12-2019

Optimal control for transboundary pollution under ecological compensation: A stochastic differential game approach

Ke JIANG

Ryan Knowles MERRILL
Singapore Management University, ryanmerrill@smu.edu.sg

Daming YOU

Pan PAN

Citation

JIANG, Ke; MERRILL, Ryan Knowles; YOU, Daming; and PAN, Pan. Optimal control for transboundary pollution under ecological compensation: A stochastic differential game approach. (2019). *Journal of Cleaner Production*. 241, 1-17.
Available at: https://ink.library.smu.edu.sg/lkcsb_research/6583

Published in Journal of Cleaner Production, Dec 2019, 241, Article 118391, 1-17.
https://doi.org/10.1016/j.jclepro.2019.118391

Optimal control for transboundary pollution under ecological compensation: A stochastic differential game approach

Ke Jiang [a, e], Ryan Merrill [b], Daming You [c], Pan Pan [c, *], Zhendong Li [d, **]

[a] Business School, Nanjing University of Information Science & Technology, Nanjing, 210044, China
[b] SMU Lee Kong Chian School of Business, Singapore Management University, Singapore, 178899, Singapore
[c] School of Business, Central South University, Changsha, 410083, China
[d] School of Management, Hangzhou Dianzi University, Hangzhou, 310018, China
[e] Development Institute of Jiangbei New Area, Nanjing University of Information Science & Technology, Nanjing, 210044, China

Keywords:
Transboundary pollution
Stochastic differential game
Ecological compensation
Optimal control strategies
Pollution governance investment

ABSTRACT

To account for previously ignored, yet widely observed uncertainty in nature's capability to replenish the natural environment in ways that should inform ideal design of ecological compensation (EC) regimes, this study constructs a stochastic differential game (SDG) model to analyze transboundary pollution control options between a compensating and compensated region. Equilibrium strategies in the stochastic, two player game inform optimal control theory and reveal a welfare distribution mechanism to form the basis of an improved cooperative game contract. A case-based numerical example serves to verify the theoretical results and supports three key insights. First, accounting for various random disturbance factors, the probabilistic pollutant stock in Stackelberg non-cooperative game exceeds that of a cooperative game situation. Second, the EC mechanism provides long-term, effective incentives only when the marginal losses of environmental damage in the compensating region are more than twice that of the compensated region. Achieving a Pareto optimal equilibrium relies upon the attainment of a dynamic allocation ratio derived from the analysis of a robust welfare allocation mechanism. Third, cross-region cooperation reliably outperforms Stackelberg non-cooperation due to either overwhelming incumbent economic interests or high abatement costs. This study illuminates the importance of balancing both parties' interests within an EC agreement while reducing uncertainty around unobserved environmental factors during ex-ante negotiations.

1. Introduction

Transboundary environmental pollution has drawn increasing interest from scholars in recent decades with rising awareness of the impacts and complexity of transnational and transregional environmental spillovers (Mäler and Zeeuw, 1998; Zhang et al., 2017; Chang et al., 2018). Discharge problems driven by pollution mobility has inspired an emerging focus on control regimes oriented to improve total welfare irrespective of administrative limits by attending to encompassing, transregional outcomes. An objective focus on mitigating both externalities and 'free riding' actions within a cross-administrative pollution context calls for mechanisms to make local governments' motivations to control contamination contingent on the interests of their "downstream" peers (Bardhan, 2002; Kahn et al., 2015).

In China, heightened public attention to policy failures in controlling environmental degradation has inspired a variety of environmental control regimes. Amid frequent occurrence of transboundary pollution events such as large-areas haze, water disputes and forest destruction (Xu et al., 2014; Liu et al., 2016a; Gu and Yim, 2016), design specifics for ecological supervision mechanisms have captured increasing attention nationwide (Wang and Shen, 2016). Experience with a continuous flow of pollutants between regions has revealed the limited effectiveness of single-jurisdiction solutions, and ecological compensation (EC) mechanisms of various types have emerged as a promising regulatory tool with which to align the interests of transregional communities.

* Corresponding author. School of Business, Central South University, Changsha, 410083, China.
** Corresponding author. School of Management, Hangzhou Dianzi University, Hangzhou, 310018, China.
E-mail addresses: jiangke@nuist.edu.cn (K. Jiang), ryanmerrill@smu.edu.sg (R. Merrill), youdaming2001@163.com (D. You), panpan0703@csu.edu.cn (P. Pan), lizhendong@hdu.edu.cn (Z. Li).

With rising interest in EC design, regulators and scholars exhibit increasing interest in tools to transfer the right to control transboundary pollution from the local level to upper-levels of government regulation (Zhao et al., 2013). Policy proposals include EC regimes supported by national horizontal transfer payments to compensate for pollution control expenditures (Yu and Xu, 2016; Liu et al., 2018) and calibrations to ensure that compensation intensity hinges on the scale and scope of external benefits (Van den Bergh, 2010; Guan et al., 2016). As a promising mechanism for aligning stakeholders' interests through economic means to maintain or improve the broader ecosystem, EC is being rapidly adopted across China and emerging as a subject of intense experimentation and learning.

Since the introduction of EC in the 1980s in China, studies on EC have sought to inform a variety of theoretical discussions and application practices (Pan et al., 2017; Shang et al., 2018; Wang et al., 2019). Game theory has served to illuminate the relationship between compensation stakeholders and informed several recent advances. In a novel application by Zhao et al. (2012a), a transfer tax model employs the typical Stackelberg game between the government and individual areas to examine an optimal transfer tax rate. Game theory aids the author in determining the optimal calibration of the rate to strengthen regional pollution reduction cooperation and minimize the cost of pollution reduction across a shared basin.

In other work, Huang et al. (2011) conduct a static Bayesian game model of one compensator and one compensated party under complete information. The authors then divide the enforcement of the compensator to implement environmental policy into two types, i.e., robust and weak, and explore the influence of the strategy choice of both parties on the size of EC under incomplete information. Additional studies by Xu et al. (2012) and Jie et al. (2012) employ an evolutionary game model to explore changing outcomes amid conflicted interests in watershed EC. The former emphasizes that the optimal solution of the evolutionary stability strategy cannot be achieved solely by upstream and downstream governments, such that basin-level EC policy requires credible intervention by the central government. The latter argues that reasonable determination of an efficient compensation level can support the evolution of EC mechanism towards a cooperative equilibrium for a variety of shared water sources.

Despite recent advances in using game theory to improve EC design, recent efforts have generally neglected to account for a variety of random disturbances in the game environment that could prove critical to outcomes (Jørgensen, 2010; Van Long, 2011; Shi et al., 2016). In real life, strategic choice attends to both the estimated quantum and timing of expected payment values (Zhao et al., 2012b) - outcomes regularly influenced by uncertain and often volatile changes in the external environment (Dellink and Finus, 2012). These may range from changes in political and cultural environments to shifts in industry performance and market structure (Yin et al., 2017). Dynamic human factors also shape the choices of decision-makers (game players) (Aklin and Urpelainen, 2013) via perceptions of psychological benefit, desire for knowledge, and novelty seeking (Lin and Huang, 2012; Czaika and Selin, 2017). Accounting for such factors requires improved method.

Generally speaking, industrial processes create two types of negative spillovers causing pollution accumulation within and across regional boundaries. First, pollutants emitted via industrial production cause a short-term local impact on the immediate environs. Second, the diffusion of pollutants, as by atmospheric motion or watershed flows, drives a transfer of effluents beyond local boundaries to accumulate in outlying regions. Many of the underlying processes are complex and difficult for decision makers to anticipate, stochastic interference factors that result in significant uncertainties in optimal equilibrium outcomes of governance mechanisms. To improve traction in this decision space, scholars have called for the development of dynamic game solutions that contend with stochastic factors (Haurie et al., 1994; Yeung, 2001; Yeung and Petrosyan, 2005).

In seminal work, Yeung (2007) employs a stochastic differential game (SDG) framework to study pollution management. Yeung's work is the first to derive time consistent solutions in a cooperative differential game environment focused on pollution control outcomes in which industries and governments constitute separate entities. A key feature of the resulting game model lies in the finding that industrial sectors remain competitive among themselves while governments cooperate in pollution abatement. Based on this model, Daskalakis et al. (2009) demonstrate the efficient emission permit price should exhibit a stochastic character, resulting from interrelated dynamics of variable scarcity of emission permits and market discipline. Funke and Paetz (2011) extend these efforts to construct a robust control framework for application to regulatory regimes to mitigate global warming. Yet despite the well-recognized potential of a SDG model to advance the study of transboundary spillovers, applications remain scarce. Perhaps this is due to inherent complexities in deriving tractable assessments of random interference factors (Yeung and Petrosyan, 2008), such as waste diffusion in river systems or air particulates or noise pollution generated by production (Dong et al., 2013; Pamen, 2015; Xu and Lin, 2016). To improve the observation of uncertain yet relevant factors, this study explores the impact of an EC mechanism on transboundary pollution within a robust SDG game.

The following queries guide the development and application of the resulting model. First, how does the introduction of a volatile, random factor interfere with the dynamic evolution of the pollutant stock? Second, how does this dynamic alter the efficient setting of the compensation level for a given level of emissions? Third, what strategies emerge when two regions become able to account for the future impact of their decisions governing emission reduction on alternative game contracts?

To explore these questions, an SDG game model of transboundary pollution must at a minimum encompass one compensated region and one compensating region under an EC mechanism. Random interference factors, such as topography, meteorological conditions, and river flows should shape outcomes from both Stackelberg non-cooperative and cooperative game contracts. Subsequently, an optimal control theory based on the maximization of the net present value of welfare outcomes should help explore the impact of several critical game elements on outcome states. These may include the respective regions' emission capacities, their investment efforts in pollution governance, and some exchange coefficient. In establishing Pareto optimality, an analysis of a welfare distribution mechanism should further elucidate the ideal scope of the welfare distribution coefficient to inform an optimal contract. Finally, the game theoretical modeling effort should be verified through a numerical illustration within a case-based example to ground the theoretical results and provide practical recommendations for regulatory policy.

In fulfilling this plan of attack, this study makes a series of significant contributions to the research of game theory and environmental management. First, this study directly addresses a gap in the literature left by a lack of work examining the optimization of transboundary pollution control through dynamic EC mechanisms. This study expands the literature on transboundary pollution control by optimizing a robust compensation framework under the premise of reasonable compensation and a volatile environment. In doing so, this study takes an important step towards overcoming a critical difficulty in establishing robust and political sustainable EC mechanisms (Liu et al., 2018).

Second, this study deviates from the leading research on transboundary pollution that focuses primarily on the dynamic model of deterministic pollutant stocks (Benchekroun and Martín-Herrán, 2016; El Ouardighi et al., 2018; Vardar and Zaccour, 2018). Here, a SDG model guarantees the stability of game equilibrium results and bridges a methodological gap between SDG applications and the traditional differential games that prevail in transboundary pollution scholarship. The findings inform practical policy prescriptions for overcoming uncertainties and contract selection problems for effective pollution control.

Third, in cross-walking the SDG simulation model to a real-world case, the analysis verifies the theoretical results and sheds light on sufficient conditions under which EC proves an effective means to coordinate the balance of interests between the compensating and the compensated regions.

To proceed, Section 2 next details methods governing the establishment of the SDG model and its two game contracts along with a comparative analysis of key equilibrium results. Section 3 reports the design of a general allocation mechanism that promises to be welfare improving. Section 4 then advances a numerical illustration, while Section 5 and 6 discuss key results and conclusions.

2. Methodology

2.1. Stochastic differential game model

SDG modeling serves to account for randomly dynamic yet relevant factors within the two-player strategic game between a compensating region and a compensated region operating under the jurisdiction of a central government. Both regions share positive spillovers of environmental services sustained by either party through investments in pollution control. This outcome follows from the fact that environmental quality exhibits public goods attributes that cannot be exclusively consumed by the producing party. The compensating region generates trans-regional pollution transmission problems in continuous time. No further administrative affiliation exists between the two principal actors.

Under equilibrium, returns to environmental quality in the compensated region exceed that of the compensator, who conversely faces higher relative returns to industrial development. These conditions give compensators lower underlying incentives to sacrifice industrial gains for marginal improvements in environmental quality. Hereafter, under the premise of satisfying the established situation of both parties, the compensator needs to share the additional cost thus increased in a timely and sufficient amount, i.e., EC, which is represented by $\varepsilon(t)$ and meets $0 \le \varepsilon(t) \le 1$. Table 1 reports the key decision variables, player-specific variables, and game parameters (both fixed and stochastic). Throughout, $h = i,j$ are subscripts of the compensating region and compensator, respectively.

