulation following the AI Economist architecture (Zheng et al., 2022), the regulator agent maximizes a social welfare function:

$$r_t^{GR} = \alpha_W \cdot GDP_t - \beta_W \cdot D(E_t^{\text{total}}) - \delta_W \cdot Inequality_t + \rho_W \cdot Revenue_t^{\text{auction}}$$

where $D(\cdot)$ is a convex climate damage function, $Inequality_t$ captures distributional effects across regions, and $Revenue_t^{\text{auction}}$ represents auction proceeds available for revenue recycling.

Mathematical Formulation of the Cap-and-Trade Mechanism

The cap-and-trade market operates as follows.

Cap setting. The regulator sets an aggregate emissions cap $\bar{E}_t$ for each compliance period t, declining over time to meet the NDC trajectory:

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 4, April, 2026
Website: usajournals.org

$$\bar{E}_t = \bar{E}_0 \cdot \left(1 - \frac{\rho^{\text{NDC}}}{T^{\text{NDC}}}\right)^t$$

where $\bar{E}_0$ is the initial cap, ρ^{NDC} is the cumulative required intensity reduction (0.35 for Uzbekistan's 35% target), and T^{NDC} is the target horizon in years. Note that this formulation converts the intensity-based NDC target into an approximate absolute cap path by holding projected GDP growth constant-a simplification acknowledged as a limitation.

Initial allocation. Two policy scenarios are analyzed:
Scenario G (Grandfathering): Free allocation based on historical emissions:

$$\omega_i^G = \bar{E}_t \cdot \frac{\bar{e}_i}{\sum_{j=1}^N \bar{e}_j}$$

Scenario A (Auctioning): All permits distributed via uniform-price sealed-bid auction. Each firm submits demand schedule $d_i(p)$; the clearing price p^* satisfies:

$$\sum_{i=1}^N d_i\,(p^*) = \bar{E}_t$$

Following Cramton and Kerr (2002), auctioning generates revenue $R_t = p^* \cdot \bar{E}_t$ available for recycling, while grandfathering creates windfall profits for recipients (Sijm et al., 2006).

Market clearing. In the secondary permit market, firms trade bilaterally or through a centralized exchange. In competitive equilibrium, the permit price equalizes marginal abatement costs across all firms (Montgomery, 1972):

$$MAC_1(a_1^*) = MAC_2(a_2^*) = \cdots = MAC_N(a_N^*) = p^*$$

The firm-level optimality condition is:

$$\min_{a_i, \Delta_i} C_i(a_i) + p^c \cdot \Delta_i \quad \text{s.t.} \quad \bar{e}_i - a_i \le \omega_i + \Delta_i$$

yielding the first-order condition for interior solutions:

$$C_i{}'(a_i^*) = p^c$$

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 4, April, 2026
Website: usajournals.org

Market power distortion. In Uzbekistan's concentrated market, the independence property (Montgomery, 1972) breaks down. Following Hahn (1984), a dominant firm recognizing its price impact sets:

$$C_i'(a_i^*) = p^c + (\bar{e}_i - a_i^* - \omega_i) \cdot \frac{\partial \Delta_i}{\partial p^c}$$

This implies systematic overpricing by net sellers and underpricing by net buyers, generating deadweight loss proportional to market concentration-a distortion the ABM-RL framework can quantify endogenously.

Results

Theoretical Outcomes from Mechanism Design Principles

The framework generates three classes of theoretical results: strategic interaction outcomes, equilibrium threshold conditions, and allocation regime comparisons. These are derived analytically from the model's formal structure rather than from simulation, consistent with the paper's theoretical nature.

Payoff Matrix for Regulator–Industry Strategic Interaction

The interaction between the regulator and industry can be formalized as a two-player game in normal form, where the regulator chooses between strict cap (S) and loose cap (L), while the representative industrial agent chooses between comply-and-invest (CI) and defect-and-pay-fines (DF).

