STOCHASTIC SWITCHING GAMES AND DUOPOLISTIC COMPETITION IN EMISSIONS MARKETS

MICHAEL LUDKOVSKI

Abstract. We study optimal behavior of energy producers under a CO_2 emission abatement program. We focus on a two-player discrete-time model where each producer is sequentially optimizing her emission and production schedules. The game-theoretic aspect is captured through a reduced-form price-impact model for the CO_2 allowance price. Such duopolistic competition results in a new type of a non-zero-sum stochastic switching game on finite horizon. Existence of game Nash equilibria is established through generalization to randomized switching strategies. No uniqueness is possible and we therefore consider a variety of correlated equilibrium mechanisms. We prove existence of correlated equilibrium points in switching games and give a recursive description of equilibrium game values. A simulation-based algorithm to solve for the game values is constructed and a numerical example is presented.

1. Introduction

In this paper we study a new class of non-zero-sum stochastic *switching games* with continuous state-space. Such games have natural applications in economics and finance, in particular for describing oligopolistic competition between large commodity producers. Our motivating example comes from the CO_2 cap-and-trade markets and our analysis provides new quantitative insight into the game-theoretic aspects of these schemes.

From a probabilistic perspective, a switching game is a repeated stopping game. It is characterized by a finite number of system states $\vec{u}$, jointly selected by the players. In addition, there are also system variables represented as controlled stochastic processes. The players dynamically react to the evolution of state variables and actions of other players by strategically modifying the system state. Overall, a switching game is a special class of dynamic non-zero-sum state-space games.

To our knowledge, such stochastic games have not been treated in detail before. Thus, our contribution is a first rigorous probabilistic analysis of switching games. The structure of game Nash equilibria in our model is delicate. In particular, we cannot guarantee *a priori* a unique equilibrium, so that an additional mechanism is needed for equilibrium selection. We propose to apply the wider concept of correlated equilibria. Our key result is the description of correlated equilibria in switching games in Section 4.4. The resulting representation in Theorem 4.4 of switching games in terms of a recursive sequence of stopping games leads to a constructive characterization of equilibrium strategies. Namely, we prove the analogue of the dynamic programming equation for the game

2000 *Mathematics Subject Classification.* 91A15, 60G40, 93E20, 91B76.

Key words and phrases. stopping games, optimal switching, correlated equilibrium, carbon trading.

values which enables numerical solution through backward recursion. Thus, the complexity of switching games is only slightly higher than for regular optimal switching problems.

In terms of existing literature this work can be seen as extending two separate strands of research. Work on stochastic *stopping games* dates back to Dynkin [15] who considered zero-sum games. Such games, now called Dynkin games, were progressively generalized, see [1, 11, 17, 20, 41]. Later extensions also treated special cases of non-zero-sum stopping games, especially the so-called monotone type [25, 34, 35]. The key tool of correlated equilibria in stochastic dynamic games was studied by [33, 37, 38, 39, 40]. Contemporaneously, the theory of *optimal switching* for a single agent was developed and extensively studied in the past decade. Let us cite [8, 12, 22, 36] for the associated results and the connections to optimal stopping. Combining the two literatures and especially following the methods of Ramsey and Szajowski [37], we construct an equilibrium point in the switching game. Simultaneously, we show that at equilibrium each player essentially faces an optimal switching problem.

Beyond the theoretical characterization, another of our contributions is a construction of numerical schemes to compute game-values and equilibrium strategies of switching games. This is achieved by combining backward recursion together with recursive solution of local 2×2 games. We suggest two approaches, one based on the Markov chain approximation method and a second algorithm that relies on least squares Monte Carlo. The latter is a novel extension of our previous work in [8, 30] and borrows ideas from standard optimal stopping theory to implement the analogue of the dynamic programming recursion on a set of Monte Carlo simulations.

A significant portion of the paper is dedicated to the application of our model to emissions trading. With imminent ramping up of CO_2-emissions markets around the world, it is crucial to understand energy producer behavior under the new frameworks. By design, the carbon allowances will be scarce. As a result, the proposed CO_2 cap-and-trade schemes will lead to (oligopolistic) producers competing for finite permit resources. Following the well-established approach of using game-theoretic methods in models of pollution and environmental impact (see e.g. the textbook [14]), our analysis sheds new light on the implications of such dynamic competition. Overall, our model is a first attempt to investigate oligopolistic behavior in CO_2 markets and can serve as a stepping-stone to more sophisticated modeling that addresses global market design (e.g. allocation of allowances) and comparative statics of our framework.

The rest of the paper is organized as follows. In Section 2 we provide a general discussion of oligopolistic competition in commodity production, followed by a description in Section 3 of the precise stochastic model we employ. Section 4 constructs the representation of switching games in terms of repeated stopping games. Section 5 describes our numerical solution algorithms and presents a computational example. Finally, Section 6 discusses extensions of the model, notably to a continuous-time setting, and points directions for future work.

2. Competitive Equilibrium among Oligopolistic Producers

In this section we expand on the nature of the competitive dynamic equilibrium between producers with market power.

2.1. **Producer Behavior under Cap-and-Trade Emission Schemes.**

The pending expansion of CO_2 emission cap-and-trade schemes will lead to new competitive phenomena in energy markets. In particular, the possibility of trading the carbon allowances will create a new major commodity market. This market will differ from classical commodities due to its limited size (the total number of allowances will be scarce) and its extreme seasonality (permits will be distributed each calendar year and expire at year-end). Also, it will be intrinsically linked to other energy markets, notably electricity whose production accounts for the bulk of regulated emissions. Last but not least, it is likely that major carbon polluters will have strong influence in the carbon market and the ability to dramatically move carbon prices based on their emission schedules. Hence, to understand equilibria in CO_2-markets it is useful to consider them from the point of view of *large traders*.

The model proposed below captures this price impact for a joint electricity-carbon market. To fix ideas, consider two heterogenous producers (henceforth termed players) who each produce commodity P (electricity) and consume commodity X (carbon allowances). These two producers generate “dirty” electricity from e.g. coal or gas and can be viewed as representative agents of a park of power plants with identical engineering characteristics. We assume that the two players are large traders in the carbon market, but small players on the electricity market. This reflects the fact that the other “green” producers (who use nuclear, hydroelectric, renewable, etc. sources) create a competitive electricity market while remaining passive in the CO_2-permits arena. Thus, the players are price-takers for P but can each influence the price of X; at date t these prices are denoted as P_t and X_t. All the other participants in the electricity and CO_2 markets are not modeled explicitly; rather we postulate that their collective actions induce *stochastic* fluctuations in the respective commodity prices.

The key feature of our model is that *both* producers can affect the dynamics of the carbon allowance price, leading to a non-zero sum stochastic game between them. The objective of the producers is to maximize their expected net profits over the planning horizon, generally corresponding to the expiry date of the carbon permits. Once the CO_2 market is running, the producers' profit will depend on their *clean dark spread* [18] which is defined as the difference between electricity price and the carbon-adjusted production cost (we assume that input fuel costs are fixed as is often the case for power generators with long-term supply contracts). The strategy of each player is described by a repeated start-up/shut-down option. Namely, if the market conditions are unfavorable, a player can stop production, eliminate CO_2 emissions and avoid losses; she can then restart production when the profit spread improves. As a first approximation we assume that these choices of production regimes are binary and denoted as “off” (0) and “on” (1). In a single-player model, such timing optionality is known as a *real option* and has been thoroughly investigated since the seminal work of [4, 13]. Repeated real options have attracted considerable attention recently, see [8, 22, 30], and others.

Moving to the duopoly setting, the presence of multiple large traders in the market leads to competitive effects. While each player is maximizing her own profits, the competitor actions will also affect her decisions. Indeed, because the total number of allowances is limited, emissions today shrink remaining permit supplies and tend to increase *future* CO_2 prices. Therefore, if player 1 is producing, player's 2 expected future profits are reduced through this feedback effect. Overall, the producers are facing a stochastic game where actions correspond to the latest choice of production regime by each player and payoffs are a function of the exogenous price P_t and the endogenous price X_t that is partly controlled by the players. Following classical game theory, competitive equilibrium is then typically described through a Nash equilibrium point (NEP) of this game. Since the game is stochastic and multi-period with Markov state variables, we restrict our attention to Markovian (feedback) Nash equilibria. We are now interested in characterizing any and all such game equilibria and then computing the corresponding game value functions (which describe expected profits) and equilibrium emission schedules.

2.2. **Types of Competitive Equilibrium.** Intuitively, the dynamic switching game is a sequence of one-period bimatrix games. At each stage t, the control or action $u_i(t) \in \{0, 1\}$ of each player $i \in \{1, 2\}$ is simply "on" or "off", leading to the classic 2×2 game. From a dynamic point of view, the relevant payoffs to the players at stage t are then the sum of the current clean spreads and the *continuation values* that correspond to the optimal game value that can be realized in the future by the respective player contingent on current state of the world. In turn, these continuation values depend on the future equilibria that will be implemented.

Let us recall that a one-shot 2×2 game that has no weakly dominant actions (see Section 3.3), can be classified as belonging to one of three possible types [5]:

(1) *Standard* game. This is the basic case where one action strictly dominates the other for at least one player and a single pure Nash equilibrium point exists.
(2) *Competitive* game. No pure Nash equilibria exist, but a single mixed-strategy equilibrium point is available.
(3) *(anti-) Coordination* game which yields two pure Nash equilibria, plus an additional mixed Nash equilibrium point.

All of the above three types of 2×2 games are relevant for the analysis of competitive equilibrium between the two producers of Section 2.1. Because the (P, X)-prices are stochastic, it is impossible to *a priori* rule out some of the above scenarios; in fact depending on current input/output prices we expect that all three game types can emerge locally. In particular, the third case of "battle-of-the-sexes" or "chicken" game is likely to appear in the competitive duopoly case when the market conditions are only able to support one producer. Thus, when the electricity-carbon spread is slightly positive, each of the players will have an incentive to emit. However, if the price impact is strong enough, it is not profitable for *both* of them to consume permits due to the associated future negative externalities. As a result, two pure Nash equilibria are possible whereby one producer *yields* the market to the other. Hence, the two players try to achieve anti-coordination, taking the opposite actions of each other.

In our repeated game setting, the players must *a priori* agree on how to implement future equilibria, otherwise the outlined computation of continuation values would not be possible. Consequently, without a clear *equilibrium selection* mechanism, the solution of the full switching game remains ill-defined.

Remark 2.1. An alternative formulation is to take the controls $u_i(t)$ to be continuous, so that the producers can choose the emissions level precisely rather than working with the binary "on/off" controls. This would lead to a model of a non-zero-sum stochastic dynamic game. Such models have been extensively studied in the past, both in discrete and continuous time, see for example [24, 32] and references therein. The problem of equilibrium selection is sometimes less severe with continuous controls thanks to the convexity of value functions. In this work we choose to focus on the timing flexibility and therefore maintain the discrete control space that can be interpreted in terms of (sequential) optimal stopping. Otherwise, the distinction is similar to that between real options models that analyze the best time to begin a project (timing option) vis-a-vis models that focus on gradual capital accumulation (singular/continuous controls).

2.3. **Equilibrium Selection.** To describe equilibrium selection, we expand the notion of game equilibria to the larger class of correlated equilibrium points. Namely, we shall assume that the players engage in communication at each stage of the game which allows them to *correlate* their decisions. This communication device is formally represented by a fictitious third party that directs the players to implement a particular strategy profile via (randomized) private signals, see Section 4.2. The meaning of the resulting *correlated equilibrium* is that conditional on the private signal, neither player has an incentive to deviate from the prescribed action, see condition (9) below.

Economically, the alluded third party is either a real entity, such as a government regulator or market watchdog, or is a proxy for market frictions that make one equilibrium most preferable. Thus, in spite of the description in the previous paragraph, no inherent collusion is required; correlated decision-making can emerge thanks to factors that are outside our modeling scope or through repeated learning, etc. The game is still non-cooperative and differs markedly from monopoly. If the players are directed by a regulator, the latter may select a socially beneficial equilibrium. For instance, a "utilitarian" communication device maximizes the (weighted) sum of the firms continuation values so that the producers as a whole have best economic health. Alternatively, a "green" device chooses the equilibrium that minimizes CO_2 emissions. On the other hand, a "preferential" communication mechanism can endogenously emerge without a third party due to extra advantages available to a given player (e.g. due to preferential regulatory treatment or other externalities).