Assumption 1. Distinct from previous researches (Yeung, 2007; Li, 2014), which disregards or abstracts away trade-offs between pollution and regional environmental quality, the present study considers regional governments as rational economic entities seeking improvements in both economic and environmental utility. Therefore, this study integrates a regional utility function in two key aspects.

First, this study takes the output $Q_h(t)$ of industrial processes to partially reflect total regional economic development and to be positively correlated with the quantity of emissions $q_h(t)$, i.e., $Q_h(t) = Q_h(q_h(t))(h = i,j)$. Following List and Mason (2001), as well as Breton et al. (2005, 2006), the resulting utility function of industrial development $R_h(Q_h)$ is characterized by the instantaneous emission $q_h(t)$ and satisfies the law of diminishing marginal utility:

Table 1
Notations and definitions.

Decision Variables	
$q_h(t)$	Emission Quantity of region h
$\varepsilon(t)$	Ecological Compensation Coefficient: Rate (%) Setting of the Transfer Payment
$u_h(t)$	Mitigation Effort: Quantity of Environmental Governance of region h
Player-specific Variables	
a_h	Utility Coefficient of Industrial Development of region h
$\varDelta_h$	Utility Coefficient of Environmental Quality of region h
δ_h	Environmental Loss Factor: Determining Damage to Natural Environment of region h
$\overline{q}_h$	Pollution Threshold: Upper Limit of Emissions of region h
κ_h	Cost Coefficient of Mitigation Effort of region h
ϖ_h	Immobile Portion of Local Emissions of region h
$1-\varpi_h$	Transboundary Spillover out of region h
State Variables and Game Parameters	
$\tau(t)$	Total Inter-regional Pollutant Stock at time t, with $\tau(0) = \tau_0 \ge 0$
$\forall$	Inter-regional Difference in Marginal Industrial Productivity
λ	Reduction to Pollution Stock due to Ecosystem Diffusion
α	Marginal Impact Coefficient of Emissions Quantities
β	Marginal Impact Coefficient of Mitigation Effort
η	Self-purification rate
ρ	Discount rates
Random Interference	
$\sigma(\tau(t))$	Random Interference Factors
$P(t)$	Standard Wiener Process

$$R_h(Q_h(t)) = \ln Q_h(t) = \ln(a_h q_h) \tag{1}$$

where $0 < a_h < 1$ represents the utility coefficient of industrial development and meets $0 < q_h(t) < a_h$. Let $a_i = a$ and $a_j = \forall a$, in which $\forall$ is a positive constant that measures the marginal difference between the industrial productivity of the two neighboring regions.

Second, this study recognizes that each region possesses an upper limit on the total amount of pollutants $\overline{q}_h(t)$, i.e., a threshold of emissions that may be sustained by the region h before spilling over to the neighboring territory. For an output of $q_h(t)$ produced by region h, some portion will be retained locally $\varpi_h q_h(t)$ to region h itself, where ϖ_h denotes the local portion. Simultaneously, region h receives the remaining short-term transboundary spillover, denoted as $1 - \varpi_h(t)$. Hereafter, let $Z_h(t)$ be the utility function of environmental quality, characterizing the reverse linear relationship between regional environmental quality and emissions quantities:

$$\begin{cases} Z_i(t) = \varDelta_i\left[\overline{q}_i(t) - \varpi_i q_i(t) - (1-\varpi_j)q_j(t)\right] \\ Z_j(t) = \varDelta_j\left[\overline{q}_j(t) - \varpi_j q_j(t) - (1-\varpi_i)q_i(t)\right] \end{cases} \tag{2}$$

where $\varDelta_h$ is the utility coefficient of environmental quality of region h that reveals the region's preference for environmental protection.

Assumption 2. Similar to numerous works on the costs of environmental governance (Breton et al., 2008; El Ouardighi et al., 2016), let $I_h(t)$ represents a mitigation function satisfying the law of rising marginal costs:

$$I_h(t) = \kappa_h u_h^2(t) \tag{3}$$

where $u_h(t)$ denotes the mitigation effort of region h at time t. The cost coefficient associated with the mitigation effort is denoted by $\kappa_h > 0$, and describes the rising resource expense (of capital, human attention, time, etc) a government must incur to achieve an additional unit of clean up.

Assumption 3. Practically, the cumulative process of pollutants generally involves the interaction between natural environment and pollutants (Bertinelli et al., 2014), and accompanied by certain random elements. For instance, the absorption of natural environment to pollutants, the mutual transmission of pollutants between regions and the impacts exerted by unpredictable factors (Kolstad, 2007; Athanassoglou and Xepapadeas, 2012; Masoudi et al., 2016), inclusive of weather, natural disasters and human damages. To process such uncertainties, a random term is added to a standard deterministic differential game to produce a stochastic differential equation (Jørgensen and Yeung, 1996):

$$\begin{cases} d\tau(t) = \left[\alpha\left(q_i + q_j\right) - \beta(u_i + u_j) - (\lambda + \eta)\tau\right]dt + \sigma(\tau(t))dP(t) \\ \tau(0) = \tau_0 \geq 0 \end{cases} \tag{4}$$

where $\sigma(\tau(t))$ represents the random interference factor and $P(t)$ refers a standard Wiener process. As such, $\sigma(\tau(t))dP(t)$ describes disturbances caused by random influences such as topography, meteorological conditions and river flows. Meanwhile, $\alpha > 0$ and $\beta > 0$ measure the marginal impact coefficients of emissions and mitigation efforts to changes in the pollutant stock. $\lambda > 0$ characterizes the impact exerted by ecosystem diffusion on the pollutant stock, and $\eta > 0$ is the natural self-purification rate (e.g. sequestration). The pollutant stock in initial planning stage is established as τ_0.

Assumption 4. At the current level of pollution control, pollutants in regional production and consumption will cause diverse damages to environmental functions, human health and crop yield. As such, supposing that this damage cost is a linear function of the current pollutant stock, characterizing the marginal loss of environmental quality stemmed from the unit pollution, also known as environmental degradation costs, denoted in δ_h. To simplify the calculation, this study assumes that the above environmental degradation costs could be overall measured by currency and such costs will merely cut the revenue of regional government down, without considering the real losses it brings to the entire social welfare.

In sum, the two governments are assumed to make rational decisions in line with the complete information and each aims to maximize the expectant welfare of their decision-making systems in an infinite time interval. As such, the welfare functions of two neighboring regions at time t form:

$$\begin{cases} \Pi_i(t) = \int_0^\infty e^{-\rho t}\{R_i(Q_i) + Z_i(t) - (1 - \varepsilon(t))I_i(t) - \delta_i\tau(t)\}dt \\ \Pi_j(t) = \int_0^\infty e^{-\rho t}\{R_j(Q_j) + Z_j(t) - I_j(t) - \varepsilon(t)I_i(t) - \delta_j\tau(t)\}dt \\ \Pi(t) = \int_0^\infty e^{-\rho t}\{R_i(Q_i) + R_j(Q_j) + Z_i(t) + Z_j(t) - I_i(t) - I_j(t) - (\delta_i + \delta_j)\tau(t)\}dt \end{cases} \tag{5}$$

where ρ is the discount rate, common to two regions.

2.2. Non-cooperative outcomes

Within a Stackelberg non-cooperative game, neither region adopts a strategy of environmental cooperation, yet the outcome of each party creates a substantial impact on the other. Thus, both sides of the game consider the other's possible reaction when making decisions. Eq. (6) describes a two-stage Stackelberg non-cooperative game model formed by the compensating region i and compensated region j. Subscript D is adopted to obtain a SDG model in which region h seeks to:

$$\begin{cases} \max\limits_{q_i,u_i} \Pi_i^D \int_0^\infty e^{-\rho t}\left\{\ln[aq_i(t)] + \Delta_i\left[\overline{q}_i(t) - \varpi_i q_i(t) - (1 - \varpi_j)q_j(t)\right] - \kappa_i[1 - \varepsilon(t)]u_i^2(t) - \delta_i\tau(t)\right\}dt \\ \max\limits_{q_j,\varepsilon,u_j} \Pi_j^D \int_0^\infty e^{-\rho t}\left\{\ln\left[\forall aq_j(t)\right] + \Delta_j\left[\overline{q}_j(t) - \varpi_j q_j - (1 - \varpi_i)q_i(t)\right] - \kappa_j(t)u_j^2(t) - \kappa_i\varepsilon(t)u_i^2(t) - \delta_j\tau(t)\right\}dt \end{cases} \tag{6}$$

$$\text{s. t.} \begin{cases} d\tau(t) = \left[\alpha\left(q_i + q_j\right) - \beta(u_i + u_j) - (\lambda + \eta)\tau(t)\right]dt + \sigma(\tau(t))dP(t) \\ \tau(0) = \tau_0 \geq 0 \end{cases}$$

2.2.1. Stackelberg non-cooperative game solutions

Proposition 1. *The optimal emission quantity and mitigation effort of two regions, as well as the optimal EC coefficient are given by:*

$$q_i^{D*} = \frac{\rho+\lambda+\eta}{\Delta_i \varpi_i(\rho+\lambda+\eta)+\alpha\delta_i};\quad q_j^{D*} = \frac{\rho+\lambda+\eta}{\Delta_j \varpi_j(\rho+\lambda+\eta)+\alpha\delta_j};$$

$$u_i^{D*} = \frac{\beta(\delta_i+2\delta_j)}{4\kappa_i(\rho+\lambda+\eta)};\quad u_j^{D*} = \frac{\beta\delta_j}{2\kappa_j(\rho+\lambda+\eta)};\quad \varepsilon^* = \begin{cases} \dfrac{2\delta_j-\delta_i}{2\delta_j+\delta_i}, & 2\delta_j \ge \delta_i \\ 0, & 2\delta_j < \delta_i \end{cases}; \tag{7}$$

and the optimal welfare functions for two regions are given by:

Proof. *See Appendix A.*

$$\begin{cases} \Pi_i^{D*} = \ln\left[\dfrac{a(\rho+\lambda+\eta)}{\Delta_i\varpi_i(\rho+\lambda+\eta)+\alpha\delta_i}\right] + \Delta_i\left[\overline{q}_i(t) - \dfrac{\varpi_i(\rho+\lambda+\eta)}{\Delta_i\varpi_i(\rho+\lambda+\eta)+\alpha\delta_i} - \dfrac{(1-\varpi_j)(\rho+\lambda+\eta)}{\Delta_j\varpi_j(\rho+\lambda+\eta)+\alpha\delta_j}\right] \\ \qquad - \dfrac{2\delta_i\beta^2(2\delta_j+\delta_i)}{16\kappa_i(\rho+\lambda+\eta)^2} - \delta_i\tau \\ \Pi_j^{D*} = \ln\left[\dfrac{\forall a(\rho+\lambda+\eta)}{\Delta_j\varpi_j(\rho+\lambda+\eta)+\alpha\delta_j}\right] + \Delta_j\left[\overline{q}_j(t) - \dfrac{\varpi_j(\rho+\lambda+\eta)}{\Delta_j\varpi_j(\rho+\lambda+\eta)+\alpha\delta_j} - \dfrac{(1-\varpi_i)(\rho+\lambda+\eta)}{\Delta_i\varpi_i(\rho+\lambda+\eta)+\alpha\delta_i}\right] \\ \qquad - \dfrac{\beta(2\delta_j-\delta_i)}{16\kappa_i(\rho+\lambda+\eta)^2} - \dfrac{\beta^2\delta_j^2}{4\kappa_j(\rho+\lambda+\eta)^2} - \delta_j\tau \end{cases} \tag{8}$$

Remark 1

(1) Under the Stackelberg non-cooperative game, both a higher discount factor and a higher natural self-purification rate increase the optimal combined emission quantity of the two regions ($\partial q_h^{D*}/\partial\eta > 0, \partial q_h^{D*}/\partial\rho > 0$). Conversely, the combined emission quantity will be much lower where the emissions cause very different levels of damage between regions ($\partial q_h^{D*}/\partial\delta_h < 0$). Here, total emissions are further reduced by increased levels of respective immobility ($\partial q_h^{D*}/\partial\varpi_h < 0$). As transboundary spillovers rise under non-cooperation, so too, unsurprisingly, does total pollution.

(2) Both the impact coefficient of mitigation effort on pollutant stock and the environmental losses positively influence the optimal mitigation efforts of both sides in the game ($\partial u_h^{D*}/\partial\beta > 0, \partial u_h^{D*}/\partial\delta_h > 0$). However, total combined mitigation efforts are negatively correlated with the impact coefficient of mitigation effort on environmental quality, the discount factor, and the self-purification rate ($\partial u_h^{D*}/\partial\kappa_h < 0$, $\partial u_h^{D*}/\partial\rho < 0$, $\partial u_h^{D*}/\partial\eta < 0$).