Payoff matrix: Regulator (row) vs. Industry (column)

	Industry: Comply/Invest (CI)	**Industry: Defect/Pay fines (DF)**
Regulator: Strict cap (S)	$(W_S^{CI},\ -C_{\text{abate}} - I_{\text{green}})$	$(W_S^{DF} + f \cdot \Delta e,\ -f \cdot \Delta e + S_{\text{avoid}})$
Regulator: Loose cap (L)	$(W_L^{CI},\ -C_{\text{abate}}^L - I_{\text{green}}^L)$	$(W_L^{DF},\ -C_{min})$

where:

- W_S^{CI} = social welfare under strict cap with full compliance: high environmental benefit, moderate economic cost
- W_S^{DF} = social welfare under strict cap with defection: penalty revenue collected but emissions target missed
- W_L^{CI} = social welfare under loose cap with compliance: low environmental benefit, low economic cost

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 4, April, 2026
Website: usajournals.org

- W_L^{DF} = social welfare under loose cap with defection: minimal environmental and economic change (status quo)
- C_{abate} = total abatement cost under compliance
- I_{green} = green technology investment expenditure
- $f \cdot \Delta e$ = penalty payment for excess emissions
- S_{avoid} = savings from avoided abatement costs

Proposition 1 (Compliance equilibrium condition). The strategy profile (Strict cap, Comply/Invest) constitutes a Nash equilibrium if and only if:

$$C_{\text{abate}} + I_{\text{green}} < f \cdot \mathbb{E}[\Delta e] - S_{\text{avoid}}$$

That is, the total cost of compliance and investment must be less than the expected penalty minus the savings from avoided abatement. This condition holds when the penalty rate f is sufficiently high relative to the marginal abatement cost, a standard result in enforcement theory (Becker, 1968; Malik, 1990).

Proposition 2 (Regulator credibility). Under repeated interaction with a finite horizon, the regulator prefers the strict cap if and only if:

$$\alpha_W \cdot \Delta GDP^{S\to L} < \beta_W \cdot [D(E_L) - D(E_S)]$$

where $\Delta GDP^{S\to L}$ is the GDP loss from tightening the cap from loose to strict. This condition is more likely to hold when the social cost of carbon is high relative to the GDP-weighting parameter α_W/β_W-a calibration choice with significant policy implications for developing countries where growth objectives compete with climate ambitions.

Equilibrium Threshold for Transition from Trading to Green Investment

A central analytical result concerns the carbon price at which a representative firm transitions from relying on permit trading to investing in green technology. This threshold determines the effectiveness of the carbon market in driving technological transformation rather than merely redistributing emissions rights.

Theorem 1 (Green investment switching threshold). Consider firm i with quadratic abatement cost $C_i(a_i) = 1/2\, c_i a_i^2$, producing output q_i with

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 4, April, 2026
Website: usajournals.org

emissions intensity ϵ_i (so baseline emissions $\bar{e}_i = \epsilon_i \cdot q_i$). A green technology investment of fixed cost F_i reduces emissions intensity to $\epsilon_i{}' = \epsilon_i \cdot (1 - \mu_i)$ where $\mu_i \in (0,1)$ is the abatement efficiency of the technology, with annual maintenance cost m_i. At discount rate r over investment horizon T, the firm prefers green investment over permit trading if and only if the carbon permit price exceeds:

$$p^* > \frac{F_i + m_i \cdot AF(r,T)}{\mu_i \cdot \epsilon_i \cdot q_i \cdot AF(r,T)} + \frac{c_i \cdot \mu_i \cdot \epsilon_i \cdot q_i}{2}$$

where $AF(r,T) = \frac{1-(1+r)^{-T}}{r}$ is the annuity factor.