In the case-study below (Section 5.3 we will examine the following possibilities:

- A *utilitarian* mechanism, maximizing the sum of game values for the players;
- An *egalitarian* (Rawls) mechanism, maximizing the minimum game value of the players;

- A fixed *preferential* mechanism, maximizing the game value for player 1 (resp. player 2).

Once a (contingent) correlated equilibrium device is fixed for each stage of the dynamic game, we will show that the full competitive producer equilibrium problem has a well-defined game value and equilibrium production schedules. Its solution can then be obtained through a variant of the dynamic programming paradigm that is used in ordinary stochastic control problems. Namely, we proceed through backward recursion, iteratively solving one-period 2×2-games to compute the required continuation values. The formal details of this construction will be presented in Section 4. In preparation, the next section completes our mathematical setup and culminates with a rigorous formulation of the game objectives.

Remark 2.2. A related concept of competitive equilibrium in the industrial organization literature is that of a stage Stackelberg game [3]. In a Stackelberg game, at each stage one player is the *leader* and has priority in making decisions; the second player then follows. This description corresponds closely to the preferential mechanism of equilibrium selection which always favors a given player.

3. Optimization Problem

3.1. **Price Dynamics.** Start with a filtered probability space $(\Omega, \mathcal{H}, (\mathcal{F}_t), \mathbb{P})$, $t \in \mathbb{T} \triangleq \{0, 1, \dots, T\}$. The full σ-algebra $\mathcal{H}$ is strictly bigger than $\mathcal{F}_T$, in particular big enough to support extra randomization parameters needed below.

The electricity price is given by the discrete-time stochastic process (P_t), for simplicity taken to be one-dimensional,

$$P_{t+1} = a(P_t, \epsilon_t^P),$$

where the innovations (ϵ_t^P) are independent. Our canonical example is the logarithmic Ornstein-Uhlenbeck process which is a log-normal stationary Gaussian process with standard Gaussian i.i.d. ϵ^P's,

$$P_{t+1} = P_t \cdot \exp\left(\kappa_P(\bar{P} - \log P_t) \; + \sigma_P \epsilon_t^P\right), \qquad \epsilon_t^P \sim \mathcal{N}(0,1), \tag{1}$$

for some positive constants $\kappa_P, \bar{P}, \sigma_P$.

Let X_t be the carbon price at date t. We model X_t as another mean-reverting process, with a policy-dependent mean and log-Gaussian increments. Namely, conditional on player actions $u_1(t), u_2(t)$,

$$\begin{cases} \quad X_{t+1} = X_t \cdot \exp\left(\kappa_X(f(u_1(t), u_2(t)) - \log X_t) \; + \sigma_X \epsilon_t^X\right) \quad \text{with} \\ f(u_1, u_2) = \log(\bar{X} + g_1 u_1 + g_2 u_2). \end{cases} \tag{2}$$

The sequence (ϵ_t^X) are again standard Gaussian random variables, with correlation ρ to (ϵ_t^P), i.e. $\epsilon_t^X = \rho \epsilon_t^P + \sqrt{1-\rho^2}\epsilon_t^\perp$ with $\epsilon_t^\perp \sim \mathcal{N}(0,1)$ independent of ϵ_t^P. Rising electricity prices are likely to increase the overall CO_2 emission rates and therefore we expect that P and X are positively correlated, $\rho > 0$.

Remark 3.1. To motivate the meaning of (2) in a carbon market, we recall from [19] that one expects "$X_t = \bar{x}\bar{\mathbb{P}}\{C_T > \bar{c}|\mathcal{F}_t\}$" where C_T is the cumulative total CO_2 emission on $[0,T]$, $\bar{x}$ is the penalty for going over the allowance limit and $\bar{c}$ is the total amount of allowances allocated. We postulate that $C_T = \sum_{s=0}^{T-1}\left\{(\sum_{i=1}^{2} b_i u_i(s)) + \tilde{u}(s)\right\}$ where $g_i u_i(s)$ are the emissions of producer i in period s and $\tilde{u}(s)$ are the emissions by all the other market participants. Assuming independent stochastic increments (due to external shocks such as weather effects, etc.) in $\tilde{u}(s)$, the dynamics (2) follow, with some complicated and time-dependent functions $f(t,\cdot)$ and coefficients $\kappa_X(t)$, $\sigma_X(t)$. Moreover, the above $\bar{\mathbb{P}}$ should be the equilibrium measure, incorporating future state-contingent equilibrium emission strategies $(u_1^*, u_2^*)(s)$ for $s > t$. To sidestep these challenges, (2) can be viewed as a simplified (reduced-form) version of this description chosen to succinctly capture the temporal feedback effect between u_i's and X_t. The form of (2) is only to fix ideas and much more complex equations can be considered; the subsequent analysis makes only minor use of (2). Finally, let us note that the form of the price impact $f(u_1,u_2)$ is crucial to obtain an "interesting" game setting. Without price impact the game becomes degenerate and reduces to standard optimization problems for each producer. On the contrary, if the supply curve for the CO_2 allowances is convex, then this price impact would be nonlinear and further magnify the competitive effects.

As discussed above, the producers have simplified *binary* production schedules, so that $u_i \equiv (u_i(t))$ are $\{0,1\}$-valued controls. Moreover, we assume that the producers have zero allowance allocations and cannot bank allowances; therefore they must purchase the requisite allowances separately at each stage of the game. The total P&L of the electricity producers then consists of (i) revenue from selling electricity, minus the (ii) cost of buying emission allowances, as well as (iii) operational costs due to adopted strategy u_i. An important case of operational costs are *fixed switching costs* $K_{\{i,j_1,j_2\}}$ that are paid each time the production regime of agent i is changed from j_1 to j_2 and corresponding to the ramping-up/winding-down costs associated with the electricity turbines [8, 18]. We set $K_{\{i,j_1,j_1\}} = 0 \forall j_1$ so costs are charged only at production regime switches.

Given a production schedule u_i the net profit of producer i up to terminal date T is then

$$\sum_{t=0}^{T-1}\left\{(a_i P_t - b_i X_t - c_i)u_i(t) \quad - K_{\{i,u_i(t-),u_i(t)\}}\right\}, \tag{3}$$

for some constants $a_i, b_i, c_i, K_{\{i,j_1,j_2\}}$, $i = 1,2$, representing amount of electricity produced by the facility, the amount of corresponding CO_2 allowances needed, fixed production costs and switching costs, respectively. The ratio a_i/b_i represents the carbon efficiency of producer i. Below, the theorems on existence of equilibria in stochastic games require bounded payoffs; therefore we formally assume that payoffs are truncated as $\psi_i(p,x,u) \triangleq \bar{N} \wedge (a_i p - b_i x - c_i)u$ for some large constant $\bar{N}$.

3.2. **Randomized Emission Schedules.** Denote by $\mathcal{F} = (\mathcal{F}_t)$, with $\mathcal{F}_t = \sigma(X_0, P_0, \dots, X_t, P_t)$, the filtration generated by the price histories. The strategies u_i may be mixed or randomized; recall that a randomized $\mathcal{F}$-stopping time is an $\mathcal{F}$-adapted stochastic process $p = (p_t)$ with $0 \le p_t \le 1$, a.s. The corresponding stopping time is

$$\tau(p) \triangleq \inf\{t : \eta_t \leq p_t\}, \qquad \eta_t \sim Unif(0,1) \perp \mathcal{F}_T, \text{ i.i.d.} \tag{4}$$

Thus, p_t is interpreted as the probability of stopping at date t, conditional on not stopping so far. When $p_t \in \{0, 1\}$ for all t, we are back in the case of regular $\mathcal{F}$-stopping times. Otherwise, $\tau(p)$ is not $\mathcal{F}$-adapted because its value depends on η's that are independent of $(\mathcal{F}_t)$. However, one may enlarge the filtration so that $\{\tau(p) \leq t\}$ is $\sigma(\mathcal{F}_t, \eta_s : 0 \leq s \leq t)$-measurable. Thus, given any enlarged filtration $(\tilde{\mathcal{G}}_t)$ satisfying $\mathcal{F}_t \subset \tilde{\mathcal{G}}_t$ and $\mathbb{P}(A|\mathcal{F}_t) = \mathbb{P}(A|\tilde{\mathcal{G}}_t)$ for all $t \leq T, A \in \mathcal{F}_T$, we may identify any $\tilde{\mathcal{G}}$-adapted stopping time with a $\mathcal{F}$-randomized stopping time. We then denote by $\mathfrak{D}^{\tilde{\mathcal{G}}}(t)$ the set of all such randomized $\mathcal{F}$-stopping times bigger than t.

In (8) below we define $\mathcal{G}^i_t \supseteq \mathcal{F}_t$ which is the private information filtration of player i; the corresponding set of stopping times (viewed as $\mathcal{F}$-randomized times) is denoted by $\mathfrak{D}^i$. An admissible production schedule pair $(u_1(t), u_2(t))$ consists of $\mathcal{G}^i$-adapted $\{0, 1\}$-valued processes. Because $u_i(t) \in \{0, 1\}$, we have a one-to-one correspondence between admissible u_i's and sequences $(\tau^u_k)_{k=1}^\infty$ satisfying $\tau^u_{k+1} \in \mathfrak{D}^i(\tau^u_k)$, via the alternative representation (using $\bar{i} \equiv 1 - i$)

$$u_i(t) = \sum_{k=0}^{\infty} u_i(0) 1_{[\tau^u_{2k}, \tau^u_{2k+1})} + \overline{u_i(0)} 1_{[\tau^u_{2k+1}, \tau^u_{2k+2})}, \qquad \tau^u_0 = 0. \tag{5}$$

The switching times τ^u_k encode the times of production regime shifts defined by u_i. The representation (5) holds because at most one regime switch can be made by each player at any given stage. Indeed, multiple simultaneous regime switches by the same producer are strongly sub-optimal if $K_{\{i,j_1,j_2\}} > 0$ and weakly suboptimal otherwise.

A third representation of u_i can be given in terms of the one-step *switching* probabilities (p_t). Namely, $u_i \equiv (p_t, \eta_t)$ where $0 \leq p_t \leq 1$ is a $\mathcal{F}$-adapted process and η_t is an i.i.d. sequence of $Unif(0, 1)$ random variables, independent of $\mathcal{F}_T$ and with $\eta_t \in \mathcal{G}^i_t$. Given such (p_t, η_t) we define the $\mathcal{G}^i$-stopping time

$$\tau^u_{k+1} = \inf\{t > \tau^u_k : \eta_t \leq p_t\}$$

and then use (5).

Let $\mathbb{P}^{\vec{u}}$ be the law of (P_t, X_t) given a strategy pair $\vec{u} \equiv (u_1(t), u_2(t))_{t=0}^T$. We also use the notation $\mathbb{P}^{\vec{\zeta}}$ to denote the law of (P_t, X_t) under a constant strategy profile $\vec{u}(t) \equiv \vec{\zeta}$. The interpretation of the randomization for the dynamics of X_t is straightforward: if p^i_t denotes the probability of changing the regime at date t by player i then

$$X_{t+1} = \begin{cases} X_t \cdot \exp(\kappa_X(f(u_1(t), u_2(t)) - \log X_t) \ + \sigma_X \epsilon^X_t) & \text{on the event} \quad \{\omega : p^1_t \leq \eta^1_t, p^2_t \leq \eta^2_t\}; \\ X_t \cdot \exp(\kappa_X(f(\overline{u_1(t)}, u_2(t)) - \log X_t) \ + \sigma_X \epsilon^X_t) & \text{on the event} \quad \{\omega : p^1_t > \eta^1_t, p^2_t \leq \eta^2_t\}; \\ X_t \cdot \exp(\kappa_X(f(u_1(t), \overline{u_2(t)}) - \log X_t) \ + \sigma_X \epsilon^X_t) & \text{on the event} \quad \{\omega : p^1_t \leq \eta^1_t, p^2_t > \eta^2_t\}; \\ X_t \cdot \exp(\kappa_X(f(\overline{u_1(t)}, \overline{u_2(t)}) - \log X_t) \ + \sigma_X \epsilon^X_t) & \text{on the event} \quad \{\omega : p^1_t > \eta^1_t, p^2_t > \eta^2_t\}. \end{cases}$$

The corresponding expected cumulative profit starting with $P_s = p_0$, $X_s = x_0$ and initial production regime $\vec{\zeta} \in \{0,1\}^2$ is

(6)

$$V_i(s, p_0, x_0, \vec{\zeta}; \vec{u}) = \mathbb{E}^{\vec{u}} \left[\sum_{t=s}^{T-1} \left\{ (a_i P_t - b_i X_t - c_i) u_i(t) \quad - K_{\{i, u_i(t-), u_i(t)\}} \right\} \Big| P_s = p_0, X_s = x_0, \vec{u}(s-) = \vec{\zeta} \right].$$

Observe that a producer may immediately change her production mode at $t = s$, so that the initial production regime is interpreted as occurring just before s.