(3) The optimal EC coefficient is greatly affected by the environmental losses in two regions. Given $0 \le \varepsilon^* \le 1$, the implied condition is apparently attained as $0 \le 2\delta_j - \delta_i \le 1$. To be specific, when $2\delta_j > \delta_i$, the compensation ratio will positively affect the mitigation effort of region *i*, and subsequent analysis and numerical illustrations will be particularly conducted within this scope. When $2\delta_j < \delta_i$, region *j* will not provide compensation accordingly, and such condition is not consistent with the foregoing assumptions. The optimal compensation ratio attains its maximum when $2\delta_j = \delta_i$.

2.2.2. The limit of expectation and variance

The welfare of two region governments is correspondingly concerned with the dynamic change in pollutant stocks, and this process of change is obviously affected by various random disturbance factors. Hence, it is necessary to reveal the expectation and variance limits of pollutant stock.

Substituting the optimal equilibrium strategy Eq. (7) into Eq. (4) yields:

$$\begin{cases} d\tau(t) = [\Omega - (\lambda+\eta)\tau(t)]dt + \sigma(\tau(t))dP(t) \\ \tau(0) = \tau_0 \ge 0 \end{cases} \tag{9}$$

where $\Omega = \frac{\alpha(\rho+\lambda+\eta)}{\Delta_i\varpi_i(\rho+\lambda+\eta)+\alpha\delta_i} + \frac{\alpha(\rho+\lambda+\eta)}{\Delta_j\varpi_j(\rho+\lambda+\eta)+\alpha\delta_j} - \frac{\beta^2(\delta_i+2\delta_j)}{4\kappa_i(\rho+\lambda+\eta)} - \frac{\beta^2\delta_j}{2\kappa_j(\rho+\lambda+\eta)}$ is a constant, and the greater the pollutant stock, the larger Ω.

According to the stochastic analysis theory, $\sigma(\tau(t))dP(t) = \sigma\sqrt{\tau}dP(t)$ is further set, informing a definition of Proposition 2 below.

Proposition 2. *The limit of expectation and variance of pollutant stock in non-cooperative game feedback equilibrium meets the condition:*

$$E\left[\tau^D(t)\right] = \frac{\Omega}{\lambda+\eta} + \left(\tau_0 - \frac{\Omega}{\lambda+\eta}\right)e^{-(\lambda+\eta)t},\quad \lim_{t\to+\infty} E\left[\tau^D(t)\right] = \frac{\Omega}{\lambda+\eta} \tag{10}$$

$$S\left[\tau^D(t)\right] = \frac{\sigma^2\{\Omega - 2[\Omega-(\lambda+\eta)\tau_0]e^{-(\lambda+\eta)t} + [\Omega - 2(\lambda+\eta)\tau_0]e^{-2(\lambda+\eta)t}\}}{2(\lambda+\eta)^2}\qquad \lim_{t\to+\infty} S\left[\tau^D(t)\right] = \frac{\sigma^2\Omega}{2(\lambda+\eta)^2} \tag{11}$$

Proof. See Appendix B.

2.3. Cooperative arrangement

This section assumes that the adjoining two regions have reached the cooperative agreement with binding force in advance, with the goal of maximizing the sum of long-term net welfare values. By using the superscript C, the common dynamic optimal solution can be acquired through jointly optimizing the welfare function of two regions, i.e.,:

$$\max_{q_h(t),u_h(t)} \Pi^C \int_0^\infty e^{-\rho t}\left\{\begin{array}{l}\ln\left[aq_i(t)\right]+\ln\left[\forall aq_j(t)\right]+\varDelta_i\left[\overline{q}_i(t)-\varpi_i q_i(t)-(1-\varpi_j)q_j(t)\right]+\varDelta_j\left[\overline{q}_j(t)\right.\\ \left.-\varpi_j q_j(t)-(1-\varpi_i)q_i(t)\right]-\kappa_i u_i^2(t)-\kappa_j u_j^2(t)\Big]-(\delta_i+\delta_j)\tau(t)\Big]\end{array}\right\}dt \tag{12}$$

$$\text{s. t.}\ \begin{cases} d\tau(t)=\left[\alpha\left(E_i+E_j\right)-\beta\left(u_i+u_j\right)-(\lambda+\eta)\tau(t)\right]dt+\sigma(\tau(t))dP(t)\\ \tau(0)=\tau_0\geq 0\end{cases}$$

2.3.1. Cooperative game solutions

Proposition 3. *The optimal conditions are necessary and sufficient, including*:

$$q_i^{C*}=\frac{\rho+\lambda+\eta}{\varDelta_i\varpi_i(\rho+\lambda+\eta)+\varDelta_j(1-\varpi_i)(\rho+\lambda+\eta)+\alpha\left(\delta_i+\delta_j\right)};$$

$$q_j^{C*}=\frac{\rho+\lambda+\eta}{\varDelta_i(1-\varpi_j)(\rho+\lambda+\eta)+\varDelta_j\varpi_j(\rho+\lambda+\eta)+\alpha\left(\delta_i+\delta_j\right)};$$

$$u_i^{C*}=\frac{\beta\left(\delta_i+\delta_j\right)}{2\kappa_i(\rho+\lambda+\eta)};\quad u_j^{C*}=\frac{\beta\left(\delta_i+\delta_j\right)}{2\kappa_j(\rho+\lambda+\eta)} \tag{13}$$

and the optimal welfare functions for two regions are as follows:

$$\Pi^{C*}=\ln\left[\frac{a(\rho+\lambda+\eta)}{\varDelta_i\varpi_i(\rho+\lambda+\eta)+\alpha\delta_i}\right]+\varDelta_i\left[\overline{q}_i(t)-\frac{\varpi_i(\rho+\lambda+\eta)}{\varDelta_i\varpi_i(\rho+\lambda+\eta)+\alpha\delta_i}-\frac{(1-\varpi_j)(\rho+\lambda+\eta)}{\varDelta_j\varpi_j(\rho+\lambda+\eta)+\alpha\delta_j}\right]$$
$$-\frac{2\delta_i\beta^2\left(2\delta_j+\delta_i\right)}{16\kappa_i(\rho+\lambda+\eta)^2}-\delta_i\tau(t) \tag{14}$$

Proof. See Appendix C.

Remark 2

Under the cooperative game condition, equilibrium strategies in both regions are insulated from any change in EC coefficient. Meanwhile, regional damages in the two neighboring regions are the focal determinant of the optimal emission quantity. The influence of the additional factors on equilibrium strategies is generally similar as in the Stackelberg non-cooperative game.

2.3.2. The limit of expectation and variance

To explore the role of ex post expectations and ex ante variance in the pollutant stock in shaping strategies over the course of a cooperative game, the results of Eq. (13) can be substituted into Eq. (4) to yield:

$$\begin{cases} d\tau(t)=[\Upsilon-(\lambda+\eta)\tau(t)]dt+\sigma(\tau(t))dP(t)\\ \tau(0)=\tau_0\geq 0\end{cases} \tag{15}$$

where

$$\Upsilon=\frac{\alpha(\rho+\lambda+\eta)}{\varDelta_i\varpi_i(\rho+\lambda+\eta)+\varDelta_j(1-\varpi_i)(\rho+\lambda+\eta)+\alpha\left(\delta_i+\delta_j\right)}$$
$$+\frac{\alpha(\rho+\lambda+\eta)}{\varDelta_i(1-\varpi_j)(\rho+\lambda+\eta)+\varDelta_j\varpi_j(\rho+\lambda+\eta)+\alpha\left(\delta_i+\delta_j\right)}$$
$$+\frac{\alpha(\rho+\lambda+\eta)}{\varDelta_j\varpi_j(\rho+\lambda+\eta)+\alpha\delta_j}-\frac{\beta^2\left(\delta_i+\delta_j\right)}{2\kappa_i(\rho+\lambda+\eta)}-\frac{\beta^2\left(\delta_i+\delta_j\right)}{2\kappa_j(\rho+\lambda+\eta)}$$

is a constant, and the greater the pollutant stock, the largerΥ.

Subsequently,$\sigma(\tau(t))dP(t)=\sigma\sqrt{\tau}dP(t)$, foundation for defining Proposition 4 below.

Proposition 4. *The limit of expectation and variance of the pollutant stock in cooperative game feedback equilibrium satisfies*:

$$E\left[\tau^C(t)\right]=\frac{\Upsilon}{\lambda+\eta}+\left(\tau_0-\frac{\Upsilon}{\lambda+\eta}\right)e^{-(\lambda+\eta)t},\ \lim_{t\to\infty}E\left[\tau^C(t)\right]=\frac{\Upsilon}{\lambda+\eta} \tag{16}$$

$$S\left[\tau^C(t)\right]=\frac{\sigma^2\{\Upsilon-2[\Upsilon-(\lambda+\eta)\tau_0]e^{-(\lambda+\eta)t}+[\Upsilon-2(\lambda+\eta)\tau_0]e^{-2(\lambda+\eta)t}\}}{2(\lambda+\eta)^2}\qquad \lim_{t\to+\infty}S\left[\tau^C(t)\right]=\frac{\sigma^2\Upsilon}{2(\lambda+\eta)^2} \tag{17}$$

Proof*. See* Appendix D.

3. Welfare allocation mechanism

To guarantee a thorough and lasting cooperation, all parties must agree to re-allocate utility willingly through time. To this end, a robust welfare allocation mechanism should form the linchpin of a governance arrangement by eliminating instability in a cooperation alliance and supporting continuous cooperation. A welfare allocation mechanism is robust over time when it is able to satisfy both overall rationality and individual rationality. Overall rationality ensures all possible welfare improvements may be obtained through a cooperative alliance. Individual rationality requires that the gains of a cooperative strategy exceed those of a Stackelberg non-cooperative strategy. The mechanism is furthermore considered dynamic in its ability to do so at any time (Yeung, 1992). Within a SDG model, these conditions should hold when all parties agree to implement the cooperative strategy over the entire game.

Abiding by Rubinstein (1982) bargaining principle, governments in each region in time duration $[0, +\infty]$ not only acquire earnings without cooperation, but also gain additional value from the sharing cooperation with other parties. Assuming that the welfare of region i accounted for ξ in the overall welfare of the two parties under cooperative case, then the share for region j is $1-\xi$, and satisfies $0 \le \xi \le 1$.Therefore, individual rationality demands:

$$\xi \Pi^C \ge \Pi_i^D; \;\; (1-\xi)\Pi^C \ge \Pi_j^D \tag{18}$$

From Eq. (18), it can be deduced $\xi \in \left[\frac{\Pi_i^D}{\Pi^C}, \frac{\Pi^C - \Pi_j^D}{\Pi^C}\right]$. For $\xi_{\max} = \frac{\Pi^C - \Pi_j^D}{\Pi^C}$, $\xi_{\min} = \frac{\Pi_i^D}{\Pi^C}$, then in the welfare allocation of $\xi \in [\xi_{\min}, \xi_{\max}]$ and both sides of the game aim to seek greater payoffs. The discount factor in the Rubinstein bargaining model can be adopted to manipulate the welfare allocation ratio ξ. Let φ_i and φ_j denote the discount factors for two regions, i.e., their 'bargaining power' or 'patience level' such that these will satisfy conditions $0 \le \varphi_i \le 1$ and $0 \le \varphi_j \le 1$.

As the dominant party in the pollution control investment, region i takes the lead in the bargaining process. Applying the Rubinstein indefinite periodic bidding game, the only subgame that may be solved to a refined Nash equilibrium:

$$R = \frac{1-\varphi_j}{1-\varphi_i\varphi_j} \tag{19}$$

Combined with $\xi \in [\xi_{\min}, \xi_{\max}]$, the following optimal allocation ratio in two regions guarantees this requirement:

$$\overline{\xi} = \frac{1-\varphi_j}{1-\varphi_i\varphi_j}(\xi_{\max} - \xi_{\min}) + \xi_{\min} \tag{20}$$

As such, the total welfares allocated between the two regions will be sustained under a robust dynamic allocation mechanism when:

$$\begin{cases} \Pi_i^C = \dfrac{1-\varphi_j}{1-\varphi_i\varphi_j}\left[\Pi^C - \left(\Pi_i^D + \Pi_j^D\right)\right] + \Pi_i^D \\ \Pi_j^C = \dfrac{\varphi_j(1-\varphi_i)}{1-\varphi_i\varphi_j}\left[\Pi^C - \left(\Pi_i^D + \Pi_j^D\right)\right] + \Pi_j^D \end{cases} \tag{21}$$

Under forced cooperation, all possible allocations could practically satisfy the optimal objective of total collective welfare irrespective of the allocation mechanism. Yet, only this robust dynamic allocation mechanism - satisfying individual rationality — promises to ensure sustained cooperation and the attainment of Pareto optimality through time.