Proof. Under permit trading alone at price p^c, the firm's optimal abatement satisfies $C_i{}'(a_i^*) = c_i \cdot a_i^* = p$, yielding $a_i^* = p/c_i$. The annual cost of compliance via trading is:

$$TC_i^{\text{trade}} = 1/2\, c_i \left(\frac{p}{c_i}\right)^2 + p \cdot (\bar{e}_i - \omega_i) - \frac{p}{c_i} = p(\bar{e}_i - \omega_i) - \frac{p^2}{2c_i}$$

Under green investment, emissions fall by $\Delta e_i = \mu_i \cdot \epsilon_i \cdot q_i$ per period, reducing annual permit requirements. The present value of net savings from investment equals:

$$NPV_i^{\text{invest}} = \left(p \cdot \mu_i \cdot \epsilon_i \cdot q_i - \frac{c_i \cdot (\mu_i \cdot \epsilon_i \cdot q_i)^2}{2} - m_i\right) \cdot AF(r,T) - F_i$$

Setting $NPV_i^{\text{invest}} > 0$ and solving for p yields the threshold.

The threshold has an intuitive economic interpretation: the first term $\frac{F_i + m_i \cdot AF}{(\mu_i \cdot \epsilon_i \cdot q_i) \cdot AF}$ represents the annualized capital cost per unit of emissions avoided, while the second term $\frac{c_i \cdot \mu_i \cdot \epsilon_i \cdot q_i}{2}$ represents the marginal abatement cost savings foregone from reducing emissions via technology rather than operational abatement.

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 4, April, 2026
Website: usajournals.org

Sensitivity of green investment threshold to key parameters (illustrative)

Parameter	**Effect on threshold**	**Interpretation**
Higher fixed cost F_i ↑	threshold ↑	Costlier technology requires higher carbon price signal
Higher abatement efficiency μ_i ↑	threshold ↓	More effective technology becomes viable at lower carbon prices
Higher output q_i ↑	threshold ↓	Scale economies in green investment favor larger firms
Higher discount rate r ↑	threshold ↑	Short-termism delays green investment
Longer horizon T ↑	threshold ↓	Policy credibility reduces threshold via extended payback period

This result carries a key implication for Uzbekistan: the large scale of enterprises like NMMC and AMMC (q_i very high) lowers their investment threshold, suggesting that the oligopolistic structure may paradoxically facilitate green technology adoption-provided the carbon price signal is credible and persistent.

Allocation Regime Comparison

Under grandfathering, incumbent firms receive windfall profits equal to $\Pi_i^{\text{windfall}} = p^* \cdot \omega_i^G$. The independence property (Montgomery, 1972) guarantees that the final allocation of emissions rights is efficient regardless of initial allocation-but only under perfect competition. In Uzbekistan's concentrated market, Hahn's (1984) result applies: the initial allocation affects efficiency because dominant firms exploit their price impact. Specifically, if NMMC or AMMC receives a grandfathered allocation exceeding their competitive equilibrium emissions level, they will strategically restrict permit supply, inflating the carbon price and transferring rents from smaller firms.

Under auctioning, revenue generation enables the "double dividend" (Goulder, 1995): carbon revenue can simultaneously reduce emissions and offset distortionary existing taxes. For Uzbekistan, where environmental pollution charges are currently set at negligible levels, auction revenue estimated at $R_t = p^* \cdot \bar{E}_t$ could fund just transition programs for monotown workers-a distributional improvement over grandfathering. However, auctioning imposes immediate costs on state-owned enterprises whose balance sheets are ultimately

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 4, April, 2026
Website: usajournals.org

government liabilities, creating a circular fiscal flow that may reduce political feasibility.

Comparative analysis of allocation regimes for Uzbekistan

Criterion	**Grandfathering**	**Auctioning**	**Hybrid (OBA for EITE)**
Efficiency (competitive market)	First-best	First-best	Second-best
Efficiency (oligopolistic market)	Distorted by market power	Less distorted (no initial endowment effect)	Intermediate
Revenue generation	None (windfall to firms)	Full ($p^* \cdot \bar{E}_t$)	Partial
Carbon leakage protection	Weak	Weak without border adjustment	Strong for EITE sectors
Political feasibility	High (no new cost to SOEs)	Low (immediate fiscal impact)	Moderate
Distributional equity	Regressive (rewards high emitters)	Depends on revenue recycling	Progressive if benchmarked to best practice