3.3. **Equilibria in 2×2 One Period Games.** To define the notion of game equilibrium with the player objectives given by (6) we first recall basic results on Nash equilibria in bimatrix games, see [3, 21].

Consider a 2-by-2 one-shot game H with payoffs $(\alpha^{ij}, \beta^{ij})$, $i, j \in \{0, 1\}$. The value $\alpha^{ij} \in \mathbb{R}$ (resp. β^{ij}) defines the payoff to player 1 (resp. player 2) when player 1 chooses action $i \in \{0,1\}$ and player 2 chooses action $j \in \{0,1\}$. In matrix notation the normal form of H is

$$H = \begin{pmatrix} (\alpha^{00}, \beta^{00}) & (\alpha^{01}, \beta^{01}) \\ (\alpha^{10}, \beta^{10}) & (\alpha^{11}, \beta^{11}) \end{pmatrix}, \tag{7}$$

where the rows of H are interpreted as action choices of player 1, and the columns correspond to the actions of player 2. A *strategy* of player i is a vector $\vec{\pi}_i \equiv (\pi_i^0, \pi_i^1)$ whence π_i^j is the probability that player i chooses action j. Thus, $\vec{\pi}_i$ is in the 2-simplex $\Delta_2 \triangleq \{(\pi^0, \pi^1) : \pi^j \geq 0, \pi^0 + \pi^1 = 1\}$. If $\pi_i^0 \cdot \pi_i^1 = 0$ then the strategy is *pure*; that is the choice of action is deterministic. Otherwise, the strategy is mixed. A *strategy profile* is a pair $(\vec{\pi}_1, \vec{\pi}_2)$ specifying the strategies of each player. We use the notation $\vec{\pi}_{-i} \equiv \vec{\pi}_{3-i}$ to denote the strategy of the player *other* than i.

Given a strategy profile, the *value* to player 1 is $A(\vec{\pi}_1, \vec{\pi}_2) \triangleq \sum_{j,k} \pi_1^j \pi_2^k \alpha^{jk}$ and is $B(\vec{\pi}_1, \vec{\pi}_2) \triangleq \sum_{j,k} \pi_1^j \pi_2^k \beta^{jk}$ to player 2. The above shows that a strategy profile $(\vec{\pi}_1, \vec{\pi}_2)$ can be viewed as a probability distribution on the corresponding payoff space, given by the product measure $\vec{\pi}_1 \times \vec{\pi}_2$ of the marginal strategies. This factorization is due to the fact that agents make *independent* decisions about their actions.

A strategy profile $(\vec{\pi}_1^*, \vec{\pi}_2^*)$ is a *Nash equilibrium point* (NEP) of H if we have

$$\begin{cases} A(\vec{\pi}_1^*, \vec{\pi}_2^*) \in \displaystyle\operatorname*{arg\,sup}_{(\pi_1^0, \pi_1^1) \in \Delta_2} \sum_{j,k} \pi_1^j (\pi_2^*)^k \alpha^{jk}, \\ B(\vec{\pi}_1^*, \vec{\pi}_2^*) \in \displaystyle\operatorname*{arg\,sup}_{(\pi_2^0, \pi_2^1) \in \Delta_2} \sum_{j,k} (\pi_1^*)^j \pi_2^k \beta^{jk}. \end{cases}$$

Hence, $\vec{\pi}_i^*$ is a best-response for player i, given that the other player uses $\vec{\pi}_{-i}^*$. We denote by

$$\mathcal{E}(H) = \{(A, B)(\vec{\pi}_1, \vec{\pi}_2) : (\vec{\pi}_1, \vec{\pi}_2) \text{ is a NEP of } H\},$$

the set of all game values corresponding to Nash equilibria points. The subset of NEP corresponding to pure strategy profiles is denoted by $\mathcal{E}_p$ and can be directly characterized as [35]

$$\mathcal{E}_p(H) = \left\{ (\alpha^{i^* j^*}, \beta^{i^* j^*}) : \alpha^{i^* j^*} \geq \alpha^{i j^*}, \beta^{i^* j^*} \geq \beta^{i^* j} \ \forall i, j \right\}.$$

We say that action i of agent 1 dominates (resp. weakly dominates) action i' if the expected payoff from implementing i is strictly bigger (resp. no smaller) than the expected payoff of i', $\alpha^{ij} > \alpha^{i'j}$ for all j. A NEP involving weakly dominant actions is subject to "trembling hand" deviations and will be ruled out in the sequel. As mentioned in Section 2, depending on three possible game types, $\mathcal{E}_p$ is then either empty, consists of a singleton or contains two pure Nash equilibrium points. The classical theorem of Nash shows that the full $|\mathcal{E}| > 0$ is always non-empty.

The above classification implies that to establish existence of NEP, one must consider *mixed* strategies; furthermore uniqueness is usually unavailable. In order to have a well-defined game value, we employ equilibrium refinement methods. Our construction relies on a correlation mechanism $\gamma \equiv (\gamma_{ij}) \in \Delta_4$ [2, 31]. A correlated equilibrium is a probability distribution γ_{ij} on the strategy profile space, which is communicated to the players using signals $\mu_i(\gamma) \in \{0, 1\}$, defined by $\mu_1(\gamma) = 1_{\{\gamma_{10}+\gamma_{11}<\eta_1\}}$ and $\mu_2(\gamma) = 1_{\{\gamma_{01}+\gamma_{11}<\eta_2\}}$ where η_1, η_2 are two independent uniform random variables on $[0, 1]$, the randomization parameters. The η's are not observed; each player only sees her signal μ_i. Conditional on the signal, the agent can impute the conditional strategy of the other player by e.g. $\vec{\pi}_2(\gamma)\big|_{\mu_1(\gamma)=0} = (\frac{\gamma_{00}}{\gamma_{00}+\gamma_{01}}, \frac{\gamma_{01}}{\gamma_{00}+\gamma_{01}})$. By definition of $\mu_i(\gamma)$, the resulting strategy profile $(\vec{\pi}_1(\gamma), \vec{\pi}_2(\gamma))$ has dependent marginals and joint distribution γ.

The meaning of equilibrium is that given a signal $\mu_i(\gamma)$ and the implied strategy $\pi_{-i}(\gamma)|\mu_i(\gamma)$ by the other player, player i has no incentive to deviate from the signalled action $\mu_i(\gamma)$. For 2×2 games, since the only deviation is choosing the opposite action, this reduces to (see [2, 5]),

Definition 3.1. *A probability distribution γ on $\{0, 1\}^2$ is a* correlated equilibrium point *(CEP) of the bimatrix game H in* (7) *if*

$$\begin{cases} \gamma_{00}\alpha^{00} + \gamma_{01}\alpha^{01} \geq \gamma_{00}\alpha^{10} + \gamma_{01}\alpha^{11}, & \gamma_{11}\alpha^{11} + \gamma_{10}\alpha^{10} \geq \gamma_{11}\alpha^{01} + \gamma_{10}\alpha^{00}, \\ \gamma_{00}\beta^{00} + \gamma_{10}\beta^{10} \geq \gamma_{00}\beta^{01} + \gamma_{10}\beta^{11}, & \gamma_{11}\beta^{11} + \gamma_{01}\beta^{01} \geq \gamma_{11}\beta^{10} + \gamma_{01}\beta^{00}. \end{cases}$$

For instance, the first inequality means that conditional on player 1 signal being $\mu_1(\gamma) = 0$, the expected payoff to player 1 from action 0 (the right-hand-side) is better than the expected payoff from action 1. In either scenario, player 2 implements the conditional strategy $(\frac{\gamma_{00}}{\gamma_{00}+\gamma_{01}}, \frac{\gamma_{01}}{\gamma_{00}+\gamma_{01}})$.

We recall that the set of correlated equilibria points, denoted $\mathcal{E}_c \supseteq \mathcal{E}$ includes the convex hull of all Nash equilibrium points and therefore allows randomization over which NEP is chosen (if more than one is available). Given a correlation device γ, the resulting game value is denoted

$$Val_\gamma(H) \triangleq \begin{pmatrix} \sum_{i,j} \gamma_{ij}\alpha^{ij} \\ \sum_{i,j} \gamma_{ij}\beta^{ij} \end{pmatrix}.$$

Except for the strictly-competitive games, one may always find correlation devices that correspond to pure Nash equilibria so that randomization (either by players or regulator) would not be needed otherwise. Nevertheless, to maintain generality we continue to work with general fully-mixed correlation devices.

Remark 3.2. For standard and strictly-competitive games, the set of CEP coincides with the unique NEP available. On the other hand, a continuum of correlated equilibria exist in the (anti-) coordination case [5]. In particular, new Pareto-efficient fully mixed equilibria can be obtained by correlating the actions of each player and reducing the wasteful non-coordinated outcomes that would arise from independent randomizations.

3.4. **Formal Objective.** We now combine the performance criteria in (6) with the concept of correlated equilibria to state the final optimization problem. Let $\gamma \equiv (\gamma_t)$ be a fixed Markovian communication device, i.e. a map $\gamma_t : (s, p, x, \vec{\zeta}) \to \Delta_4$. As before, γ produces a private correlation signal $\mu_i(t; \gamma)$ for agent $i \in \{1, 2\}$ at each stage t. We denote the resulting strategy profile as $\vec{u}(t; \gamma) = (\mu_1(t; \gamma), \mu_2(t; \gamma))$ and assume a full-information setting, whereby the emission schedules of each agent are publicly known. Define

$$\mathcal{G}_t^i = \sigma(\mathcal{F}_t, \mu_i(s; \gamma), 0 \le s \le t). \tag{8}$$

The set of admissible strategies for player i is $\mathcal{U}_i$, consisting of all $\mathcal{G}^i$-adapted switching randomized controls u_i. Since the equilibrium strategy of player i at each stage is pure given the signal μ_i, one can in fact restrict admissible strategies to just $\mathcal{G}^i$-adapted switching controls (in the same way as one can restrict attention to pure stopping times in optimal stopping problems).

Definition 3.2. *A correlated equilibrium point (CEP) for the switching game is a communication device γ and a joint production schedule $\vec{u}^*$ such that $\forall (s, p_0, x_0, \vec{\zeta})$ (recall definition of V_i in (6))*

$$\begin{cases} V_1(\cdot; u_1^*, u_2^*) \ge V_1(\cdot; u_1, u_2^*) & \forall u_1 \in \mathcal{U}_1, \\ V_2(\cdot; u_1^*, u_2^*) \ge V_2(\cdot; u_1^*, u_2) & \forall u_2 \in \mathcal{U}_2. \end{cases} \tag{9}$$

The resulting game value is denoted simply as $V_i(\cdot; \gamma)$.

Note that in (9), even if a player chooses to deviate from the recommendation $\mu_i(t; \gamma)$ she continues to receive future signals $\mu_i(s; \gamma)$, $s > t$ and therefore information about the implied strategy of the other player. Existence of CEP of switching games will be established in Theorem 4.4. We will also provide a recursive construction of $V_i(t, \cdot)$ in terms of conditional expectations of $V_i(t + 1, \cdot)$ and one-shot 2×2 games.

4. Sequential Stopping Game

Our analysis of the switching game will consist of building up the solution in several steps. We begin with analyzing the single-agent objective. Next, in Section 4.2 we move on to the one-shot non-zero-sum stopping game that is built iteratively from the one-period 2×2 games, following the methods of [37]. Finally, in Section 4.4 we describe the sequential stopping game that in the limit leads to our original model and definition (9).

4.1. **Single Producer Problem.** Before tackling the stochastic duopoly game, let us briefly review the solution of the single-player model. Since the control $u_i(t)$ takes on a finite number of values, we have an optimal switching model that can be viewed as a sequence of optimal stopping problems. Such models (including price impact) were studied in [8, 30].