4. Numerical illustrations

The results presented thus far, while analytically compelling, may further benefit from practical grounding in a real-world example. This section thus assesses the results of the game theoretic models against a realistic backdrop of the first, inter-provincial watershed EC mechanism deployed in China. A numerical illustration serves to illustrate (i) how dynamic state variables shape expectations and variance in the stock of pollutants and (ii) how related and key parameter values affect equilibrium strategies of both sides of the game.

4.1. Xin'an River Basin — Anhui and Zhejiang Province, China

Xin'an River Basin (XRB) spans Anhui and Zhejiang provinces, originating from Xiuning County of Huangshan City in Anhui Province (Fig. 1). XRB's primary watercourse flows through Jiekou Town to the form the largest inbound river to Zhejiang Province and the primary feed to Qiandao Lake in Chun'an County. In addition to providing drinking water for many communities of Zhejiang, Qiandao Lake is a strategic water reserve for the broader Yangtze River Delta. Since the beginning of the 21st century, degradation of inflowing water from the upper reaches of XRB in Anhui has produced increasingly serious levels of eutrophication of Qiandao Lake. Upstream pollution control thus holds profound practical significance for ensuring the quality of this critical, downstream reserve.

For the downstream government, high levels of economic and social development produce relatively higher requirements for water quality that those experienced in the relatively undeveloped municipalities upstream. Downstream communities have very low incentives to sacrifice environmental quality for additional economic benefit. Thus, to achieve a better environment for both living and industrial production, upstream and downstream management agencies have sought an effective EC mechanism to incentivize reductions of upstream pollution.

Since 2012, under the guidance of the Ministry of Finance and the Ministry of Environmental Protection of China, Anhui and Zhejiang provinces have carried out two rounds of pilot projects for an EC regime linking the upper and lower reaches of XRB. Each pilot has lasted three years and linked Huangshan City in the upper reaches with Jixi County in Xuancheng City and Chun'an County in Hangzhou City below.

For the sake of SDG analysis, the key stakeholders are simplified as the upper XRB compensated region i and the lower XRB compensating region j. Based on in-person investigations and data collection from the relevant government departments and enterprises, we establish respective pollution thresholds as $\overline{q}_i = 30$ and $\overline{q}_j = 40$, and the utility coefficients of environmental quality of two regional government meet the condition $\Delta_i = 2$, $\Delta_j = 6$. Further, according to the value range of local emission immobile portion and environmental loss factors by Yeung and Petrosyan (2008) and Huang et al. (2016), the immobile portion of local emission as established as $\varpi_i = 0.4$ and $\varpi_j = 0.8$, and the environmental loss factor lower for the upstream region, i.e., $\delta_i = 1$ and $\delta_j = 5$. Referring to Heutel (2012), we set environmental self-degradation parameters of pollutant emissions at $\eta = 0.25$. Cost coefficients of mitigation effort are $\kappa_i = 0.7$ and $\kappa_j = 0.4$, while a constellation of additional parameters is held at: $a = 3$, $\forall = 1$, $\alpha = 0.7$, $\beta = 0.4$, $\rho = 0.5$, $\lambda = 0.4$, $\tau_0 = 60$, $\varphi_i = 0.5$, $\varphi_j = 0.6$, $\sigma = 0.07$.

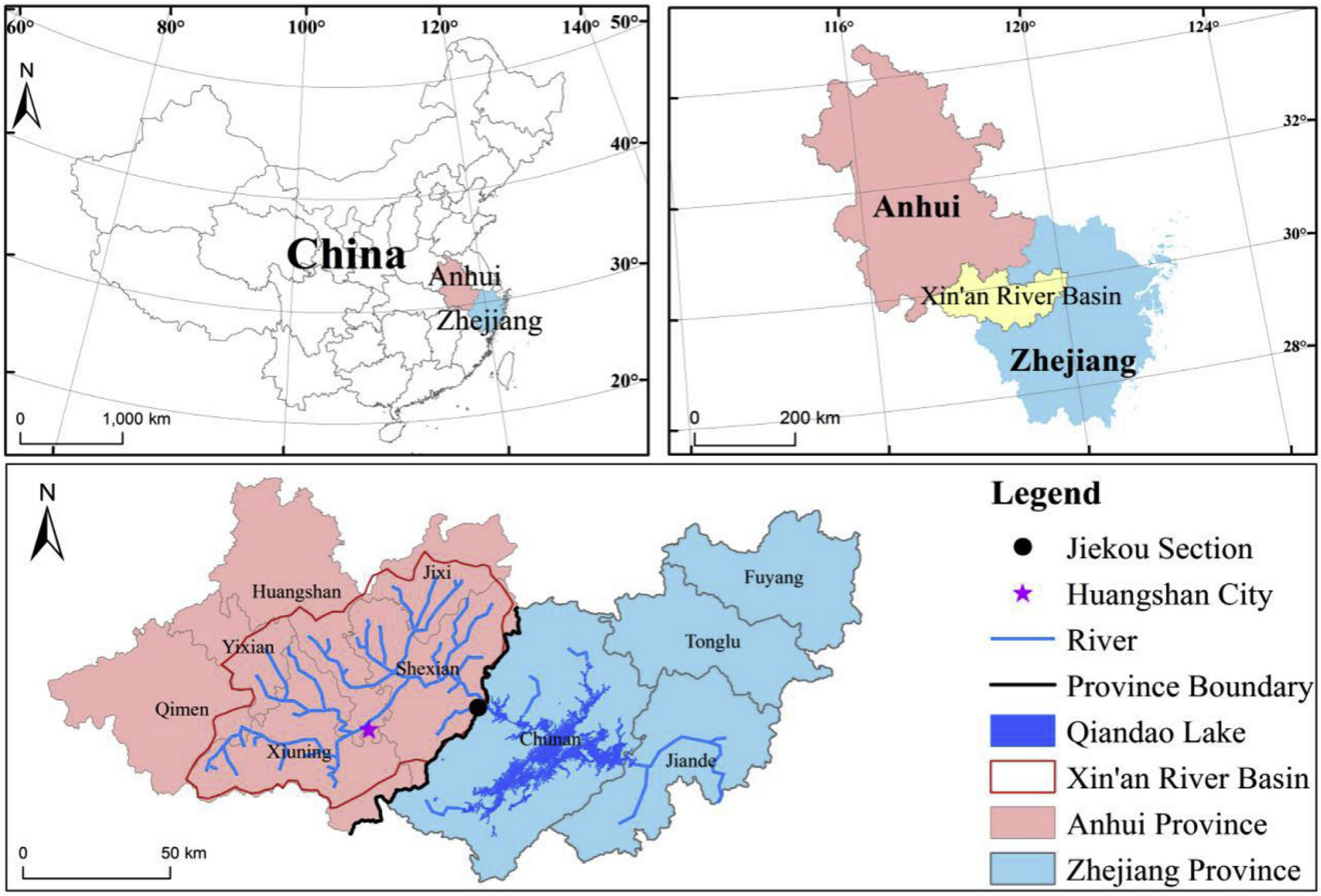


Fig. 1. The location of Xin'an River Basin.

4.2. Analysis of stochastic pollutant stock

Using the simulation method of Prasad and Sethi (2004), for Eqs. (9 and 15), the numerical approximation follows:

$$\tau(t+\Theta)=\tau(t)+(\Omega-(\lambda+\eta)\tau(t))\Theta+\sigma\sqrt{\tau(t)}\sqrt{\Theta}\zeta(t) \tag{22}$$

$$\tau(t+\Theta)=\tau(t)+(\Upsilon-(\lambda+\eta)\tau(t))\Theta+\sigma\sqrt{\tau(t)}\sqrt{\Theta}\zeta(t) \tag{23}$$

where the random variable $\zeta(t)\sim N(0,1)$, i.e., $\zeta(t)$ is the independent and identically distributed (i.i.d.) standard normal random variable with a time step of $\Theta=0.001$.

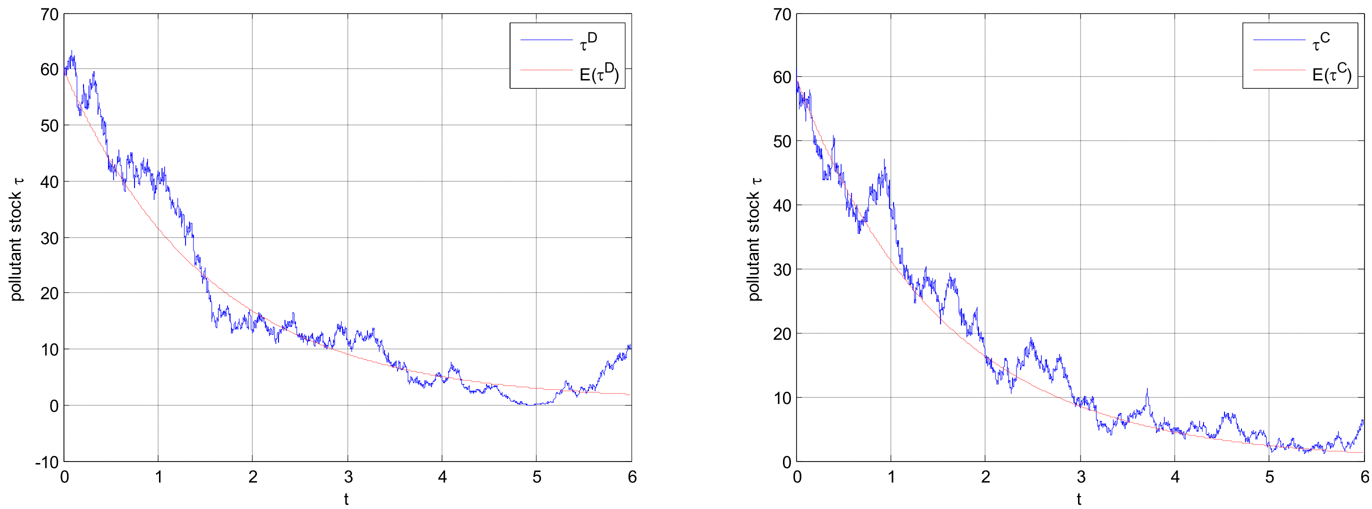
As shown graphically in Fig. 2 above, the state variable (i.e., pollutant stock) fluctuates continuously due to the Brownian motion, making it difficult for either the upstream or downstream player to obtain an exact value of watershed pollutant stocks at any given time. Note the pollutant stock of cooperative game is not necessarily lower than that of Stackelberg non-cooperative game at any time.

Hereafter, following procedures established in Zwillinger (1998), a confidence interval can be adopted to describe the variation range of the pollutant stock. At a 95% confidence level, the confidence interval of pollutant stocks should be $(E[\tau(t)]-1.96\sqrt{S[\tau(t)]},E[\tau(t)]+1.96\sqrt{S[\tau(t)]})$. Setting a confidence interval improves the predictive power of diagnostic tools for watershed management departments who deal with fluctuation in effluent matter. Whereas minor fluctuations in observed ranges range require a rough check, major overshoots of the interval indicate a need for detailed review.

Further, the simulation holds quite true to observed, real-world conditions recalled in Fig. 2. Indeed, Chinese policymakers regularly fail to accurately predict actual, observed values of pollutant stocks in the XRB basin's primary river. The use of an interval improves predictive power, in line with Huang et al. (2016), who argue that optimal management strategies should take emission levels as a control to be optimized in tandem with dynamic states of the wider ecosystem. We concur that analysts too often attend solely to the difficulty of coordinating optimal emission levels and neglect to account for random disturbances in the environment driving outcomes. Accounting for the absorption rate of the natural environment, as well as the reciprocal transmission rates of pollutants across borders, are critical to setting an optimal compensation rate.

4.3. Sensitivity analysis of optimal equilibrium feedback

Due to the multi-directional and interactive relationship between the upper XRB and lower XRB, relevant shifts in keys parameter each differ across various game contracts. To better observe changes of relevant parameters, let the expected value of pollutant stock is set to 10 in advance. Next, by fixing other parameters unchanged and changing any of the parameter according to the pattern of −50%, −25%, +25%, +50%, a sensitivity analysis of optimal equilibrium strategies for key parameters appears in Table 2. Within the value range of the given model parameters employed in the analysis of the XRB regime, the EC coefficient



(a) Stackelberg non-cooperative game (b) Cooperative game

Fig. 2. Expectations and variance in actual pollution stocks between two game contracts.