Discussion

Carbon Leakage Prevention in Uzbekistan's Regional Context

Carbon leakage-the increase in emissions outside a regulated jurisdiction caused by emissions regulation within it-poses a particularly acute risk for Uzbekistan. The country borders four economies without comprehensive carbon pricing (Turkmenistan, Tajikistan, Kyrgyzstan, and Afghanistan), while Kazakhstan's ETS has produced near-zero carbon prices with limited environmental impact. Russia, a major trading partner, lacks comprehensive carbon pricing. The leakage rate, formally defined as

$$L = \frac{\Delta E_{\text{outside}}}{-\Delta E_{\text{inside}}} \times 100\%$$

(Böhringer et al., 2010), could be substantial through three channels.

The competitiveness channel is most relevant for Uzbekistan's EITE sectors. AMMC's copper and zinc exports compete directly with producers in Russia and Kazakhstan; Uzkimyosanoat's fertilizers compete with Russian and

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 4, April, 2026
Website: usajournals.org

Turkmen producers; and Bekabad steel competes with Chinese and Russian imports. Babiker (2005) demonstrated that in oligopolistic energy-intensive industries, leakage rates can exceed 100%-meaning unilateral carbon pricing increases global emissions. For Uzbekistan's highly concentrated sectors, this finding is directly applicable.

The energy price channel operates differently for Uzbekistan as a net gas exporter. Domestic carbon pricing that reduces gas consumption could lower domestic gas prices further (amplifying existing subsidy distortions) or free gas for export (a positive fiscal effect). This channel's net direction depends on the elasticity of domestic gas demand relative to export capacity constraints.

Output-based allocation (OBA) provides the most theoretically grounded leakage protection mechanism for Uzbekistan's EITE sectors. Under OBA, free allocation is proportional to current output using sector-specific benchmarks:

$$\omega_i^{OBA} = \beta_j \cdot q_{i,t}$$

where β_j is the benchmark emissions intensity for sector j. Fischer and Fox (2007) showed that OBA is equivalent to a combination of emissions tax and output subsidy, maintaining production incentives while preserving abatement incentives on the emissions-intensity margin. For Uzbekistan, we propose a hybrid allocation regime: auctioning for the power sector (which serves the domestic market and faces no leakage risk) combined with OBA benchmarking for metals, chemicals, and cement.

Distributional Effects on Monotowns and EITE Labor Markets

The most politically salient distributional concern is the impact on Uzbekistan's monotown economies. Almalyk (population ~130,000) depends almost entirely on AMMC's copper smelting and mining operations. Navoi (~180,000–200,000) exists because of NMMC's gold and uranium extraction. Angren (~130,000) centers on coal mining and thermal power generation. Bekabad (~85,000) depends on steel production. In these cities, the dominant employer accounts for the overwhelming majority of formal employment and municipal revenue.

Carbon pricing imposes costs through two mechanisms on these communities. The direct employment channel operates if carbon costs reduce production or

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 4, April, 2026
Website: usajournals.org

trigger facility rationalization. The employment elasticity with respect to carbon price, $\varepsilon_{L,\tau} = \partial lnL/\partial ln\tau$, is sharply negative in monotowns because labor mobility is constrained-workers possess industry-specific human capital (metallurgists, miners, furnace operators) with limited transferability, and geographic relocation faces housing market and social network barriers (Vona et al., 2018). The indirect fiscal channel operates through reduced municipal tax revenues and social service provision when the dominant enterprise faces higher costs.

These distributional effects create a credible political economy constraint: if carbon pricing threatens monotown viability, the regulator faces pressure to weaken the cap-undermining policy credibility and the green investment threshold derived in Section 3.3. The ABM-RL framework can model this feedback loop explicitly by incorporating the regulator's political feasibility constraint into its reward function, testing whether revenue recycling from permit auctions to monotown transition funds can maintain policy credibility while protecting vulnerable communities.