Let us consider the optimization for producer 1. For the remainder of this section we fix a production schedule u_2 of the second producer, as well as a communcation device γ that sends private signals $\mu_1(\gamma)$ to player 1. In the single-producer problem, the objective is to maximize the expected profit

$$\sup_{(u(t))\in\mathcal{U}_1} \mathbb{E}^{(u,u_2,\gamma)}\left[\sum_{t=0}^{T-1}\left\{(a_1P_t - b_1X_t - c_1)u(t) \;\; - K_{\{1,u(t-),u(t)\}}\right\}\right]. \tag{10}$$

Because producer 1 receives private signals $\mu_1(t;\gamma)$, the set of admissible controls is again $\mathcal{U}_1$, i.e. $\mathcal{G}^1$-adapted.

Consider initial conditions $P_s = p, X_s = x, u_2(s) = \zeta_2$ and let $V(s,p,x,\zeta_2)$ be the value function corresponding to (10) conditional on starting in the "on"-production regime, and $W(s,p,x,\zeta_2)$ the value function starting offline. Furthermore, using same initial conditions define recursively

(11)

$$\begin{cases} V^0(s,p,x,\zeta_2) = \mathbb{E}^{(1,u_2,\gamma)}\left[\sum_{t=s}^{T-1}(a_1P_t - b_1X_t - c_1)\right], \qquad \text{as well as} \quad W^0(s,p,x,\zeta_2) = 0; \\ V^n(s,p,x,\zeta_2) = \sup_{\tau\in\mathfrak{D}^1(s)} \mathbb{E}^{(1,u_2,\gamma)}\left[\sum_{t=s}^{\tau-1}(a_1P_t - b_1X_t - c_1) + (W^{n-1}(\tau,P_\tau,X_\tau,u_2(\tau)) - K_{\{1,1,0\}})\right]; \\ W^n(s,p,x,\zeta_2) = \sup_{\tau\in\mathfrak{D}^1(s)} \mathbb{E}^{(0,u_2,\gamma)}\left[V^{n-1}(\tau,P_\tau,X_\tau,u_2(\tau)) - K_{\{1,0,1\}}\right], \qquad n \ge 1 \end{cases}$$

where under $\mathbb{P}^{(i,u_2,\gamma)}$ the drift of the carbon allowance price is $f(i,u_2(t))$.

Proposition 4.1. *Let* $\mathcal{U}_1^n \triangleq \{u \in \mathcal{U}_1 : u \text{ has at most } n \text{ switches}\}$. *Then,*

$$V^n(s,p,x,\zeta_2) = \sup_{(u(t))\in\mathcal{U}_1^n, u(s-)=1} \mathbb{E}^{(u,u_2,\gamma)}\left[\sum_{t=s}^{T-1}\left\{(a_1P_t - b_1X_t - c_1)u(t) \;\; - K_{\{1,u(t-),u(t)\}}\right\}\right],$$

and as $n\to\infty$, $V^n(s,p,x,\zeta_2) \to V(s,p,x,\zeta_2)$, $W^n(s,p,x,\zeta_2) \to W(s,p,x,\zeta_2)$ *uniformly on compacts.*

Proof. This is an analogue of [8, Theorem 1]. Compared to our earlier work, the only new feature is that the payoffs to producer 1 are randomized. Indeed, from her perspective, the strategy of player 2 (implied through the private signal $\mu_1(\gamma)$) may be mixed. In the latter case her continuation value is unknown at decision time, depending as it is on the action of player 2. Formally, allowing for a relaxed switching control p_s^1 at date s (representing probability of a switch) the dynamic programming principle implies that in (11)

$$V^n(s,p,x,\zeta_2) = \mathbb{E}^{\gamma_s}\Bigg[\sup_{p_s^1\in[0,1]}\Big\{(1-p_s^1)(a_1p-b_1x-c_1)+p_s^1K_{\{1,1,0\}}$$
$$+\mathbb{E}^{\mu_1(s;\gamma)}\Big[p_s^1p_s^2W^{n-1}(s+1,P_{s+1},X_{s+1}^{(0,\bar{\zeta}_2)},\bar{\zeta}_2)+p_s^1(1-p_s^2)W^{n-1}(s+1,P_{s+1},X_{s+1}^{(0,\zeta_2)},\zeta_2)$$
$$+(1-p_s^1)p_s^2V^n(s+1,P_{s+1},X_{s+1}^{(1,\bar{\zeta}_2)},\bar{\zeta}_2)+(1-p_s^1)(1-p_s^2)V^n(s+1,P_{s+1},X_{s+1}^{(1,\zeta_2)},\zeta_2)\Big]\Big\}\Bigg].$$

The outer expectation is averaging over the signal μ_1 whose law is specified by the communication device γ; however the decision-maker has access to $\mu_1(t;\gamma)$ and therefore makes the switching decision p_s^1 based on the conditional strategy $(p_s^2)|\mu_1(s)$ of player 2. The inner optimization is linear in p_s^1 and therefore the optimizer must be an end point of $[0,1]$. Thus, as expected, we can continue to work with pure $\mathcal{G}^i$-adapted controls. Note, that from the perspective of an observer who has access only to $\mathcal{F}_s$, the strategy of both players appears randomized.

The rest of the proof proceeds exactly as in [8] by iterating over the control decisions of producer 1 using the strong Markov property of (P,X) and the Snell envelope characterization of optimal stopping problems. □

Proposition 4.1 shows that the solution to (10) can be represented in terms of the sequence (V^n,W^n) which correspond to optimal stopping problems defined in (11). Taking the limit $n\to\infty$ we obtain the corollary

Corollary 4.1. (V,W) *satisfy the coupled dynamic programming equation:*

(12)
$$\begin{cases} V(s,p,x,\zeta_2) = \sup\limits_{\tau\in\mathfrak{D}^1(s)} \mathbb{E}^{(1,u_2,\gamma)}\left[\sum_{t=s}^{\tau-1}(a_1P_t-b_1X_t-c_1) \; + (W(\tau,P_\tau,X_\tau,u_2(\tau))-K_{\{1,1,0\}})\right] \\ W(s,p,x,\zeta_2) = \sup\limits_{\tau\in\mathfrak{D}^1(s)} \mathbb{E}^{(0,u_2,\gamma)}\Big[V(\tau,P_\tau,X_\tau,u_2(\tau))-K_{\{1,0,1\}}\Big], \end{cases}$$

Moreover, an optimal strategy $u_1^*\in\mathcal{U}_1$ *exists.*

4.2. **Non-Zero-Sum Stopping Games.** In this section we recall existing results on two-player non-zero sum stopping games in discrete time, and finite horizon. We refer to [20, 34, 35, 37, 39] for further references.

Let $\mathcal{Z}\equiv(Z_i^{jk}(t))$, $i\in\{1,2\},j,k\in\{0,1\}$ be a octuple of bounded $(\mathcal{H}_t)$-adapted stochastic processes on a filtered probability space $(\Omega,\mathcal{H},(\mathcal{H}_t),\overline{\mathbb{P}})$. Consider a finite horizon stochastic game with player $i\in\{1,2\}$ maximizing the reward

$$\tilde{J}_i(s,\tau_1,\tau_2) \triangleq \left(\sum_{t=s}^{\tau_i\wedge\tau_{-i}-1} Z_i^{00}(t)\right) + Z_i^{10}(\tau_i)1_{\{\tau_i<\tau_{-i}\}} + Z_i^{01}(\tau_{-i})1_{\{\tau_{-i}<\tau_i\}} + Z_i^{11}(\tau_i)1_{\{\tau_i=\tau_{-i}\}}, \quad (13)$$

where the (randomized) $(\mathcal{H}_t)$-stopping times $\tau_i\le T$ are chosen by player i. In words, Z_i^{00} is the ongoing reward for staying in the game, Z_i^{10} is the reward if the player stops first; Z_i^{01} is the reward if the other player stops first and Z_i^{11} is the reward if both players stop simultaneously. Thus, continuing is associated with action '0' and stopping with action '1'.

The theory of game equilibria of (13) has been divided according to the relationships between Z^{jk}'s. The first case of the Dynkin zero-sum stopping game corresponds to $Z_1^{10} = -Z_2^{01}, Z_1^{01} = -Z_2^{10}, Z_1^{11} = -Z_2^{11}$ and was recently fully analyzed by Ekstrom and Peskir [17]. Also, the monotone cases $Z_i^{01} \le Z_i^{11} \le Z_i^{10}$ $\bar{\mathbb{P}}$-a.s. (where both players prefer to stop late) and $Z_i^{01} \ge Z_i^{11} \ge Z_i^{10}$ where both players prefer to stop first, were considered by Ohtsubo [34]. In these special cases, the assumptions on the payoff structure ensure the existence of a unique pure Markov NEP. The fundamental result of [34, 35] states that if one can find a pair of $\mathcal{H}$-adapted random processes (V_1, V_2) such that $\bar{\mathbb{E}}[\sup_{0\le t\le T} V_i(t)] < \infty$ and

$$(V_1(t), V_2(t)) \in \mathcal{E}\begin{pmatrix} (\bar{\mathbb{E}}[V_1(t+1)|\mathcal{H}_t] + Z_1^{00}(t), \bar{\mathbb{E}}[V_2(t+1)|\mathcal{H}_t] + Z_2^{00}(t)) & (Z_1^{01}(t), Z_2^{01}(t)) \\ (Z_1^{10}(t), Z_2^{10}(t)) & (Z_1^{11}(t), Z_2^{11}(t)) \end{pmatrix}, \tag{14}$$

for all $0 \le t \le T$ then (V_1, V_2) form a pair of game value functions for the stopping game $\mathcal{Z}$. The functions $V_i(t)$ can be sequentially constructed using backward recursion, starting with the terminal condition $V_i(T) = 0$, and the usual conditional expectation paradigm. This reduces computation of game values to iterative solution of one-shot 2×2 games, in complete analogy to standard discrete-time dynamic programming.

Without any assumptions on the structure of $\mathcal{Z}$ appearing in (13), the existence of a pure NEP is not guaranteed. However, as shown by [20] a two-person stopping game always admits a mixed Nash equilibrium point. This is consistent with our discussion in Section 3.3 where we have shown that $|\mathcal{E}| > 0$ for the 2-by-2 game (14) arising at each stage of the stopping game. A randomized stopping strategy profile is a pair $(\tau(p^1), \tau(p^2))$ of randomized $(\mathcal{H}_t)$-stopping times, recall (4). It can also be viewed as a sequence of one-period strategy profiles defined by $\vec{\pi}^i(t) = (1 - p^i(t), p^i(t))$.