Table 2
Sensitivity analysis.

Variables	q_i^{D*}/q_i^{C*}		q_j^{D*}/q_j^{C*}		u_i^{D*}/u_i^{C*}		u_j^{D*}/u_j^{C*}		ε^*	$\Pi_i^{D*}/\Pi_j^{D*}/\Pi_i^{C*}/\Pi_j^{C*}$			
Benchmark	0.76/0.13		0.14/0.12		0.34/0.37		0.54/0.65		0.82	30.15/85.54/30.67/85.92			
δ_i=(0.50 → 1.50)	−	−	×	−	+	+	×	+	−	−	+	−	+
δ_j=(2.50 → 7.50)	×	−	−	−	+	+	+	+	+	−	−	−	−
ϖ_i=(0.20 → 0.60)	−	−	×	×	×	×	×	×	×	−	+	−	+
ϖ_j=(0.40 → 1.20)	×	×	−	−	×	×	×	×	×	+	−	+	−
Δ_i=(1.00 → 3.00)	−	−	×	−	×	×	×	×	×	+	+	+	+
Δ_j=(3.00 → 9.00)	×	−	−	−	×	×	×	×	×	+	+	+	+
β=(0.05 → 0.15)	×	×	×	×	+	+	+	+	×	−	−	−	−
κ_i=(0.35 → 1.05)	×	×	×	×	−	×	−	×	×	+	+	+	+
κ_j=(0.20 → 0.60)	×	×	×	×	×	−	×	−	×	×	+	+	+

$\varepsilon^* = 0.82$ may be attained to satisfy the condition of $2\delta_j > \delta_i$. The impacts exerted by different parameters are specific presented below.

4.3.1. Changes in δ_h

Sensitivity analysis indicates the slope of optimal emission quantities at the same point (i.e., at any δ_i from 0.5 to 1.5 and δ_j from 2.5 to 7.5) presents a declining trend in both game situations. Meanwhile, respective mitigation efforts reveal an opposing trend. Perhaps unsurprisingly, increasing environmental losses enable both upstream and downstream players to reduce emissions within their jurisdictions through greater levels of proactive mitigation. Yet increasing δ_h also means higher levels of environmental damage at lower optimal emission quantities. In particular, environmental losses in the economically underdeveloped region, i.e, the upper reach of XRB, remain relatively high, indicating upstream regions are apt to persist in following incumbent patterns of laissez-faire control.

4.3.2. Changes in Δ_h

Under an increasing utility coefficient of environmental quality (Δ_h), optimal welfare of both upper and lower XRB improves under both games. However, it is noteworthy that within the Stackelberg non-cooperative situation, the trend of optimal emission quantities in two regions runs counter to welfare. Taking the upper XRB as an example, as a rational "economic participant", Anhui possesses local incentive to cut emissions, resulting in a slight drop to 0.58. Yet such environmental preference-induced changes have no direct impact on emission capacities downstream. Under cooperation, instantaneous emissions on both sides are more apt to decrease in sequence, with net increases to both upstream and downstream welfare.

4.3.3. Changes in ϖ_h

Variation in the immobile portion of local emissions (ϖ_h) does not directly impact regional investment behavior or the EC coefficient in either two game situations, yet it leads to important changes to welfare outcomes. For example, when upstream immobility of emissions increases from 0.2 to 0.6, - at any level downstream ϖ_i - upstream welfare decreases accordingly, i.e., Π_i^{D*} falls from 30.15 to 29.79. The downstream player is less affected by the transboundary spillover $1-\varpi_i$, and enjoys a slight increase in local welfare Π_j^{D*} from 85.54 to 86.87. The same dynamic holds in vice versa. Under cooperation, it is precisely this impacts of the transboundary spillover $1-\varpi_h$ that motivate both the upper and lower reaches to improve governance and encourage local industry to reduce emissions.

4.3.4. Changes in β and κ_h

Under the Stackelberg non-cooperative game, a tripling in the value of β (from 0.05 to 0.15) raises upstream mitigation effort u_i^{D*} from 0.34 to 0.51. Conversely, a tripling of κ_i decreases upstream mitigation effort from 0.34 to 0.23. The lower the input cost of mitigation, the more effective the effort as government more vigorously implements reform. Sensitivity analysis further highlights that cooperation encourages an upstream, compensated region to invest more in mitigation in ways that increase total welfare. This finding is consistent with Assumption 2, which describes the rising resource expenses of capital, attention, time, etc. a government incurs to achieve an additional unit of clean up. Because both parties remain bound by the compensation agreement, an economically developed region has incentives to support both sustained eco-environmental protection and industrial investment in less-industrialized, upstream neighbors. Incentives to this end should serve to promote more balanced development.

5. Results and discussion

In this section, we discuss the equilibrium strategies given by the investigated game models. On the whole, optimal equilibrium feedback analysis and simulation analysis of the two games provide useful reference values for dealing with transboundary pollution and drive a series of important discussion points.

Proposition 5. *Optimal emission quantities of two regions meet the conditions $q_i^{C*} < q_i^{D*}$ and $q_j^{C*} < q_j^{D*}$.*

This implies that when a pollution spillover occurs, and the two sides influence each other's incentives to invest in mitigation, the optimal total emission quantity of the Stackelberg non-cooperative situation exceeds that of the cooperative alternative. In other words, the presence of a pollution spillover and its influence on the behavioral incentives of the counterparty leads to significantly lower optimal emission quantities under cooperation. Decision-making differences are further affected by multiple coefficients (e.g. the utility coefficient of environmental quality, immobile portion of local emissions, and environmental damage losses), which are in turn shaped by the strength of the spillover (Proof, see appendix E).

More precisely, when neither side selects a strategy of joint governance, regional governments make decisions seeking the maximization of their discrete welfare, i.e., considering solely the impact of emissions on economic output and the local environment, while ignoring the negative utilities generated by the external spillover. To the contrary, in the case of cooperative game, both sides emphasize the impact of pollution in surrounding regions on their internal utilities, cognizant that the transmission of pollutants across regional lines erodes overall regional welfare. This result differs from a key conclusion of Yeung (2007), who focuses on the global impact of pollution but largely neglects the role of localized environmental dynamics in shaping outcomes. To wit, reductions in winter haze and water pollution in China have often hinged more on favorable diffusion conditions than the imposition of stricter controls (Hao and Liu, 2016; Liu et al., 2016a, b).

Proposition 6. *In the cooperative game, the optimal mitigation efforts in both regions is higher than that in the Stackelberg non-cooperative case, i.e., $u_i^{C*} > u_i^{D*}$ and $u_j^{C*} > u_j^{D*}$.*

Here, the difference between the two sides of the game are not only affected by the cost coefficient of mitigation effort and marginal impact coefficient of mitigation effort, but also by the environmental damage losses of the compensated region. Given $\varepsilon^* = (2\delta_j - \delta_i)/(2\delta_j + \delta_i) > 0$, $2\delta_j > \delta_i$ holds (Proof, see appendix E).

It is worth noting this condition possesses practical significance. The environmental damage losses of compensating region should be more than twice that of compensated region, i.e.,$2\delta_j > \delta_i$. If such condition is not met, the cooperation process will not attain Pareto improvement, depriving the compensator of motivation to ameliorate total environmental losses caused by the trans-regional spillover. This is particularly relevant for watershed eco-compensation in China, where an upper-level entity, often a higher-level government, has a clear role to bring upstream (water-supplying) and downstream (water-receiving) jurisdictions together within an efficient mechanism. In practice, with the increase in mitigation effort, the compensating region will inevitably suffer from economic losses if long-term investments are unrewarded. Therefore, to ensure the sustainable development of the entire regional environment, it is necessary for the compensator to encourage the compensating region to encourage greater governance through a compensation payment.

As one of the largest polluters in the world, with more than 40,000 local jurisdictions in varying stages of development, China has seen growing collaboration among local governments amid rising pressure from the Central Government for environmental protection (Yang, 2017). While the state has expended much effort using environmental finance tools like EC to drive abatement, decision-makers have invested only modest consideration establishing best-practices for tool selection and calibration. Better policymaking should possess sound scientific basis in attending to implementation risks, particularly where policy failure may prove particularly costly to both large state-led projects and local communities. In China's growth-driven economy, where environmental protection resources are in short supply, calibrating effective and adaptive compensation standards should prove critical to securing positive and sustained policy outcomes.

Proposition 7. *The expected value and variance of the pollutant stock and its corresponding stable value in the cooperative game are smaller than those in the Stackelberg non-cooperative situation, i.e.,:*

$$\begin{cases} E\left[\tau^C(t)\right] < E\left[\tau^D(t)\right], & \lim\limits_{t\to\infty} E\left[\tau^C(t)\right] < \lim\limits_{t\to\infty} E\left[\tau^D(t)\right] \\ S\left[\tau^C(t)\right] < S\left[\tau^D(t)\right], & \lim\limits_{t\to\infty} S\left[\tau^C(t)\right] < \lim\limits_{t\to\infty} S\left[\tau^D(t)\right] \end{cases}$$

This indicates that in the non-cooperative game, the stock of pollutant is higher than that in the cooperative game. Further, variance in the pollutant stock is significantly affected by various random interference factors, as both sides must synchronously bear greater risks to achieve higher economic and environmental utilities (Proof, see appendix E).

By modelling uncertainty in nature's capability to replenish the environment, we thus show that the ultimate pollutant stock in the non-cooperative game is probabilistically higher than under cooperation. Thus, although the effectiveness of environmental pollution control still depends to a key extent on the subjective initiative of administrators, commitments remain contingent on multiple uncertainties including the degree of pollution transfer, terrain conditions and hydrological characteristics. This finding supports expanded investment in monitoring and evaluation of natural phenomena and processes, and in skills for deciphering and disseminating such relevant data to public managers and their private sector counterparts.

6. Conclusions

6.1. Conclusions and implications

This study develops a SDG model to examine the transboundary pollution issue between one compensating region and one ecological compensated region - in the absence of administrative subordination - via an EC mechanism. This robust, game theoretical investigation of compensation processes contends with negative spillovers of a regional accumulative pollutant and incorporates stochastic processes to account for uncertain dynamics in pollution stocks and environmental factors that could otherwise confound an efficient remedial contract. The SDG model allows equilibrium strategies of two neighboring governments to evolve and it informs a more optimal control theory based on a general welfare allocation mechanism. By determining and contrasting outcomes across both cooperative and non-cooperative game scenarios, with validation via a numerical simulation based on the XRB case, this study returns several insights:First, equilibrium results show that as a long-term mechanism, an EC contract is most effective when marginal losses due to additional environmental damage in the compensating region are more than twice those of the compensated region. This finding should aid help decision makers identify relative ecological and economic thresholds for deploying EC contracts across China.

Second, the analysis shows the ultimate pollutant stock in the non-cooperative game is probabilistically higher than that in a cooperative game, underscoring the importance of pro-cooperation incentives to avoid conditions in which both sides are forced to bear risks synchronously in mutually-undermining pursuit of higher utilities. Where non-cooperation prevails, policymakers will likely experience significantly greater difficulty assessing the actual pollutant levels, such that establishing reliable confidence intervals for likely effluent values may provide a useful diagnostic tool for environmental managers.

Third, for an EC contract to remain credible, a degree of time consistency is required, i.e. compliance must remain economically optimal for all parties to the agreement at all times. An inter-temporal decomposition of the welfare allocation mechanism allows us to obtain time consistent outcomes and inform a grounded simulation of outcomes. Based on this advance, our investigation demonstrates that cross-regional cooperation constantly outperforms Stackelberg non-cooperation, irrespective of variations in economic interests and cost-sensitivities to emission abatement.

6.2. Limitations

This study retains several limitations that call for further research. Practically, transboundary spillovers are multi-faceted, such that efficient control may be best accomplished through the participation and interaction of multiple entities. In China, these may include regional governments, major enterprises, and the public at large. Inclusion of these additional players introduces engaging problems for robust game analyses and future studies. Research extensions should also consider incorporating the EC compensation coefficient into prevailing models and adapting this rate to the regulation of supply and demand on the basis of fulfilling a robust welfare allocation mechanism.

Acknowledgment

This study was supported by the National Natural Science Foundation of China (Nos. 71573283, 71701217); the National Social Science Fund of China (No. 19BJY028); the Tier 1 Academic Research Fund of Singapore's Ministry of Education (No. #16-C207-SMU-023); the Startup Foundation for Introducing Talent of NUIST (No. 2019r045). The authors are grateful to the editor and the anonymous referees for their constructive comments and profound insights to improve this study.