Bauer et al. (2020) demonstrated that uniform carbon prices impose disproportionately high costs on developing economies, quantifying the efficiency-sovereignty trade-off at $4.4 trillion in present-value international transfers needed for equalization. For Uzbekistan, this finding reinforces the case for differentiated carbon pricing-lower initial prices with pre-announced escalation schedules-rather than immediate adoption of international-level carbon prices.

Limitations of the Theoretical Approach

This paper presents a purely theoretical framework, and several limitations must be explicitly acknowledged. First, the model lacks empirical parameterization. The abatement cost functions $C_i(a_i)$, technology adoption parameters (F_i, μ_i), and behavioral parameters for SOE agents require firm-level data that is not publicly available for Uzbekistan's industrial sector. Without facility-level emissions data, monitoring-reporting-verification (MRV) system outputs, or

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 4, April, 2026
Website: usajournals.org

detailed technology cost assessments, the quantitative predictions of the model remain illustrative rather than calibrated.

Second, the MDP formulation assumes stationarity of the transition function, whereas real carbon markets exhibit non-stationary dynamics driven by policy changes, technological breakthroughs, and macroeconomic shocks. The single-agent MDP formulation for each firm does not fully capture the game-theoretic interactions among firms-a multi-agent RL (MARL) extension using stochastic games (Zhang et al., 2021) would be more appropriate but substantially increases computational complexity.

Third, the model does not endogenize the political economy of cap-setting. The regulator's choice between strict and loose caps is modeled as a strategic decision within the game, but the deeper political constraints-including energy subsidy reform sequencing, SOE governance reform, and international climate diplomacy-operate outside the model's boundaries.

Fourth, the conversion of Uzbekistan's intensity-based NDC target to an absolute emissions cap requires GDP growth projections that introduce substantial uncertainty. An intensity-based trading system (similar to China's output-based ETS design) may be more appropriate but introduces different incentive properties.

Fifth, the framework treats each major installation as a single agent, abstracting from intra-firm decision-making complexities within large state conglomerates where investment decisions may involve multiple government ministries, regional authorities, and enterprise management with potentially conflicting objectives.

Future Research: Computational Implementation

The theoretical framework developed here is designed for computational implementation as a critical next step. The proposed simulation architecture follows the two-level deep multi-agent RL design of Zheng et al. (2022):

Inner loop (Industrial agents): Each firm trains a policy network $\pi_i(a \mid s; \theta_i)$ using a hybrid actor-critic algorithm (Fan et al., 2019) to handle the mixed continuous-discrete action space. The continuous policy head outputs ϕ_t^i (green

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 4, April, 2026
Website: usajournals.org

investment proportion) as a Beta-distributed random variable, while the discrete policy head outputs Δ_t^i (trading volume) as a Categorical distribution conditioned on ϕ_t^i and state s_t^i. Training uses Proximal Policy Optimization (Schulman et al., 2017) for stability in the multi-agent setting.

Outer loop (Regulator): The regulator agent trains on a slower timescale, updating its policy $\pi^{GR}\left(\bar{E}, \alpha, f \mid \mathcal{S}_{global}; \theta_{GR}\right)$ after observing the equilibrium behavior of industrial agents under each policy configuration. Structured curriculum learning (Zheng et al., 2022) is recommended to stabilize the bi-level optimization.

Empirical parameterization priorities include: (a) obtaining facility-level emissions data through Uzbekistan's State Ecology Committee or World Bank Partnership for Market Implementation (PMI) program; (b) estimating sectoral marginal abatement cost curves using proxy data from Kazakhstan's ETS and China's pilot carbon markets; (c) calibrating SOE behavioral parameters using institutional analysis of Uzbekistan's state enterprise governance structure; and (d) validating the model against Kazakhstan's ETS outcomes as an out-of-sample test.

Conclusion

This paper establishes that the standard toolkit of competitive carbon market theory is insufficient for Uzbekistan's industrial structure, where a handful of state-affiliated conglomerates dominate emissions. The ABM-RL framework proposed here addresses this gap by providing a formal architecture in which bounded-rational agents learn strategic behaviors within a cap-and-trade institution, while a regulatory agent simultaneously optimizes market design parameters.