4.3. **Correlated Equilibria in Stopping Games.** In order to have a well-defined concept of a game value for a stopping game, the problem of potentially multiple equilibria must be addressed. Following [37] we use the concept of *weak (stepwise) communication device* γ. Such γ is an $(\mathcal{H}_t)$-adapted stochastic process taking values in Δ_4, i.e. $\gamma(t)$ is a probability measure on the stopping action pair, where $\gamma_{10}(t)$ (resp. $\gamma_{01}(t)$) specifies the probability that the first (resp. second) player stops at stage t, $\gamma_{00}(t)$ is the probability that both continue and $\gamma_{11}(t)$ is the probability that both stop. Thus, conditional on the game still continuing, the stage-t expected payoff to player i is $\sum_{j,k} \gamma_{jk}(t) Z_i^{jk}(t)$. Given $(\gamma_{jk}(t))$, the randomized stopping rules are defined by

$$\begin{cases} \tau_1(\gamma) \triangleq \inf\{t : \eta_1'(t) \le \gamma_{10}(t) + \gamma_{11}(t)\}, \\ \tau_2(\gamma) \triangleq \inf\{t : \eta_2'(t) \le \gamma_{01}(t) + \gamma_{11}(t)\}, \end{cases} \qquad \eta_1'(t), \eta_2'(t) \sim Unif[0,1] \text{ i.i.d.}$$

Overall, the total expected payoff to player i is

$$\bar{\mathbb{E}}^\gamma \left[\tilde{J}_i(s, \tau_1(\gamma), \tau_2(\gamma)) \right] = \bar{\mathbb{E}} \left[\sum_{t=s}^{T-1} \sum_{j,k} \left\{ \Big(\prod_{r=0}^{t-1} \gamma_{00}(r) \Big) \gamma_{jk}(t) Z_i^{jk}(t) \right\} \right]. \tag{15}$$

The correlation is implemented by a third-party simulating the pair (η_1', η_2') and then communicating the signal $\mu_i(t;\gamma) = 1_{\{\tau_i(\gamma) \geq t\}} 1_{\{\eta_i'(t) \leq \gamma_{i(-i)}(t) + \gamma_{11}(t)\}}$ to player i. In comparison to independent randomizations giving rise to $\tau(p^1), \tau(p^2)$ above, the stopping times $\tau_i(\gamma)$ may be dependent. The communication device is active on the full horizon $[0,T]$, regardless of deviations. Consequently, given γ, the payoff functionals $\tilde{J}_1(t, \tilde{\tau}_1, \tau_2(\gamma))$ and $\tilde{J}_2(t, \tau_1(\gamma), \tilde{\tau}_2)$, as defined in (13) are well-defined for any $\tilde{\tau}_i \in \mathfrak{D}^i(t)$. This leads to

Definition 4.1. *Consider a stopping game* $(\mathcal{Z}_i^{jk})$ *with payoffs in* (13)*. A stopping strategy* $\vec{\tau}(\gamma) \triangleq (\tau_1(\gamma), \tau_2(\gamma)) \in \mathfrak{D}^1 \times \mathfrak{D}^2$ *is a correlated equilibrium of* $\mathcal{Z}$ *with communication device* γ *if for* $i = 1, 2$ *and all* $0 \leq t < T$ *we have*

$$V_i(t;\gamma,\mathcal{Z}) \triangleq \bar{\mathbb{E}}^{\gamma}[\tilde{J}_i(t, \vec{\tau}(\gamma))|\mathcal{H}_t] \geq \bar{\mathbb{E}}^{\gamma}[\tilde{J}_i(t, \tilde{\tau}_i, \tau_{-i}(\gamma))|\mathcal{H}_t], \quad \forall \tilde{\tau}_i \in \mathfrak{D}^i(t). \tag{16}$$

Observe that given a device γ leading to a correlated equilibrium, it must be that

$$V_i(t;\gamma,\mathcal{Z}) = \sup_{\tau \in \mathfrak{D}^i(t)} \bar{\mathbb{E}}^{\gamma}\Big[\left(\sum_{s=t}^{(\tau \wedge \tau_{-i}(\gamma))-1} Z_i^{00}(s) \right) + Z_i^{10}(\tau) 1_{\{\tau < \tau_{-i}\}} + Z_i^{01}(\tau_{-i}) 1_{\{\tau_{-i} < \tau\}} + Z_i^{11}(\tau) 1_{\{\tau = \tau_{-i}\}} \Big| \mathcal{H}_t \Big] \tag{17}$$

which is a standard optimal stopping problem for player i in the enlarged filtration $\mathcal{G}^i$. Let us also recall the following result (compare to (14)).

Lemma 4.2. [37, Theorem 2.3] *Consider a CEP with communication device* γ *of a stopping game* $\mathcal{Z}$*. Then for all* $t \in \{0, 1, \ldots, T-1\}$ *we have*

(18)
$$\begin{cases} \gamma_{00}(t)(\bar{\mathbb{E}}[V_1(t+1)|\mathcal{H}_t] + Z_1^{00}(t)) + \gamma_{01}(t) Z_1^{01}(t) \geq \gamma_{00}(t) Z_1^{10}(t) + \gamma_{01}(t) Z_1^{11}(t)); \\ \gamma_{00}(t)(\bar{\mathbb{E}}[V_2(t+1)|\mathcal{H}_t] + Z_2^{00}(t)) + \gamma_{10}(t) Z_2^{10}(t) \geq \gamma_{00}(t) Z_2^{01}(t) + \gamma_{10}(t) Z_2^{11}(t); \\ \gamma_{10}(t) Z_1^{10}(t) + \gamma_{11}(t) Z_1^{11}(t) \geq \gamma_{10}(t)(\bar{\mathbb{E}}[V_1(t+1)|\mathcal{H}_t] + Z_1^{00}(t)) + \gamma_{11}(t) Z_1^{01}(t); \\ \gamma_{01}(t) Z_2^{01}(t) + \gamma_{11}(t) Z_2^{11}(t) \geq \gamma_{01}(t)(\bar{\mathbb{E}}[V_2(t+1)|\mathcal{H}_t] + Z_2^{00}(t)) + \gamma_{11}(t) Z_2^{10}(t). \end{cases}$$

As shown by [37, Theorem 2.4], any finite-horizon stopping game with bounded payoffs admits a CEP; in fact outside the zero-sum and monotone cases we expect that a large number of CEPs are possible. It is convenient to think of communication device γ leading to a CEP as a measurable selector of local correlated equilibrium points in the one-shot 2×2 games. Thus, let $\Gamma : \mathbb{T} \times \Omega \times \mathbb{R}^{2\times2\times2} \to \Delta_4$ be a measurable map such that for any 2×2 game H, $\Gamma(t, \omega, H) \in \mathcal{E}_c(H)$. Then using Γ, one may construct a communication device γ by using the CEP $\Gamma(t, \omega, \mathcal{Z}(t, \omega))$ on (14) and applying backward induction. Observe that for most H's, $\Gamma(\cdot, H)$ is simply the unique NEP available, so that the selection feature is "silent", and the device is only really activated when considering the coordination game. With this perspective in mind, we call a *correlation law* Γ a communication device which is based on the same local criterion (for instance "minimize today's emissions" or "maximize today's value of player 1").

Remark 4.1. A variety of correlated decision-making is possible in sequential games [31]. Here we focus on the stepwise weak communication device which means that players and the regulator communicate before each stage; such a formulation allows the most flexibility and fits our economic description. However, in practice much weaker correlation could suffice. For instance, players can agree at date 0 to use the preferential-i correlation law which means that in any "tie-break" case, player i "wins". Once this rule is fixed, no further communication (or randomization by regulator) would be necessary. Similarly, if Γ is such that the implied strategy $\vec{\pi}_{-i}(\Gamma)|\mu_i(\gamma)$ of the other player is always pure, then a public randomization (representing external circumstances not part of the model) is sufficient at each step and no private signals are needed.

4.4. **Recursive Construction.** We return to the CO_2-duopoly market setup. The emission schedules of the two agents are interpreted as a sequence of regime-changes. Thus, the single-stopping game in the previous section is viewed as the sub-game for making the next regime-switch. The stopping game in Section 4.2 is accordingly denoted as a $(1,1)$-fold switching game and we now will consider (n,m)-fold switching games with game value functions $V^{n,m}$. These games have a restricted set of possible production strategies; namely the total number of regime switches over the game horizon is bounded by n and m, respectively. Using the Markov property of the game state and actions it is not surprising that these various switching games are related to each other.

In terms of the notation of Section 4.2, we identify the running profit with $Z_i^{00}(t) = (a_i P_t - b_i X_t - c_i) u_i(t)$ and the other Z_i^{jk}'s with various game continuation-values. The generic filtration $(\mathcal{H}_t)$ of the previous section will be replaced with the filtration $(\mathcal{F}_t)$ generated by (P_t, X_t), and we will make explicit the resulting dependence of game values on current prices and initial regime $\vec{\zeta}$.

For the remainder of the section, we make a standing assumption that a communication device γ is chosen and fixed. Let us fix an initial state $P_s = p, X_s = x$ and initial production regime $\vec{\zeta} = (\zeta_1, \zeta_2)$. Define a double cascade of stopping games indexed by n and m via

$$V_i^{n,m}(s,p,x,\vec{\zeta}) \triangleq V_i(s;\gamma,\tilde{\mathcal{Z}}^{n,m}(\vec{\zeta})), \qquad n,m \geq 1 \tag{19}$$

which uses the notation of (16) based on the recursive payoff structure

$$\begin{cases} (\tilde{Z}^{n,m})_i^{00}(t,\vec{\zeta}) = (a_i P_t - b_i X_t - c_i)\zeta_i; \\ (\tilde{Z}^{n,m})_i^{01}(t,\vec{\zeta}) = V_i^{n,m-1}(t,P_t,X_t,\zeta_1,\bar{\zeta}_2) - 1_{\{i=2\}} K_{\{2,\zeta_2,\bar{\zeta}_2\}}; \\ (\tilde{Z}^{n,m})_i^{10}(t,\vec{\zeta}) = V_i^{n-1,m}(t,P_t,X_t,\bar{\zeta}_1,\zeta_2) - 1_{\{i=1\}} K_{\{1,\zeta_1,\bar{\zeta}_1\}}; \\ (\tilde{Z}^{n,m})_i^{11}(t,\vec{\zeta}) = V_i^{n-1,m-1}(t,P_t,X_t,\bar{\zeta}_1,\bar{\zeta}_2) - K_{\{i,\zeta_i,\bar{\zeta}_i\}}. \end{cases} \tag{20}$$

The boundary cases are first

$$V_i^{0,0}(s,p,x,\vec{\zeta}) \triangleq \mathbb{E}^{\vec{\zeta}}\left[\sum_{t=s}^{T-1} (a_i P_t - b_i X_t - c_i)\zeta_i\right];$$

next, $V_1^{n,0}(s,p,x,\vec{\zeta})$ and $V_2^{0,m}(s,p,x,\vec{\zeta})$ are identified with the single-player optimization problems as in (11) (keeping the emission regime of the other player fixed at ζ_{-i}). Finally, we take

$$V_2^{n,0}(s,p,x,\vec{\zeta}) = \mathbb{E}^{(u_1^{n,*},\zeta_2,\gamma)}\left[\sum_{t=s}^{T-1}(a_2 P_t - b_2 X_t - c_2)\zeta_2\right]$$

where $u_1^{n,*}$ is an optimal control for the problem defining $V_1^{n,0}$, and similarly for $V_1^{0,m}(s,p,x,\vec{\zeta})$.

4.5. **Switching Game Equilibrium as Sequential Stopping Game Equilibrium.** We now proceed to *glue* the sequential stopping games of $V_i^{n,m}$ and re-interpret the latter as value functions of a switching game. Recall that for $n \geq 0$, $\mathcal{U}_i^n \subset \mathcal{U}_i$ is the set of all production strategies for player i with at most n switches. Consider the restricted repeated game with payoffs (6) where we require $u_1 \in \mathcal{U}_1^n$ and $u_2 \in \mathcal{U}_2^m$, so that the first producer may change her production regime at most n times, and the second producer at most m times.

Our first task is to obtain a switching-game CEP that matches the definition of $V^{n,m}$. To do so we pick a correlation law Γ; Γ gives rise to a CEP of any stopping game, in particular it leads to well-defined game values $V_i^{n,m}$ in (19). We now construct a communication device $\gamma^{n,m}$ for the (n,m)-switching game. Let $k_i(t)$ be the number of production switches used by player i by stage t. We define $\gamma^{n,m}(t)$ at stage t by applying the device $\Gamma(\tilde{\mathcal{Z}}^{n-k_1(t),m-k_2(t)}(t,\vec{u}(t))$ defined in terms of (20) and the latest regime $\vec{u}(t)$. Note that the overall $\gamma^{n,m}$ is no longer Markovian since it has memory of the number of switches made by each player, which is necessary in the constrained game. The above construction is well-defined for all paths of $(P,X,\vec{u})$, even outside equilibrium.