Appendix A. Proof of Proposition 1

With the inverse induction method, the optimization problem of region i is attained as $\Pi_i^{D*}(q_i, u_i) = e^{-\rho t} V_i^D(\tau)$ after time t. Abiding by the optimal control theory, $V_i^D(\tau)$ for any $\tau \geq 0$ satisfies the Hamilton-Jacobi-Bellman (HJB) equation below:

$$\rho V_i^D(\tau) = \max_{E_i, u_i} \left\{ \begin{array}{l} \ln(aq_i) + \Delta_i\left[\overline{q}_i - \varpi_i q_i - (1-\varpi_j)q_j\right] - \kappa_i(1-\varepsilon)u_i^2 - \delta_i \tau \\ + V_i^{D'}(\tau)\left[\alpha\left(q_i + q_j\right) - \beta(u_i + u_j) - (\lambda + \eta)\tau\right] + \dfrac{\sigma^2(\tau)}{2} V_i^{D''}(\tau) \end{array} \right\} \tag{24}$$

where $V_i^{D'}(\tau) = \frac{dV_i^D(\tau)}{d\tau}$ and $V_i^{D''}(\tau) = \frac{d^2V_i^D(\tau)}{d\tau^2}$.

Differentiate the above HJB equation with respect to q_i and u_i, the optimal feedback strategies are then acquired as:

$$q_i^* = \frac{1}{\Delta_i \varpi_i - \alpha V_i^{D'}}; \qquad u_i^* = -\frac{\beta V_i^{D'}}{2\kappa_i(1-\varepsilon)} \tag{25}$$

Likewise, the optimization problem of region j is $\Pi_j^{D*}(q_j, u_j, \varepsilon) = e^{-\rho t} V_j^D(\tau)$ after time t. The HJB equation associated with such optimal control problem is expressed as:

$$\rho V_j^D(\tau) = \max_{q_j, u_j, \varepsilon} \left\{ \begin{array}{l} \ln\left(\forall aq_j\right) + \Delta_j\left[\overline{q}_j - \varpi_j q_j - (1-\varpi_i)q_i\right] - \kappa_j u_j^2 - \varepsilon\kappa_i u_i^2 - \delta_j \tau \\ + V_j^{D'}\left[\alpha\left(q_i + q_j\right) - \beta(u_i + u_j) - (\lambda + \eta)\tau\right] + \dfrac{\sigma^2(\tau)}{2} V_j^{D''}(\tau) \end{array} \right\} \tag{26}$$

One can usually characterize the optimal feedback strategies of q_j, u_j and ε from the first-order condition:

$$q_j^* = \frac{1}{\Delta_j \varpi_j - \alpha V_j^{D'}}; \;\; u_j^* = -\frac{\beta V_j^{D'}}{2\kappa_j}; \;\; \varepsilon^* = \frac{2V_j^{D'} - V_i^{D'}}{2V_j^{D'} + V_i^{D'}}; \tag{27}$$

which when substituting back into the above two HJB equations yields:

$$\begin{cases} \rho V_i^D(\tau) = -\left[\delta_i + (\lambda+\eta)V_i^{D'}(\tau)\right]\tau + \frac{\sigma^2(\tau)}{2}V_i^{D''}(\tau) + \ln\left[\frac{a}{\Delta_i\varpi_i - \alpha V_i^{D'}(\tau)}\right] + \Delta_i\bar{q}_i - 1 \\ \qquad -\frac{\Delta_i(1-\varpi_j) - \alpha V_i^{D'}(\tau)}{\Delta_j\varpi_j - \alpha V_j^{D'}(\tau)} + \frac{\beta^2 V_i^{D'}\left[2V_j^{D'}(\tau) + V_i^{D'}(\tau)\right]}{8\kappa_i} + \frac{\beta^2 V_i^{D'}(\tau)V_j^{D'}(\tau)}{2\kappa_j} \\ \rho V_j^D(\tau) = -\left[\delta_j + \left(\lambda+\eta\right)V_j^{D'}\left(\tau\right)\right]\tau + \frac{\sigma^2\left(\tau\right)}{2}V_j^{D''}\left(\tau\right) + \ln\left[\frac{\forall a}{\Delta_j\varpi_j - \alpha V_j^{D'}\left(\tau\right)}\right] + \Delta_j\bar{q}_j - 1 \\ \qquad -\frac{\Delta_j(1-\varpi_i) - \alpha V_j^{D'}(\tau)}{\Delta_i\varpi_i - \alpha V_i^{D'}(\tau)} + \frac{\beta^2\left[2V_j^{D'}(\tau) + V_i^{D'}(\tau)\right]^2}{16\kappa_i} + \frac{\beta^2 V_j^{D'2}(\tau)}{4\kappa_j} \end{cases} \tag{28}$$

Following the structure of Eq. (28), the linear analytical formula of $V_i^D(\tau)$ and $V_j^D(\tau)$ for τ is assumed as:

$$\begin{cases} V_i^D(\tau) = m_1\tau + m_2 \\ V_j^D\left(\tau\right) = n_1\tau + n_2 \end{cases} \tag{29}$$

where m_1, m_2, n_1 and n_2 are constant.

Substituting the above general function forms, $V_i^{D'}(\tau) = m_1$ and $V_j^{D'}(\tau) = n_1$ into Eq. (28) yields:

$$\begin{cases} \rho(m_1\tau + m_2) = -[\delta_i + (\lambda+\eta)m_1]\tau + \ln\left[\frac{a}{\Delta_i\varpi_i - \alpha m_1}\right] + \Delta_i\bar{q}_i - 1 - \frac{\Delta_i(1-\varpi_j) - \alpha m_1}{\Delta_j\varpi_j - \alpha n_1} \\ \qquad + \frac{\beta^2 m_1[m_1 + 2n_1]}{8\kappa_i} + \frac{\beta^2 m_1 n_1}{2\kappa_j} \\ \rho(n_1\tau + n_2) = -[\delta_j + (\lambda+\eta)n_1]\tau + \ln\left[\frac{\forall a}{\Delta_j\varpi_j - \alpha n_1}\right] + \Delta_j\bar{q}_j - 1 - \frac{\Delta_j(1-\varpi_i) - \alpha n_1}{\Delta_i\varpi_i - \alpha m_1} \\ \qquad + \frac{\beta^2[m_1 + 2n_1]^2}{16\kappa_i} + \frac{\beta^2 n_1^2}{4\kappa_j} \end{cases} \tag{30}$$

Let $\tau \geq 0$, the parameter values of Eq. (29) can be calculated:

$$\begin{cases} m_1 = -\frac{\delta_i}{\rho+\lambda+\eta} \\ m_2 = \ln\left(\frac{a}{\Delta_i\varpi_i - \alpha m_1}\right)\Big/\rho + (\Delta_i\bar{q}_i - 1)\Big/\rho - [\Delta_i(1-\varpi_j) - \alpha m_1] \\ \qquad \Big/[(\Delta_j\varpi_j - \alpha n_1)\rho] + \left[\beta^2 m_1(m_1 + 2n_1)\right]\Big/8\rho\kappa_i + \beta^2 m_1 n_1\Big/2\rho\kappa_j \end{cases} \tag{31}$$

$$\begin{cases} n_1 = -\dfrac{\delta_j}{\rho+\lambda+\eta} \\ n_2 = \ln\left(\dfrac{\forall a}{\varDelta_j\varpi_j+\alpha n_1}\right)\Big/\rho + \left(\varDelta_j\bar{q}_j - 1\right)\Big/\rho - \left[\varDelta_j\left(1-\varpi_i\right) - \alpha n_1\right] \\ \quad \Big/\left[\left(\varDelta_i\varpi_i - \alpha m_1\right)\rho\right] + \left[\beta^2\left(m_1+2n_1\right)^2\right]\Big/16\rho\kappa_i + \beta^2 n_1^2\Big/4\rho\kappa_j \end{cases} \tag{32}$$

By substituting the results of Eqs. (31 and 32) into Eqs. (25 and 27), respectively, the conclusion of Proposition 1 can be attained, thus completing the Proof.

Appendix B. Proof of Proposition 2

Itô's lemma is an identity adopted in Itô calculus to search the change of function over time in a stochastic process. If $F(t)$ is a quadratic continuous differential function, then t satisfies the following Itô stochastic integral equation:

$$\begin{cases} \tau^2(t) = \tau_0^2 + \int\limits_0^t \left[\left(2\Omega+\sigma^2\right)\tau(s) - 2(\lambda+\eta)\tau^2(s)\right]ds + \int\limits_0^t 2\tau(s)\sigma\sqrt{\tau(s)}dP(s) \\ \tau^2(0) = (\tau_0)^2 \geq 0 \end{cases} \tag{38}$$

$$F(B(t)) = F(0) + \int\limits_0^t F'(s, B(s))dB(s) + \frac{1}{2}\int\limits_0^t F'(s, B(s))ds \tag{33}$$

where $\{B(t) : t \in [0, \infty)\}$ is the Brownian motion.

Accordingly, Eq. (9) can be rewritten as a random integral:

$$\begin{cases} \tau(t) = \tau_0 + \int\limits_0^t [\Omega - (\lambda+\eta)\tau(s)]ds + \int\limits_0^t \sigma\sqrt{\tau(s)}dP(s) \\ \tau(0) = \tau_0 \geq 0 \end{cases} \tag{34}$$

$$\begin{cases} dE\left[\tau^2(t)\right]\Big/dt = \left(2\Omega+\sigma^2\right)\left[\dfrac{\Omega}{\lambda+\eta} + \left(\tau_0 - \dfrac{\Omega}{\lambda+\eta}\right)e^{-(\lambda+\eta)t}\right] - 2(\lambda+\eta)E\left[\tau^2(t)\right] \\ E\left[\tau^2(t)\right] = (\tau_0)^2 \end{cases} \tag{40}$$

where $\Omega = \frac{\alpha(\rho+\lambda+\eta)}{\varDelta_i\varpi_i(\rho+\lambda+\eta)+\alpha\delta_i} + \frac{\alpha(\rho+\lambda+\eta)}{\varDelta_j\varpi_j(\rho+\lambda+\eta)+\alpha\delta_j} - \frac{\beta^2(\delta_i+2\delta_j)}{4\kappa_i(\rho+\lambda+\eta)} - \frac{\beta^2\delta_j}{2\kappa_j(\rho+\lambda+\eta)}$.

It is easy to judge that the expectation of pollutant stock is independent of random disturbance factor, then:

$$\begin{cases} E[\tau(t)] = \tau_0 + \int\limits_0^t \{\Omega - (\lambda+\eta)E[\tau(s)]\}ds \\ E[\tau(0)] = \tau_0 \end{cases} \tag{35}$$

The above equation can be seen as the ordinary differential equation with respect to $E[\tau(t)]$, and the solution is:

$$\begin{cases} E\left[\tau^D(t)\right] = \dfrac{\Omega}{\lambda+\eta} + \left(\tau_0 - \dfrac{\Omega}{\lambda+\eta}\right)e^{-(\lambda+\eta)t} \\ E\left[\tau^D(0)\right] = \tau_0 \end{cases} \tag{36}$$

Subsequently, to obtain the variance of pollutant stock, Itô lemma is supposed to be applied to Eq. (9):

$$\begin{cases} d\tau^2(t) = \left[\left(2\Omega+\sigma^2\right)\tau - 2(\lambda+\eta)\tau^2\right]dt + 2\tau\sigma\sqrt{\tau}dP(t) \\ \tau^2(0) = (\tau_0)^2 \geq 0 \end{cases} \tag{37}$$

Accordingly, the above equation can be rewritten as the following random integral form:

and its expectation is,

$$\begin{cases} E\left[\tau^2\left(t\right)\right] = \tau_0^2 + \int\limits_0^t \left\{\left(2\Omega+\sigma^2\right)E\left[\tau\left(s\right)\right] - 2\left(\lambda+\eta\right)E\left[\tau^2\left(s\right)\right]\right\}ds \\ E\left[\tau^2\left(0\right)\right] = \left(\tau_0\right)^2 \end{cases} \tag{39}$$

Substituting the result of Eq. (36) into Eq. (39) and rewritten it as the differential form:

Solving the above linear non-homogeneous differential equation yields:

$$\begin{cases} E\left[\tau^2(t)\right] = \tau_0^2 e^{-2(\lambda+\eta)t} + \dfrac{\Omega\left(2\Omega+\sigma^2\right)}{2(\lambda+\eta)^2}\left(1 - e^{-2(\lambda+\eta)t}\right) + \dfrac{\left(2\Omega+\sigma^2\right)[\tau_0(\lambda+\eta) - \Omega]}{(\lambda+\eta)^2}\left(e^{-(\lambda+\eta)t} - e^{-2(\lambda+\eta)t}\right) \\ E\left[\tau^2(0)\right] = (\tau_0)^2 \end{cases} \tag{41}$$