Three findings carry immediate policy relevance. The green investment switching threshold (Theorem 1) reveals that Uzbekistan's large enterprise scale paradoxically lowers the carbon price needed to trigger technology adoption-but only if the price signal is credible over a multi-year horizon, which requires institutional commitment mechanisms that developing economies often lack.

The payoff matrix analysis demonstrates that the Nash equilibrium in the regulator–industry game depends critically on penalty rate calibration: penalties must substantially exceed expected permit prices to make compliance dominant, yet excessively high penalties on state-owned enterprises create fiscal circularity. The allocation regime comparison shows that neither pure grandfathering nor pure auctioning is optimal for Uzbekistan-a hybrid approach combining auctioning for the non-traded power sector with output-based benchmarking for EITE sectors (metals, chemicals, cement) best balances efficiency, leakage protection, and political feasibility.

The distributional analysis reveals a deeper structural tension. Monotown dependence on carbon-intensive employers creates a political economy trap: the very communities most affected by carbon pricing are those least able to absorb transition costs, generating political pressure that can undermine regulatory credibility before the carbon price reaches the green investment threshold. Breaking this trap requires embedding just transition financing directly into the carbon market mechanism-channeling auction revenues to retraining programs and economic diversification in Almalyk, Navoi, Angren, and Bekabad-rather than treating distributional concerns as separate from market design.

The framework's most significant limitation is its purely theoretical nature. Without empirical parameterization from facility-level emissions data, technology cost assessments, and behavioral calibration of SOE decision-making, the model's quantitative predictions remain illustrative. Computational implementation using multi-agent reinforcement learning, validated against Kazakhstan's ETS experience as a regional benchmark, represents the essential next step toward actionable policy guidance. Uzbekistan's engagement with the World Bank's Partnership for Market Implementation creates a timely opportunity to generate the empirical foundations this framework requires.

References

1. Arthur, W. B. (2006). Out-of-equilibrium economics and agent-based modeling. In K. L. Judd & L. Tesfatsion (Eds.), Handbook of Computational Economics (Vol. 2, pp. 1551–1564).

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 4, April, 2026
Website: usajournals.org

2. Babiker, M. H. (2005). Climate change policy, market structure, and carbon leakage. Journal of International Economics, 65(2), 421–445.
3. Bauer, N., Bertram, C., Schultes, A., Klein, D., Luderer, G., Kriegler, E., McCollum, D. L., & Edenhofer, O. (2020). Quantification of an efficiency–sovereignty trade-off in climate policy. Nature, 588(7837), 261–266.
4. Becker, G. S. (1968). Crime and punishment: An economic approach. Journal of Political Economy, 76(2), 169–217.
5. Böhringer, C., Balistreri, E. J., & Rutherford, T. F. (2012). The role of border carbon adjustment in unilateral climate policy: Overview of an Energy Modeling Forum study (EMF 29). Energy Economics, 34(Supplement 2), S97–S110.
6. Calvano, E., Calzolari, G., Denicolò, V., & Pastorello, S. (2020). Artificial intelligence, algorithmic pricing, and collusion. American Economic Review, 110(10), 3267–3297.
7. Chappin, E. J. L., & Dijkema, G. P. J. (2010). On the impact of CO2 emission-trading on power generation emissions. Technological Forecasting and Social Change, 77(6), 999–1010.
8. Cramton, P., & Kerr, S. (2002). Tradeable carbon permit auctions: How and why to auction not grandfather. Energy Policy, 30(4), 333–345.
9. Fan, Z., Su, R., Zhang, W., & Yu, J. (2019). Hybrid actor-critic algorithm for parameterized action space. Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI-2019), 2279–2285.
10. Farmer, J. D., & Foley, D. (2009). The economy needs agent-based modelling. Nature, 460(7256), 685–686.
11. Fischer, C., & Fox, A. K. (2007). Output-based allocation of emission permits for mitigating tax and trade interactions. Land Economics, 83(4), 575–599.
12. Goulder, L. H. (1995). Environmental taxation and the double dividend: A reader's guide. International Tax and Public Finance, 2(2), 157–183.
13. Hahn, R. W. (1984). Market power and transferable property rights. Quarterly Journal of Economics, 99(4), 753–765.