Using $\gamma^{n,m}$ we proceed to construct switching controls $u_i^{n,m}$ for the (n,m)-switching game. To simplify notation we write $\underline{\tau}^{n,m} = \tau_1(\gamma^{n,m}) \wedge \tau_2(\gamma^{n,m})$ which is interpreted as the equilibrium first stopping time for the game defined by (19) under the correlation law Γ. Given the starting production regime $\vec{\zeta} = (\zeta_1,\zeta_2)$, let us define the switching controls $u_i^{n,m}(s)$ for this game by

$$u_1^{n,m}(s) = \zeta_1 \qquad \text{for} \quad s < \underline{\tau}^{n,m};$$

$$u_1^{n,m}(s) = \begin{cases} \bar{\zeta}_1 \text{ for} & \underline{\tau}^{n,m} \leq s < \underline{\tau}^{n-1,m} \qquad \text{when } \tau_1^{n,m} < \tau_2^{n,m}; \\ \zeta_1 \text{ for} & \underline{\tau}^{n,m} \leq s < \underline{\tau}^{n,m-1} \qquad \text{when } \tau_2^{n,m} < \tau_1^{n,m}; \\ \bar{\zeta}_1 \text{ for} & \underline{\tau}^{n,m} \leq s < \underline{\tau}^{n-1,m-1} \qquad \text{when } \tau_1^{n,m} = \tau_2^{n,m}, \end{cases} \tag{21}$$

... and so on,

and similarly for $u_2^{n,m}(t)$. In words, $u_i^{n,m}$ keeps track of the production regime of the i-th agent following the decision rules defined sequentially by descending through the family of the $V^{n,m}$-stopping subgames (one stopping game at a time). Then by definition of (19) we have $u_1^{n,m} \in \mathcal{U}_1^n$ and $u_2^{n,m} \in \mathcal{U}_2^m$. It can also be seen through an easy induction argument that

$$V_i^{n,m}(s,p,x,\vec{\zeta}) = V_i(s,p,x,\vec{\zeta};\vec{u}^{n,m}), \tag{22}$$

so that the switching control $\vec{u}^{n,m}$ of (21) allows to achieve the game values $V^{n,m}$ defined recursively in (19). Moreover, the next theorem shows that the pair $(u_1^{n,m}, u_2^{n,m})$ is in fact a correlated equilibrium (using correlation device $\gamma^{n,m}$) for the game (6) over the control set $\mathcal{U}_1^n \times \mathcal{U}_2^m$.

Theorem 4.3. *For all $n > 0$ and $u_1 \in \mathcal{U}_1^n$ we have $V_1(t, \cdot; u_1^{n,m}, u_2^{n,m}) \geq V_1(t, \cdot; u_1, u_2^{n,m})$. Similarly for all $m > 0$ and $u_2 \in \mathcal{U}_2^m$ we have $V_2(t, \cdot; u_1^{n,m}, u_2^{n,m}) \geq V_2(t, \cdot; u_1^{n,m}, u_2)$.*

Proof. The idea of the proof is to make use of the Markov structure of our problem and apply induction. The other key tool is that given $\gamma^{n,m}$, we can look at one player at a time which essentially reduces to a single-player problem studied before, see (17).

Due to symmetry, it suffices to prove the result for player 1. When $m = 0$ the other player cannot act, the game becomes trivial and Theorem 4.3 is just a re-statement of Proposition 4.1. Conversely, when $n = 0$, the first player cannot act and there is nothing to prove. Using induction we assume that the theorem has been shown for the pairs $(n-1, m-1)$, $(n, m-1)$ and $(n-1, m)$; let us show it for the case (n, m). Given an arbitrary $u_1 \in \mathcal{U}_1^n$, write it as $u_1 = (\tau^1, \hat{u}_1)$ where $\hat{u}_1 \in \mathcal{U}_1^{n-1}$ denotes the remainder of u_1 after the first switch time τ^1. Let $\tau^{2,*} \equiv \tau^2(\gamma^{n,m})$ be the first switch for the second player dictated through $\gamma^{n,m}$. Define $\underline{\tau} = \tau^1 \wedge \tau^{2,*}$. Also for notational convenience we omit all the arguments of $V^{n,m}$ except for the time variable. Then the strong Markov property of (P, X) and the way $u_2^{n,m}(t)$ was constructed show that

$$\mathbb{E}^{(u_1, u_2^{n,m}, \gamma^{n,m})}\left[\sum_{s=\underline{\tau}}^{T-1} (a_1 P_s - b_1 X_s - c_1)\hat{u}_1(s)\right] = \mathbb{E}^{(u_1, u_2^{n,m}, \gamma^{n,m})}\Big[V_1(\underline{\tau}; \hat{u}_1, u_2^{n-1,m}) 1_{\{\tau^1 < \tau^{2,*}\}}$$
$$+ V_1(\underline{\tau}; \hat{u}_1, u_2^{n,m-1}) 1_{\{\tau^1 > \tau^{2,*}\}} + V_1(\underline{\tau}; \hat{u}_1, u_2^{n-1,m-1}) 1_{\{\tau^1 = \tau^{2,*}\}}\Big].$$

Conditioning on τ^1 and $\tau^{2,*}$ we therefore have

$$\begin{aligned}
V_1(t; u_1, u_2^{n,m}) = \mathbb{E}^{(u_1, u_2^{n,m}, \gamma^{n,m})}\Big[&\left(\sum_{s=t}^{\underline{\tau}-1} (a_1 P_s - b_1 X_s - c_1) u_1(t)\right) + \left(\sum_{s=\tau^1}^{T-1} (a_1 P_s - b_1 X_s - c_1)\hat{u}_1(s)\right) 1_{\{\tau^1 < \tau^{2,*}\}} \\
&+ \Big(\sum_{s=\tau^{2,*}}^{T-1} (a_1 P_s - b_1 X_s - c_1)\hat{u}_1(s)\Big) 1_{\{\tau^1 > \tau^{2,*}\}} + \left(\sum_{s=\tau^1}^{T-1} (a_1 P_s - b_1 X_s - c_1)\hat{u}_1(s)\right) 1_{\{\tau^1 = \tau^{2,*}\}}\Big] \\
= \mathbb{E}^{(u_1, u_2^{n,m}, \gamma^{n,m})}\Big[&\left(\sum_{s=t}^{\underline{\tau}-1} (a_1 P_s - b_1 X_s - c_1) u_1(t)\right) + V_1(\underline{\tau}; \hat{u}_1, u_2^{n-1,m}) 1_{\{\tau^1 < \tau^{2,*}\}} \\
&+ V_1(\underline{\tau}; \hat{u}_1, u_2^{n,m-1}) 1_{\{\tau^1 > \tau^{2,*}\}} + V_1(\underline{\tau}; \hat{u}_1, u_2^{n-1,m-1}) 1_{\{\tau^1 = \tau^{2,*}\}}\Big]
\end{aligned}$$

by induction hypothesis we have the inequality

$$\begin{aligned}
&\leq \mathbb{E}^{(u_1, u_2^{n,m}, \gamma^{n,m})}\Big[\left(\sum_{s=t}^{\underline{\tau}-1} (a_1 P_s - b_1 X_s - c_1) u_1(t)\right) + V_1(\tau^1; u_1^{n-1,m}, u_2^{n-1,m}) 1_{\{\tau^1 < \tau^{2,*}\}} \\
&\qquad + V_1(\tau^{2,*}; u_1^{n,m-1}, u_2^{n,m-1}) 1_{\{\tau^1 > \tau^{2,*}\}} + V_1(\tau^1; u_1^{n-1,m-1}, u_2^{n-1,m-1}) 1_{\{\tau^1 = \tau^{2,*}\}}\Big] \\
&\leq \sup_{\tau^1 \in \mathfrak{D}^1(t)} \mathbb{E}^{(u_1(t), u_2^{n,m}, \gamma^{n,m})}\Big[\left(\sum_{s=t}^{\underline{\tau}-1} (a_1 P_s - b_1 X_s - c_1) u_1(t)\right) + V_1(\tau^1; \vec{u}^{n-1,m}) 1_{\{\tau^1 < \tau^{2,*}\}} \\
&\qquad + V_1(\tau^{2,*}; \vec{u}^{n,m-1}) 1_{\{\tau^1 > \tau^{2,*}\}} + V_1(\tau^1; \vec{u}^{n-1,m-1}) 1_{\{\tau^1 = \tau^{2,*}\}}\Big] \\
&= V_1(t; \vec{u}_1^{n,m}),
\end{aligned}$$

where the last line uses the relationship (22), the construction of the stopping game defining $V^{n,m}$ in (20), and property (17). □

The above construction leads to the key result of this section that characterizes CEP of switching games, establishes their existence, and gives a recursive formula for the resulting game value functions.

Theorem 4.4. *Fix a correlation law* Γ. *Then* Γ *gives rise to a CEP of the switching game* (6). *Moreover, the corresponding value functions* $V_i(t, p, x, \vec{\zeta}; \Gamma)$ *solve*

$$(23)\quad (V_1(t,\vec{\zeta}), V_2(t,\vec{\zeta})) = Val_{\Gamma(t)} \begin{pmatrix} (Z_1(t,\zeta_1,\zeta_2), Z_2(t,\zeta_1,\zeta_2)) & (Z_1(t,\zeta_1,\bar{\zeta}_2), Z_2(t,\zeta_1,\bar{\zeta}_2) - K_{2,\zeta_2,\bar{\zeta}_2}) \\ (Z_1(t,\bar{\zeta}_1,\zeta_2) - K_{1,\zeta_1,\bar{\zeta}_1}, Z_2(t,\bar{\zeta}_1,\zeta_2)) & (Z_1(t,\bar{\zeta}_1,\bar{\zeta}_2) - K_{1,\zeta_1,\bar{\zeta}_1}, Z_2(t,\bar{\zeta}_1,\bar{\zeta}_2) - K_{2,\zeta_2,\bar{\zeta}_2}) \end{pmatrix}$$

where $Z_i(t,\vec{\zeta}) = \mathbb{E}^{\vec{\zeta}}[V_i(t+1,\vec{\zeta})|\mathcal{F}_t] + (a_i P_t - b_i X_t - c_i)\zeta_i$. *Moreover the equilibrium controls can be taken as* $u_i^* \equiv u_i^{T,T}$, *as defined in* (21).

Proof. We wish to take $n, m \to \infty$ in Theorem 4.3. Because for $n > m$, $\mathcal{U}^m \subseteq \mathcal{U}^n$, it follows that for a fixed m, $V_1^{n,m}$ is increasing in n (and for a fixed n, $V_2^{n,m}$ is increasing in m). For our finite horizon discrete-time game, at most T regime switches are possible for each player. Therefore $u_i^* \in \mathcal{U}_i^T$ and it follows that $V_i^{n,n} \equiv V_i$ for all $n > T$. In particular, a switching CEP based on Γ results by using $\gamma^{T,T}$.

Moreover, at equilibrium at most one switch is made at any given stage. Therefore, suppose it is optimal to switch at stage t from $\vec{\zeta}$ to $\vec{u}$; then already starting at regime $\vec{u}$ at t (and same state variables) it is optimal to make no changes, so that $V_i(t,\vec{u}) = \mathbb{E}^{\vec{u}}[V_i(t+1,\vec{u})|\mathcal{F}_t] + (a_i P_t - b_i X_t - c_i)u_i$ for that scenario. Combining these facts with the form of (14) and dropping the constraints on the number of switches, we may express all payoffs in terms of next-stage game values. The recursion (23) is now obtained by making this substitution in (14). □

5. Numerical Implementation

Theorem 4.4 shows that a game value and equilibrium strategy profile can be obtained recursively by solving the 1-period 2-by-2 games in (23). The payoffs of those games are given iteratively in terms of *conditional expectations* of next-stage game values. Therefore, a numerical implementation hinges on accurate evaluation of these expectations. Since our state-space in (P, X) is continuous, it is impossible to make this computation exactly. Instead, we must resort to an approximation; below we present two possible approaches.

5.1. **Markov Chain Approximation Algorithm.** Our model would be simplified if the continuous state space of (P, X) is discretized. Let $(\tilde{P}, \tilde{X})$ be an approximating discrete-state process with $(\tilde{P}_t, \tilde{X}_t)$ living on a finite subset $D_t \subset \mathbb{R}_+^2$. If the pair $(\tilde{P}, \tilde{X})$ is furthermore chosen to be again Markov, this is known as the Markov Chain Approximation (MCA) method of [27]. With such $(\tilde{P}, \tilde{X})$, a conditional expectation $\mathbb{E}[f(\tilde{P}_{t+1}, \tilde{X}_{t+1})|\tilde{P}_t, \tilde{X}_t]$ for any measurable function f is just a weighted sum based on the transition probability matrix of $(\tilde{P}, \tilde{X})$. The backward recursion in (23) for $\tilde{V}_i$, the corresponding approximation of V_i, can now be implemented directly for each stage t and each possible state of $(\tilde{P}_t, \tilde{X}_t) \in D_t$. A well-known procedure constructs $(\tilde{P}, \tilde{X})$ by taking D_t to be a 2-dimensional regular grid or lattice and allowing state transitions only between neighboring grid points. Moreover, the transition probabilities of $(\tilde{P}, \tilde{X})$ are chosen so that to have *local consistency* in the first two moments with the 1-step transition densities of (P, X); see [27, Chapter 5].