Hence, the variance of the pollutant stock can be computed as:

$$S\left[\tau^D(t)\right]=E\left[\tau^2(t)\right]-\left[E\left[\tau^D(t)\right]\right]^2 =\frac{\sigma^2\{\Omega-2[\Omega-(\lambda+\eta)\tau_0]e^{-(\lambda+\eta)t}+[\Omega-2(\lambda+\eta)\tau_0]e^{-2(\lambda+\eta)t}\}}{2(\lambda+\eta)^2} \tag{42}$$

$$\rho(s_1\tau+s_2)=-\left[\delta_i+\delta_j+(\lambda+\eta)s_1\right]\tau+\ln\left[\frac{a}{\Delta_i\varpi_i+\Delta_j(1-\varpi_i)-\alpha s_1}\right] +\ln\left[\frac{\forall a}{\Delta_i(1-\varpi_j)+\Delta_j\varpi_j-\alpha s_1}\right]+\Delta_i\bar{q}_i+\Delta_j\bar{q}_j-2 +\frac{\beta^2 s_1^2}{4\kappa_i}+\frac{\beta^2 s_1^2}{4\kappa_j} \tag{47}$$

Hereafter, the parameter values satisfy:

$$\begin{cases} s_1=-\dfrac{\delta_i+\delta_j}{\rho+\lambda+\eta} \\ s_2=\ln\left[\dfrac{a}{\Delta_i\varpi_i+\Delta_j(1-\varpi_i)-\alpha s_1}\right]\Big/\rho+\ln\left[\dfrac{\forall a}{\Delta_i(1-\varpi_j)+\Delta_j\varpi_j-\alpha s_1}\right]\Big/\rho \\ \quad+\left(\Delta_i\bar{q}_i+\Delta_j\bar{q}_j-2\right)\Big/\rho+\beta^2 s_1^2\Big/4\rho\kappa_i+\beta^2 s_1^2\Big/4\rho\kappa_j \end{cases} \tag{48}$$

Appendix C. Proof of Proposition 3

Under cooperative arrangement, the optimization problem of two regions after time t is required as $\Pi^C=e^{-\rho t}V^C(\tau)$. For any $\tau\geq 0$, $V^C(\tau)$ satisfies the HJB equation below:

$$\rho V^C(\tau)=\max_{E_h,u_h}\left\{\begin{array}{l}\ln(aq_i)+\ln\left(\forall aq_j\right)+\Delta_i\left[\bar{q}_i-\varpi_i q_i-(1-\varpi_j)q_j\right]+\Delta_j\left[\bar{q}_j-\varpi_j q_j-(1-\varpi_i)q_i\right]-\kappa_i u_i^2 \\ -\kappa_j u_j^2-(\delta_i+\delta_j)\tau+V^{C'}(\tau)\left[\alpha\left(q_i+q_j\right)-\beta(u_i+u_j)-(\lambda+\eta)\tau\right]+\dfrac{\sigma^2(\tau)}{2}V^{C''}(\tau)\end{array}\right\} \tag{43}$$

where $V^{C'}(\tau)=\frac{dV^C(\tau)}{d\tau}$ and $V^{C''}(\tau)=\frac{d^2V^C(\tau)}{d\tau^2}$.

Similarly, one can generate the optimal feedback strategies of q_i, q_j, u_i and u_j from the first-order condition:

$$q_i^*=\frac{1}{\Delta_i\varpi_i+\Delta_j(1-\varpi_i)-\alpha V^{C'}(\tau)};\quad q_j^*=\frac{1}{\Delta_i(1-\varpi_j)+\Delta_j\varpi_j-\alpha V^{C'}(\tau)};$$

$$u_i^*=-\frac{\beta V^{C'}}{2\kappa_i};\quad u_j^*=-\frac{\beta V^{C'}}{2\kappa_j} \tag{44}$$

and form:

$$\rho V^C(\tau)=-\left[\delta_i+\delta_j+(\lambda+\eta)V^{C'}(\tau)\right]\tau+\frac{\sigma^2(\tau)}{2}V^{C''}(\tau) +\ln\left[\frac{a}{\Delta_i\varpi_i+\Delta_j(1-\varpi_i)-\alpha V^{C'}(\tau)}\right]+\ln\left[\frac{\forall a}{\Delta_i(1-\varpi_j)+\Delta_j\varpi_j-\alpha V^{C'}(\tau)}\right] +\Delta_i\bar{q}_i+\Delta_j\bar{q}_j+\frac{\beta^2 V^{C'2}(\tau)}{4\kappa_i}+\frac{\beta^2 V^{C'2}(\tau)}{4\kappa_j}-2 \tag{45}$$

Similarly, the linear analytical formula of $V^C(\tau)$ for τ is:

$$V^C(\tau)=s_1\tau+s_2 \tag{46}$$

where s_1 and s_2 are constant.

Substituting the above general function forms and $V^{C'}(\tau)=s_1$ into Eq. (45) provides:

Using the similar approach, Eq. (48) is submitted into Eq. (44), and the results in Proposition 3 can be attained. This accomplishes the Proof.

Appendix D. Proof of Proposition 4

Similarly to the Proof of Proposition 2, Eq. (9) can be rewritten as a random integral:

$$\begin{cases} \tau(t)=\tau_0+\displaystyle\int_0^t[\Upsilon-(\lambda+\eta)\tau(s)]ds+\int_0^t\sigma\sqrt{\tau(s)}dP(s) \\ \tau(0)=\tau_0\geq 0 \end{cases} \tag{49}$$

where

$$\Upsilon=\frac{\alpha(\rho+\lambda+\eta)}{\Delta_i\varpi_i(\rho+\lambda+\eta)+\Delta_j(1-\varpi_i)(\rho+\lambda+\eta)+\alpha(\delta_i+\delta_j)} +\frac{\alpha(\rho+\lambda+\eta)}{\Delta_i(1-\varpi_j)(\rho+\lambda+\eta)+\Delta_j\varpi_j(\rho+\lambda+\eta)+\alpha(\delta_i+\delta_j)} +\frac{\alpha(\rho+\lambda+\eta)}{\Delta_j\varpi_j(\rho+\lambda+\eta)+\alpha\delta_j}-\frac{\beta^2(\delta_i+\delta_j)}{2\kappa_i(\rho+\lambda+\eta)}-\frac{\beta^2(\delta_i+\delta_j)}{2\kappa_j(\rho+\lambda+\eta)}.$$

Accordingly, the expectation of pollutant stock satisfies:

$$\begin{cases} E[\tau(t)]=\tau_0+\displaystyle\int_0^t\{\Upsilon-(\lambda+\eta)E[\tau(s)]\}ds \\ E[\tau(0)]=\tau_0 \end{cases} \tag{50}$$

Similarly, the solution of the above equation can be obtained as follows:

$$\begin{cases} E\left[\tau^C(t)\right] = \frac{\Upsilon}{\lambda+\eta} + \left(\tau_0 - \frac{\Upsilon}{\lambda+\eta}\right)e^{-(\lambda+\eta)t}, \lim_{t\to\infty} E\left[\tau^C(t)\right] = \frac{\Upsilon}{\lambda+\eta} \\ E\left[\tau^C(0)\right] = \tau_0 \end{cases} \tag{51}$$

Further, applying Itô lemma to Eq. (9) gives:

$$\begin{cases} d\tau^2(t) = \left[\left(2\Upsilon + \sigma^2\right)\tau - 2(\lambda+\eta)\tau^2\right]dt + 2\tau\sigma\sqrt{\tau}dP(t) \\ \tau^2(0) = (\tau_0)^2 \geq 0 \end{cases} \tag{52}$$

The above equation can be rewritten as a random integral form below:

$$\begin{cases} \tau^2(t) = \tau_0^2 + \int_0^t \left[\left(2\Upsilon + \sigma^2\right)\tau(s) - 2(\lambda+\eta)\tau^2(s)\right]ds + \int_0^t 2\tau(s)\sigma\sqrt{\tau(s)}dP(s) \\ \tau^2(0) = (\tau_0)^2 \geq 0 \end{cases} \tag{53}$$

and its expectation is,

$$\begin{cases} E\left[\tau^2\left(t\right)\right] = \tau_0^2 + \int_0^t \left\{\left(2\Upsilon+\sigma^2\right)E\left[\tau\left(s\right)\right] - 2\left(\lambda+\eta\right)E\left[\tau^2\left(s\right)\right]\right\}ds \\ E\left[\tau^2\left(0\right)\right] = \left(\tau_0\right)^2 \end{cases} \tag{54}$$

Substituting the result of Eq. (51) into Eq. (54) and rewritten it as the differential form:

$$\begin{cases} dE\left[\tau^2(t)\right]\big/dt = \left(2\Upsilon + \sigma^2\right)\left[\frac{\Upsilon}{\lambda+\eta} + \left(\tau_0 - \frac{\Upsilon}{\lambda+\eta}\right)e^{-(\lambda+\eta)t}\right] - 2(\lambda+\eta)E\left[\tau^2(t)\right] \\ E\left[\tau^2(t)\right] = (\tau_0)^2 \end{cases} \tag{55}$$

Solving the above linear non-homogeneous differential equation yields:

$$\begin{cases} E\left[\tau^2(t)\right] = \tau_0^2 e^{-2(\lambda+\eta)t} + \frac{\Upsilon\left(2\Upsilon+\sigma^2\right)}{2(\lambda+\eta)^2}\left(1 - e^{-2(\lambda+\eta)t}\right) + \frac{\left(2\Upsilon+\sigma^2\right)\left[\tau_0(\lambda+\eta) - \Upsilon\right]}{(\lambda+\eta)^2}\left(e^{-(\lambda+\eta)t} - e^{-2(\lambda+\eta)t}\right) \\ E\left[\tau^2(0)\right] = (\tau_0)^2 \end{cases} \tag{56}$$

Then, the variance of the pollutant stock can be computed as:

$$S\left[\tau^C(t)\right] = E\left[\tau^2(t)\right] - \left[E\left[\tau^C(t)\right]\right]^2 = \frac{\sigma^2\left\{\Upsilon - 2[\Upsilon - (\lambda+\eta)\tau_0]e^{-(\lambda+\eta)t} + [\Upsilon - 2(\lambda+\eta)\tau_0]e^{-2(\lambda+\eta)t}\right\}}{2(\lambda+\eta)^2} \tag{57}$$

Appendix E. Proof of Propositions 5 to 7

According to Eqs. (6 and 12), there exists:

$$\begin{cases} q_i^{C*} - q_i^{D*} = -\frac{(\rho+\lambda+\eta)\left[\Delta_j(1-\varpi_i)(\rho+\lambda+\eta) + \alpha\delta_j\right]}{\left[\Delta_i\varpi_i(\rho+\lambda+\eta) + \Delta_j(1-\varpi_i)(\rho+\lambda+\eta) + \alpha(\delta_i+\delta_j)\right]\left[\Delta_i\varpi_i(\rho+\lambda+\eta) + \alpha\delta_i\right]} \\ q_j^{C*} - q_j^{D*} = -\frac{(\rho+\lambda+\eta)\left[\Delta_i(1-\varpi_j)(\rho+\lambda+\eta) + \alpha\delta_i\right]}{\left[\Delta_i(1-\varpi_j)(\rho+\lambda+\eta) + \Delta_j\varpi_j(\rho+\lambda+\eta) + \alpha(\delta_i+\delta_j)\right]\left[\Delta_j\varpi_j(\rho+\lambda+\eta) + \alpha\delta_j\right]} \end{cases} \tag{58}$$

Clearly, $q_i^{C*} < q_i^{D*}$ and $q_j^{C*} < q_j^{D*}$ hold. Proposition 5 is proved.

Moreover, let $\varepsilon > 0$, according to Eqs. (10 and 16), there exists:

$$\begin{cases} u_i^{C*} - u_i^{D*} = \dfrac{\beta \delta_i}{4\kappa_i(\rho + \lambda + \eta)} \\ u_j^{C*} - u_j^{D*} = \dfrac{\beta \delta_i}{2\kappa_j(\rho + \lambda + \eta)} \end{cases} \tag{59}$$

As evidenced, $u_i^{C*} > u_i^{D*}$ and $u_j^{C*} > u_j^{D*}$ hold. Proposition 6 is proved.