14. Han, Z., Liu, Y., & Wang, J. (2025). Agent-based simulation of carbon market dynamics in China: Strategic behavior under heterogeneous agent interactions. Energy Economics, 131, 107345.
15. Kornai, J. (1986). The soft budget constraint. Kyklos, 39(1), 3–30.
16. Malik, A. S. (1990). Markets for pollution control when firms are noncompliant. Journal of Environmental Economics and Management, 18(2), 97–106.
17. Masson, S., Denoyer, L., Muller, P., & Gallinari, P. (2016). Reinforcement learning with parameterized actions. arXiv preprint arXiv:1610.01626.
18. Montgomery, W. D. (1972). Markets in licenses and efficient pollution control programs. Journal of Economic Theory, 5(3), 395–418.
19. Narassimhan, E., Gallagher, K. S., Koester, S., & Alejo, J. R. (2018). Carbon pricing in practice: A review of existing emissions trading systems. Climate Policy, 18(8), 967–991.
20. Puterman, M. L. (1994). Markov Decision Processes: Discrete Stochastic Dynamic Programming. John Wiley & Sons.
21. Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347.
22. Sijm, J., Neuhoff, K., & Chen, Y. (2006). CO2 cost pass-through and windfall profits in the power sector. Climate Policy, 6(1), 49–72.
23. Simon, H. A. (1955). A behavioral model of rational choice. Quarterly Journal of Economics, 69(1), 99–118.
24. Stavins, R. N. (1995). Transaction costs and tradeable permits. Journal of Environmental Economics and Management, 29(2), 133–148.
25. Stranlund, J. K., & Dhanda, K. K. (1999). Endogenous monitoring and compliance with a transferable emissions permit system. Journal of Environmental Economics and Management, 38(3), 267–282.
26. Sutton, R. S., & Barto, A. G. (2018). Reinforcement Learning: An Introduction (2nd ed.). MIT Press.
27. Tang, B.-J., Shen, C., & Gao, C. (2013). The efficiency analysis of the European CO2 futures market. Applied Energy, 112, 1544–1547.
28. Tesfatsion, L. (2006). Agent-based computational economics: A constructive approach to economic theory. In L. Tesfatsion & K. L. Judd (Eds.),

Modern American Journal of Business, Economics, and Entrepreneurship
ISSN (E): 3067-7203
Volume 2, Issue 4, April, 2026
Website: usajournals.org

Handbook of Computational Economics, Volume 2: Agent-Based Computational Economics (pp. 831–880).
29. UNFCCC. (2021). Uzbekistan's Nationally Determined Contribution (NDC). United Nations Framework Convention on Climate Change.
30. Vona, F., Marin, G., Consoli, D., & Popp, D. (2018). Environmental regulation and green skills: An empirical exploration. Journal of the Association of Environmental and Resource Economists, 5(4), 713–753.
31. World Bank. (2022). Uzbekistan country climate and development report. World Bank Group.
32. World Resources Institute. (2023). Climate Watch: Historical GHG Emissions — Uzbekistan. World Resources Institute.
33. Zhang, K., Yang, Z., & Basar, T. (2021). Multi-agent reinforcement learning: A selective overview of theories and algorithms. In K. G. Vamvoudakis, Y. Wan, F. L. Lewis, & D. Cansever (Eds.), Handbook of Reinforcement Learning and Control (pp. 321–384).
34. Zhang, Y.-J., & Wei, Y.-M. (2010). An overview of current research on EU ETS: Evidence from its operating mechanism and economic effect. Applied Energy, 87(6), 1804–1814.
35. Zheng, S., Trott, A., Srinivasa, S., Parkes, D. C., & Socher, R. (2022). The AI Economist: Taxation policy design via two-level deep multiagent reinforcement learning. Science Advances, 8(18), eabk2607.