To use this approach in our model, one must take into account the price impact. Therefore, we construct four approximations $(\tilde{P}, \tilde{X}^{\vec{\zeta}})$ indexed by the possible joint production regimes $\vec{\zeta} \in \{0,1\}^2$ that induce different local dynamics of $\tilde{X}^{\vec{\zeta}}$, see (2). In other words, our effective state variables are $(\tilde{P}, \tilde{X}, \vec{\zeta})$. For every possible combination $(t, \tilde{p}, \tilde{x}, \vec{\zeta}) \in \mathbb{T} \times D_t \times \{0,1\}^2$ the relation (23) is then solved through backward recursion. A generic convergence proof (as the grid spacing tends to zero) of this procedure for finite-horizon non-zero-sum stochastic games was obtained in [26]. Note that in our model the controls $\vec{u}(t)$ are discrete and finite-valued and therefore all the compactness conditions in [26] for the control space are automatically satisfied.

5.2. **Least Squares Monte Carlo Approach.** Like classical dynamic programming, the MCA method above suffers from the curse of dimensionality. Indeed, the size of the approximating grid grows exponentially in the dimension of the state variables. In our basic model (P, X) are two-dimensional; however realistic implementations are likely to take multi-dimensional factor models for P and (possibly) X. Thus, it is helpful to seek a more robust algorithm.

A seminal idea due to [9, 16, 29] is to use a cross-sectional regression combined with a Monte Carlo simulation to make the recursive computations of the relevant conditional expectations. The key step is a global approximation of the maps $(t, p, x, \vec{\zeta}) \mapsto V_i(t, p, x, \vec{\zeta})$ and equilibrium one-step strategies $(t, p, x, \vec{\zeta}) \mapsto \vec{u}(t, p, x, \vec{\zeta})$ (based on a fixed correlation law Γ) via a random sample of (P_t, X_t). The construction is iterative and backward in time.

Suppose that the current date is t and we already know all the approximations $v_i(s, p, x, \vec{\zeta}) \simeq V_i(s, p, x, \vec{\zeta})$ for $s > t$ and the corresponding equilibrium strategy profiles. Then given a collection of initial points (p^n_t, x^n_t), for $n = 1, \ldots, N$, and an arbitrary starting emission regime $\vec{\zeta} = \vec{u}^n(t)$ we simulate the future *cashflows* on $[t+1, T]$ for each scenario n by iteratively computing the equilibrium actions $u^n_i(s)$ of each player for $s = t+1, \ldots, T$ based on the estimated future game values $v_i(s, p^n_s, x^n_s, \cdot)$ and the chosen communication device Γ. If Γ leads to randomized strategies, such a randomization is naturally implemented as part of this simulation. Given a strategy profile $\vec{u}^n(s)$, we update (p^n_{s+1}, x^n_{s+1}) by an independent draw from the conditional law $\mathbb{P}^{\vec{u}^n(s)}$. Eventually the simulation procedure reaches terminal date T, whereupon we have for each path n a realized equilibrium cashflow pair $\vartheta^n_i(t+1, \vec{\zeta})$. Each ϑ^n_i represents an empirical draw from $V_i(t+1, P_{t+1}, X^{\vec{\zeta}}_{t+1}, \vec{\zeta})$ conditional on $P_t = p^n_t, X_t = x^n_t$. We now perform a cross-sectional regression of $(\vartheta^n_i(t+1, \vec{\zeta}))^N_{n=1}$ against $(p^n_t, x^n_t)^N_{n=1}$ to compute the continuation values

$$\hat{v}_i(t, p^n_t, x^n_t, \vec{\zeta}) \simeq \mathbb{E}^{\vec{\zeta}} \left[V_i(t+1, P_{t+1}, X_{t+1}, \vec{\zeta}) | P_t = p^n_t, X_t = x^n_t \right].$$

Finally, using $\hat{v}_i$ together with the current payoffs and the correlation law Γ we solve for the equilibrium game values $v_i(t, p^n_t, x^n_t, \vec{u})$ for each production regime $\vec{u}$, taking into account the switching costs $K_{\{i,u_i,\zeta_i\}}$. The computed game equilibrium also provides the map $(t, p^n_t, x^n_t, \vec{u}) \mapsto \vec{u}^*(t)$ for the equilibrium strategies. The regression results allow to further extend this to arbitrary initial condition $(t, p, x, \vec{u})$.

The initial collection (p^n_t, x^n_t) is obtained by simulation. Since, X is affected by the price impact of $\vec{u}$, to implement this simulation we need to select some anterior auxiliary strategy profile $\vec{u}^0$. While in theory $\vec{u}^0$ can be arbitrary, in practice it should be close to the equilibrium $\vec{u}^*$. Indeed, the collection $(v_i(t, p^n_t, x^n_t, \vec{\zeta}))_{n=1}^N$ is supposed to approximate $V_i(t, P_t, X^*_t, \vec{u}(t))$ where X^*_t is the equilibrium CO_2 allowance price. Because v_i's are computed by employing regression, the resulting approximation cannot be uniformly good on $\mathbb{R}^2_+$. From the point of view of accurate solutions, it needs to be good around the region of interest for X^*_t. Thus, we need most of the x^n_t's to be in the same (*a priori* unknown) neighborhood. To overcome this difficulty, as the algorithm works back through time, the future paths (p^n_s, x^n_s), $s > t$ are re-computed using the now-available (approximately) equilibrium strategies $u^*(s)$. To further mitigate the problem, we iteratively re-do the whole simulation and subsequent backward recursion a few times (in practice three iterations suffice), using the computed $\vec{u}^*$ from one iteration as the anterior $\vec{u}^0$ in the next one.

Selection of basis functions should reflect the expected shape of $(p, x) \mapsto V_i(t, p, x, \vec{\zeta})$. A typical choice is to use low-degree polynomial basis functions, such as p, p^2, x, x^2, etc. In practice, $r = 5-7$ basis functions and $N = 32000 - 50000$ paths suffice. A large degree of customization, such as time-varying bases, constrained least-squares regression, variance reduction methods, etc., is possible to speed up the computations. Also, note that as for standard optimal stopping [16], it is not necessary to actually store the values $v_i(t, \cdot)$, but only the realized cashflows $\vartheta^n_i(t, \vec{\zeta})$ and the regression coefficients $\vec{\alpha}_i(t, \vec{\zeta})$. Below, we summarize the above scheme in pseudo-code in Algorithm 1. It calls as a sub-routine Algorithm 2 that carries out the forward simulations of ϑ^n_i.

The cost of simulations in Algorithm 1 is roughly $N \cdot (1 + 2 + \ldots + T) = \mathcal{O}(N \cdot T^2)$ which consists of re-simulating N paths on $[t, T]$ as t goes from $T - 1$ to zero (see Algorithm 2). The cost of doing regression against r basis functions on each path and for each stage is $\mathcal{O}(N \cdot T \cdot r^3)$ and the cost of computing continuation values is $\mathcal{O}(N \cdot T^2 \cdot r)$. The memory requirements are $\mathcal{O}(N \cdot T)$ which comes from storing all the simulation paths during the backward recursion.

Algorithm 1 Computing Correlated Equilibrium Game Values

input: $N > 0$ (number of paths); $B_\ell(p, x)$, $\ell = 1, \dots, r$ (r regression basis functions)
input: Correlation law Γ
Select anterior strategy profile $\vec{u}^0$
for each regime $\vec{\zeta} \in \{0,1\}^2$ **do**
Set $p_0^n = p_0$, $x_0^{\vec{\zeta},n} = x_0$ for $n = 1, \dots$,N
Simulate N independent paths $(p_t^n, x_t^{n,\vec{\zeta}})_{n=1}^N$ under $\mathbb{P}^{\vec{u}^0}$ using Algorithm 2
end for
Initialize $\vartheta_i^n(T, \vec{\zeta}) \leftarrow 0$, $n = 1, \dots, N$
for $t = (T-1), \dots, 1, 0$ **do**
for each regime $\vec{\zeta}$ **do**
Evaluate $B_\ell(p_t^n, x_t^{\vec{\zeta},n})$ for $\ell = 1, \dots, r$ and $n = 1, \dots, N$
Regress

$$\vec{\alpha}_i(t, \vec{\zeta}) \leftarrow \arg\min_{\vec{\alpha} \in \mathbb{R}^r} \sum_{n=1}^N \Big| \vartheta_i^n(t+1, \vec{\zeta}) - \sum_{\ell=1}^r \alpha^\ell B_\ell(p_t^n, x_t^{\vec{\zeta},n}) \Big|^2$$

end for
for each current regime $\vec{u}$ **do**
for each $\vec{\zeta} \in \{0,1\}^2$, and each $n = 1, \dots, N$ **do**
// Compute the predicted continuation value for each player from taking action $\vec{\zeta}$
// Note that evaluate on the $x^{\vec{u},n}$-paths using $\vec{\alpha}(t, \vec{\zeta})$-coefficients
Set $\hat{q}_i^n(t, \vec{u}, \vec{\zeta}) \leftarrow \sum_{\ell=1}^r \alpha_i^\ell(t, \vec{\zeta}) B_\ell(p_t^n, x_t^{\vec{u},n})$.
// Add the switching costs and current payoff (based on $\vec{\zeta}$)
$\hat{q}_i^n(t, \vec{u}, \vec{\zeta}) \leftarrow \hat{q}_i^n(t, \vec{u}, \vec{\zeta}) - K_{\{i, u_i, \zeta_i\}} + (a_i p_t^n - b_i x_t^{\vec{u},n} - c_i)\zeta_i$
end for
for each path $n = 1, \dots, N$ **do**
Compute the stage-t game values based on $\hat{q}_\cdot^n(t, \vec{u}, \cdot)$ and Γ, see (23)
Obtain the equilibrium policy $\vec{u}^{n,*}(t, \vec{u})$
Using Algorithm 2 update $\vartheta_i^n(t, \vec{u})$ using $\vec{u}^{n,*}(t, \vec{u})$ and $\vec{\alpha}_i(t, \vec{\zeta})$
end for
end for
end for
return $V_i(0, p_0, x_0, \vec{\zeta}) \simeq \frac{1}{N} \sum_{n=1}^N \vartheta_i^n(0, \vec{\zeta})$
return Regression coefficients $\vec{\alpha}_i(t, \vec{\zeta})$ summarizing equilibrium strategies

5.3. **Numerical Examples.** In this section we illustrate our analysis with a numerical case-study. The selected model parameters are listed in Table 1. The example represents emission scheduling of two producers over one calendar year; all the parameters of (P, X) are in annualized units and we use $T' = 26$ bi-weekly periods to model the actual scheduling flexibility. Note that the electricity price P_t is more volatile than the CO_2 allowance price X_t; also the mean-reversion strength κ_X is quite

Algorithm 2 Simulating one realized cashflow path ϑ

input: Basis functions $B_\ell(p,x)$, $\ell = 1,\dots,r$
input: Regression coefficients $\vec{\alpha}_i(t,\vec{\zeta})$; correlation law Γ
input: Initial condition $(p_0, x_0, \vec{u}_0)$; horizon T
  Initialize $\vartheta_i(0) \leftarrow 0$ // Realized cashflows
  for $t = 0,\dots,T-1$ **do**
    for each $\vec{\zeta} \in \{0,1\}^2$ **do**
      // Evaluate the predicted continuation values from taking action $\vec{\zeta}$
      Set $\hat{q}_i(t,\vec{\zeta}) \leftarrow \sum_{\ell=1}^{r} \alpha_i^\ell(t,\vec{\zeta}) B_\ell(p_t,x_t) - K_{\{i,(u_t)_i,\zeta_i\}} + (a_i p_t - b_i x_t - c_i)\zeta_i$
    end for
    Compute the stage-t game values based on $\hat{q}_\cdot(t,\cdot)$ and Γ, see (23)
    Obtain the correlated equilibrium strategy $\tilde{\mathbf{u}}_t$.
    if $\tilde{\mathbf{u}}_t$ is mixed **then**
      Perform randomization to obtain the realized action pair $\vec{u}_{t+1}$
    else
      Set $\vec{u}_{t+1} \leftarrow \vec{\mathbf{u}}_t$ // $\vec{\mathbf{u}}_t$ is pure
    end if
    Update $\vartheta_i(t+1) \leftarrow \vartheta_i(t) - K_{\{i,(u_t)_i,(u_{t+1})_i\}} + (a_i p_t - b_i x_t - c_i)(u_{t+1})_i, \quad i = 1,2$
    Update (p_{t+1}, x_{t+1}) through an independent draw from the law $\mathbb{P}^{\vec{u}_{t+1}}(\cdot|p_t,x_t)$
  end for
  return Simulated cumulative cashflow pair $\vartheta_i(T)$.

large, implying a significant price impact. For the latter we take $f(\zeta_1,\zeta_2) = \log(12+8\zeta_1+4\zeta_2)$ in (2), so that the mean-reversion level of $\log X$ is linear in the production regimes of producers 1 and 2, with producer 1 having more influence due to emitting twice as much carbon, $b_1 = 2b_2 \Rightarrow g_1 = 2g_2$. The stylized production/emission parameters are supposed to represent a dirty "coal" producer 1 who has low input costs but needs lots of allowances and a clean "natural gas" producer 2 who has high fixed costs but small sensitivity to allowance prices (and can generate twice as much electricity). Observe that if both producers emit simultaneously for a long period of time, then we expect $P_t \sim \bar{P} = 45, X_t \sim f(1,1) = 24$ meaning that both producers will be losing money. Therefore, extended joint emissions are not *sustainable*.