Lastly, according to Propositions 2 and 4, there exists:

$$E\left[\tau^C(t)\right] - E\left[\tau^D(t)\right] = \frac{\Upsilon - \Omega}{\lambda + \eta}\left(1 - e^{-(\lambda+\eta)t}\right),$$
$$\lim_{t\to\infty} E\left[\tau^C(t)\right] - \lim_{t\to\infty} E\left[\tau^D(t)\right] = \frac{\Upsilon - \Omega}{\lambda + \eta} \tag{60}$$

Apparently, $\Upsilon < \Omega$ holds, then for any $t \in (0, \ \infty)$ exists:

$$E\left[\tau^C(t)\right] - E\left[\tau^D(t)\right] < 0, \qquad \lim_{t\to\infty} E\left[\tau^C(t)\right] - \lim_{t\to\infty} E\left[\tau^D(t)\right] < 0 \tag{61}$$

Similarly, it follows:

$$S\left[\tau^C(t)\right] - S\left[\tau^D(t)\right] = \frac{\sigma^2(\Upsilon - \Omega)\left(1 - 2e^{-(\lambda+\eta)t} + e^{-2(\lambda+\eta)t}\right)}{2(\lambda + \eta)^2}$$
$$\lim_{t\to\infty} S\left[\tau^C(t)\right] - \lim_{t\to\infty} S\left[\tau^D(t)\right] = \frac{\sigma^2(\Upsilon - \Omega)}{2(\lambda + \eta)^2} < 0 \tag{62}$$

Moreover, for $t \in (0, \ \infty)$, the first derivative of $1 - 2e^{-(\lambda+\eta)t} + e^{-2(\lambda+\eta)t}$ is greater than 0. Its value equals to 0 when $1 - 2e^{-(\lambda+\eta)t} + e^{-2(\lambda+\eta)t} = 0$, and then $S[\tau^C(t)] < S[\tau^D(t)]$ can be obtained. Proposition 7 is proved.

References

Aklin, M., Urpelainen, J., 2013. Political competition, path dependence, and the strategy of sustainable energy transitions. Am. J. Pol. Sci. 57 (3), 643–658.
Athanassoglou, S., Xepapadeas, A., 2012. Pollution control with uncertain stock dynamics: when, and how, to be precautious. J. Environ. Econ. Manag. 63 (3), 304–320.
Bardhan, P., 2002. Decentralization of governance and development. J. Econ. Perspect. 16 (4), 185–205.
Benchekroun, H., Martín-Herrán, G., 2016. The impact of foresight in a transboundary pollution game. Eur. J. Oper. Res. 251 (1), 300–309.
Bertinelli, L., Camacho, C., Zou, B., 2014. Carbon capture and storage and transboundary pollution: a differential game approach. Eur. J. Oper. Res. 237 (2), 721–728.
Breton, M., Martín-Herrán, G., Zaccour, G., 2006. Equilibrium investment strategies in foreign environmental projects. J. Optim. Theory Appl. 130 (1), 23–40.
Breton, M., Sokri, A., Zaccour, G., 2008. Incentive equilibrium in an overlapping-generations environmental game. Eur. J. Oper. Res. 185 (2), 687–699.
Breton, M., Zaccour, G., Zahaf, M., 2005. A differential game of joint implementation of environmental projects. Automatica 41 (10), 1737–1749.
Chang, S., Qin, W., Wang, X., 2018. Dynamic optimal strategies in transboundary pollution game under learning by doing. Physica A 490, 139–147.
Czaika, E., Selin, N.E., 2017. Model use in sustainability policy making: an experimental study. Environ. Model. Softw 98, 54–62.
Daskalakis, G., Psychoyios, D., Markellos, R.N., 2009. Modeling CO2 emission allowance prices and derivatives: evidence from the European trading scheme. J. Bank. Financ. 33 (7), 1230–1241.
Dellink, R., Finus, M., 2012. Uncertainty and climate treaties: does ignorance pay? Resour. Energy Econ. 34 (4), 565–584.
Dong, F., Li, X., Long, R., Liu, X., 2013. Regional carbon emission performance in China according to a stochastic frontier model. Renew. Sustain. Energy Rev. 28, 525–530.
El Ouardighi, F., Kogan, K., Gnecco, G., Sanguineti, M., 2018. Transboundary pollution control and environmental absorption efficiency management. Ann. Oper. Res. 1–29.
El Ouardighi, F., Sim, J.E., Kim, B., 2016. Pollution accumulation and abatement policy in a supply chain. Eur. J. Oper. Res. 248 (3), 982–996.
Funke, M., Paetz, M., 2011. Environmental policy under model uncertainty: a robust optimal control approach. Clim. Change 107 (3–4), 225–239.
Gu, Y., Yim, S.H.L., 2016. The air quality and health impacts of domestic transboundary pollution in various regions of China. Environ. Int. 97, 117–124.
Guan, X., Liu, W., Chen, M., 2016. Study on the ecological compensation standard for river basin water environment based on total pollutants control. Ecol. Indicat. 69, 446–452.
Hao, Y., Liu, Y.M., 2016. The influential factors of urban PM2. 5 concentrations in China: a spatial econometric analysis. J. Clean. Prod. 112, 1443–1453.
Haurie, A., Krawczyk, J.B., Roche, M., 1994. Monitoring cooperative equilibria in a stochastic differential game. J. Optim. Theory Appl. 81 (1), 73–95.
Heutel, G., 2012. How should environmental policy respond to business cycles? Optimal policy under persistent productivity shocks. Rev. Econ. Dyn. 15 (2), 244–264.
Huang, B.B., Wang, X.J., Gui, F.L., et al., 2011. The two-stage dynamic game of subject and object in ecological compensation under incomplete information. Syst. Eng-Theory. Pract. 31 (12), 2419–2424.
Huang, X., He, P., Zhang, W., 2016. A cooperative differential game of transboundary industrial pollution between two regions. J. Clean. Prod. 120, 43–52.
Jie, Y.M., Ge, Y.X., Xu, G.L., 2012. Cooperation evolution of ecological compensation between waterhead region and downstream based on evolutionary game. Oper. Res. Manag. Sci. 3, 023.
Jørgensen, S., 2010. A dynamic game of waste management. J. Econ. Dyn. Control 34 (2), 258–265.
Jørgensen, S., Yeung, D.W.K., 1996. Stochastic differential game model of a common property fishery. J. Optim. Theory Appl. 90 (2), 381–403.
Kahn, M.E., Li, P., Zhao, D., 2015. Water pollution progress at borders: the role of changes in China's political promotion incentives. Am. Econ. J. 7 (4), 223–242.
Kolstad, C.D., 2007. Systematic uncertainty in self-enforcing international environmental agreements. J. Environ. Econ. Manag. 53 (1), 68–79.
Li, S., 2014. A differential game of transboundary industrial pollution with emission permits trading. J. Optim. Theory Appl. 163 (2), 642–659.
Lin, P.C., Huang, Y.H., 2012. The influence factors on choice behavior regarding green products based on the theory of consumption values. J. Clean. Prod. 22 (1), 11–18.
List, J.A., Mason, C.F., 2001. Optimal institutional arrangements for transboundary pollutants in a second-best world: evidence from a differential game with asymmetric players. J. Environ. Econ. Manag. 42 (3), 277–296.
Liu, J., Liu, Q., Yang, H., 2016a. Assessing water scarcity by simultaneously considering environmental flow requirements, water quantity, and water quality. Ecol. Indicat. 60, 434–441.
Liu, G., Yang, Z., Chen, B., et al., 2016>b. Prevention and control policy analysis for energy-related regional pollution management in China. Appl. Energy 166, 292–300.
Liu, M., Yang, L., Min, Q., 2018. Establishment of an eco-compensation fund based on eco-services consumption. J. Environ. Manag. 211, 306–312.
Mäler, K.G., De Zeeuw, A., 1998. The acid rain differential game. Environ. Resour. Econ. 12 (2), 167–184.
Masoudi, N., Santugini, M., Zaccour, G., 2016. A dynamic game of emissions pollution with uncertainty and learning. Environ. Resour. Econ. 64 (3), 349–372.
Pamen, O.M., 2015. Optimal control for stochastic delay systems under model uncertainty: a stochastic differential game approach. J. Optim. Theory Appl. 167 (3), 998–1031.
Pan, X., Xu, L., Yang, Z., Yu, B., 2017. Payments for ecosystem services in China: policy, practice, and progress. J. Clean. Prod. 158, 200–208.
Prasad, A., Sethi, S.P., 2004. Competitive advertising under uncertainty: a stochastic differential game approach. J. Optim. Theory Appl. 123 (1), 163–185.
Rubinstein, A., 1982. Perfect equilibrium in a bargaining model. Econometrica: J. Econom. Soc. 97–109.
Shang, W., Gong, Y., Wang, Z., et al., 2018. Eco-compensation in China: theory, practices and suggestions for the future. J. Environ. Manag. 210, 162–170.
Shi, G.M., Wang, J.N., Zhang, B., et al., 2016. Pollution control costs of a transboundary river basin: empirical tests of the fairness and stability of cost allocation mechanisms using game theory. J. Environ. Manag. 177, 145–152.
Van den Bergh, J.C., 2010. Externality or sustainability economics? Ecol. Econ. 69 (11), 2047–2052.
Van Long, N., 2011. Dynamic games in the economics of natural resources: a survey. Dyn. Games. Appl. 1 (1), 115–148.
Vardar, B., Zaccour, G., 2018. The strategic impact of adaptation in a transboundary pollution dynamic game. Environ. Model. Assess. 23 (6), 653–669.
Wang, X., Berman, E.M., Chen, D.Y., et al., 2019. Strategies to improve environmental networks for pollution control: evidence from eco-compensation programs in China. J. Environ. Manag. 234, 387–395.
Wang, Y., Shen, N., 2016. Environmental regulation and environmental productivity: the case of China. Renew. Sustain. Energy Rev. 62, 758–766.
Xu, B., Lin, B., 2016. Regional differences of pollution emissions in China: contributing factors and mitigation strategies. J. Clean. Prod. 112, 1454–1463.
Xu, D.W., Tu, S.Y., Chang, L., 2012. Interest conflict of river basin ecological compensation based on evolutionary game theory. China. Popul. Resour. Environ. 22, 8–14.
Xu, J., Grumbine, R.E., Beckschäfer, P., 2014. Landscape transformation through the use of ecological and socioeconomic indicators in Xishuangbanna, Southwest China, Mekong Region. Ecol. Indicat. 36, 749–756.
Yang, L., 2017. Types and institutional design principles of collaborative governance

in a strong-government society: the case study of desertification control in northern China. Int. Public Manag. J. 20 (4), 586–623.
Yeung, D.W.K., 2001. Infinite-horizon stochastic differential games with branching payoffs. J. Optim. Theory Appl. 111 (2), 445–460.
Yeung, D.W.K., 1992. A differential game of industrial pollution management. Ann. Oper. Res. 37 (1), 297–311.
Yeung, D.W.K., 2007. Dynamically consistent cooperative solution in a differential game of transboundary industrial pollution. J. Optim. Theory Appl. 134 (1), 143–160.
Yeung, D.W.K., Petrosyan, L.A., 2005. Subgame consistent solutions of a cooperative stochastic differential game with nontransferable payoffs. J. Optim. Theory Appl. 124 (3), 701–724.
Yeung, D.W.K., Petrosyan, L.A., 2008. A cooperative stochastic differential game of transboundary industrial pollution. Automatica 44 (6), 1532–1544.
Yin, H., Pizzol, M., Xu, L., 2017. External costs of PM2.5 pollution in Beijing, China: uncertainty analysis of multiple health impacts and costs. Environ. Pollut. 226, 356–369.
Yu, B., Xu, L., 2016. Review of ecological compensation in hydropower development. Renew. Sustain. Energy Rev. 55, 729–738.
Zhang, Q., Jiang, X., Tong, D., et al., 2017. Transboundary health impacts of transported global air pollution and international trade. Nature 543 (7647), 705.
Zhao, L., Li, C., Huang, R., et al., 2013. Harmonizing model with transfer tax on water pollution across regional boundaries in a China's lake basin. Eur. J. Oper. Res. 225 (2), 377–382.
Zhao, L., Qian, Y., Huang, R., et al., 2012a. Model of transfer tax on transboundary water pollution in China's river basin. Oper. Res. Lett. 40 (3), 218–222.
Zhao, R., Neighbour, G., Han, J., et al., 2012b. Using game theory to describe strategy selection for environmental risk and carbon emissions reduction in the green supply chain. J. Loss. Prevent. Proc. 25 (6), 927–936.
Zwillinger, D., 1998. Handbook of Differential Equations. Academic Press, San Diego, California.