We proceed to look at the duopoly setting where the two producers above compete against each other. A large variety of CEP are possible in our model; for the sake of illustration we consider the four choices of Utilitarian, Egalitarian, Preferential 1 and Preferential 2 correlation laws, see Section 2.3. Table 2 shows the game values corresponding to these different correlation laws. These values were obtained by running Algorithm 1 of Section 5.2 using $N = 40000$ paths, and the basis functions $\{1, p, x, x^2, (2p-x-80)_+, (p-2x-10)_+\}$.

We find that the correlation law modifies the expected profit of the producers by $3\% - 5\%$. As expected, individual producer values are maximized by the preferential equilibria that always

κ_X	3	σ_X	0.25
κ_P	2	σ_P	0.4
T	1	ρ	0.6
$\bar{P}$	45	$\bar{X}$	12
P_0	45	X_0	15

Producer 1		Producer 2	
a_1	1	a_2	2
b_1	2	b_2	1
c_1	10	c_2	80
g_1	8	g_2	4
K_1	0.2	K_2	0.2

TABLE 1. Model Parameters for the Examples in Section 5.3.

Correlation Law	$V_1(0, P_0, X_0)$	$V_2(0, P_0, X_0)$
Utilitarian	5.30	4.14
Egalitarian	5.33	4.20
Preferential 1	5.39	4.11
Preferential 2	5.02	4.24

TABLE 2. Comparison of equilibrium game values for different correlation laws Γ. Standard errors of the Monte Carlo scheme are about 1%. Parameters are as given in Table 1.

favor the respective player. However, counterintuitively, the egalitarian CEP produces larger game values to both producers than the utilitarian CEP. This occurs because the correlation law is applied *stage-wise* and therefore optimizes a local criterion; there is no guarantee that the corresponding global criterion is respected. A similar phenomenon was observed in [37, Section 5.4].

To illustrate the equilibrium strategy profiles, Figure 1 shows the empirical regions in the (P, X)-space corresponding to different equilibrium strategies at a fixed date $t = 7$ (i.e. about three months into the year) using the Preferential-1 correlation law that always favors producer 1. As expected, when the current P&L of both producers is strongly negative (upper-left corner), the equilibrium action is $\vec{u}^*(t) = (0, 0)$; when it is strongly positive (large P_t) the equilibrium is to generate electricity $\vec{u}^*(t) = (1, 1)$. Because of the differing carbon-efficiencies of the producers, there are also large regions where exactly one producer can generate profit (e.g. around $\{P_t \in [40, 45], X_t^* \in [10, 12]\}$ only producer 2 is profitable). However at the border regions, the price impact and competition create new effects. In Figure 1, we observe the emergence of a local anti-coordination game around $\{(P_t, X_t^*) = (50, 15)\}$, and a competitive game around $\{(P_t, X_t^*) = (50, 12)\}$. We do not have any tools to analytically check whether a particular type of game may emerge locally; thus the competitive game region in Figure 1 could be either a true phenomenon or an aberration due to numerical errors (e.g. poor regression fit in that region). Note that most simulated equilibrium paths for X_t^* stay above $x = 13$ so the competitive game scenario at $t = 7$ is very unlikely to be realized (i.e. very few paths hit that region).

To better illustrate the optimal strategy over time, Figure 2 shows a sample path of the equilibrium (X_t^*) for one simulated ω, obtained using Algorithm 2. Analogously to single-player problems, the CO_2 allowance price undergoes hysteresis cycles [13]. Thus, when (X_t^*) is small, production becomes profitable. This leads to increased emissions and X_t^* rises through the price impact mechanism. In turn, this rise in emission costs eventually curtails production and X_t^* starts to fall. The presence of switching costs K_i lowers the scheduling flexibility of the producers and further amplifies this cycle through inertia. The pathwise cycle is of course also strongly influenced by the stochastic shocks in (P, X).

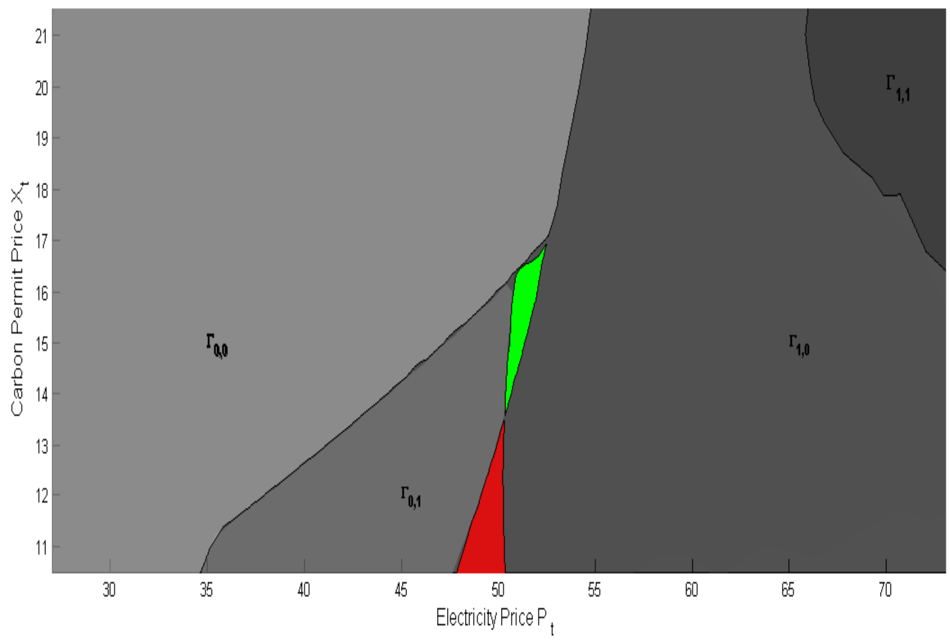

FIGURE 1. Equilibrium game strategy $u^*(t)$ as a function of (P_t, X_t^*) for $t = 7$. Here $\vec{\zeta} = (0, 0)$. The green region denotes the anti-coordination game-type where the Preferential-1 correlation law is used, and the red region denotes the competitive game-type where the unique mixed NEP is chosen.

6. Conclusion

In this paper we studied a new type of stochastic games which were motivated by dynamic emission schedules of energy producers under cap-and-trade schemes. Because multiple game equilibria can emerge, we explored various correlated equilibria. It is an interesting economic policy question which equilibrium is likely/desirable to be implemented and how the regulator can steer market participants towards that choice. The related issue of unwanted producer collusion is also highly relevant. Charging for emission externalities is supposed to promote cleaner energy generation and partially drive out "dirty" producers. It would be an interesting exercise to study how much these effects are amplified by the competitive equilibrium and especially equilibrium selection. It can be imagined that under some equilibria dirty producers could be entirely *blockaded* out of the electricity market.

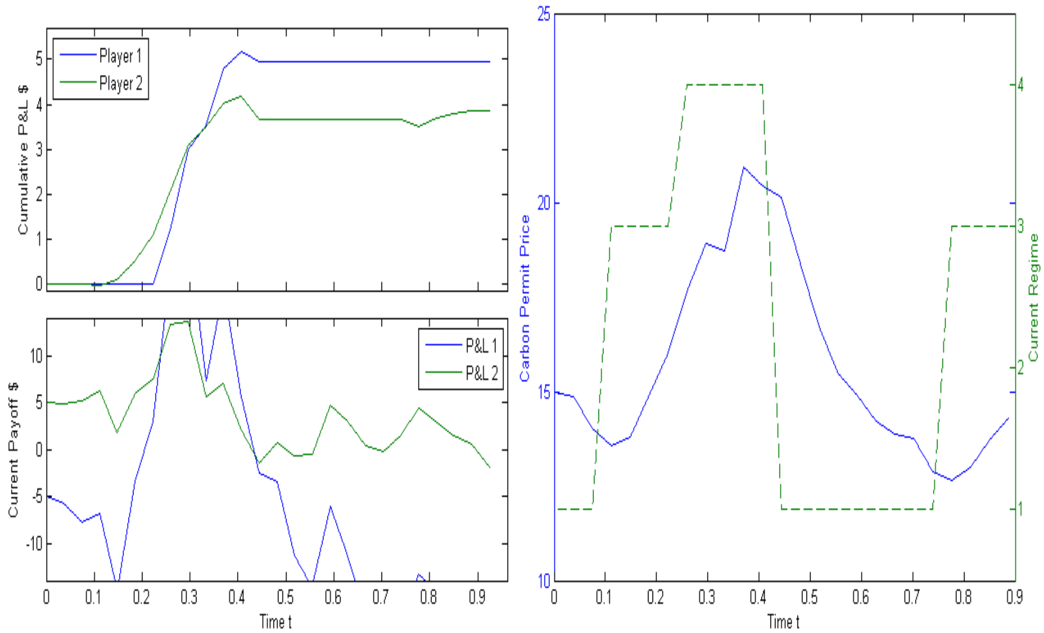

FIGURE 2. Sample path of equilibrium X_t^*, including the corresponding strategy $u^* \in \{00, 01, 10, 11\} \equiv \{1, \ldots, 4\}$. Top left panel: cumulative realized P&L of the players. Bottom left panel: the electricity-carbon spread of each producer for the current time step. Right panel: evolution of the controlled equilibrium allowance price X_t^*, as well as the implemented strategy $\vec{u}^*(t)$. Note that as $\vec{u}^*(t)$ increases (in the lexicographic order), emissions rise and X_t^* tends to increase.

In our simplified model, the producers only made binary emission decisions at each stage. On a practical level, much finer granularity is available. It would be straightforward to extend our problem and allow a more general finite-state control set of size $|\mathcal{A}|$. The only modification would be to replace the 2×2 bimatrix games with a more general $\mathcal{A} \times \mathcal{A}$ bimatrix. The theory for more than two producers is still not developed and it is an open problem to establish existence of CEP/NEP for multi-player stopping games.

6.1. **Further Extensions.** Several aspects of our model merit further analysis. First, our model for CO_2 allowance prices in (2) was highly stylized. It was selected to capture succinctly the price impact of each producer. However, many other features were left out. As described in the introduction, as the permit expiration date T approaches, the CO_2 price should converge either to zero (if excess permits remain) or to a fixed upper bound $\overline{x}$ (the penalty for emitting without an allowance). New (time-dependent) stochastic models are needed to mimic this property, see [7, 19]. Also, some cap-and-trade proposals will allow free trading of allowances by financial participants; if so, then *no-arbitrage* restrictions might have to be imposed on the dynamics of X. All these possibilities can in theory be handled straightforwardly, since the main construction is for arbitrary X-dynamics.

Ideally, a fully endogenous model is desired for allowance prices; namely X_t should be a function of total expected emissions until T compared to total current supply, i.e. have a characterization in terms of conditional expectations of future equilibrium emission schedules. See [6, 7, 10] for such price-formation models and related general equilibrium frameworks. These extensions will be considered in forthcoming papers.

Our formulation was in discrete-time; while this is sufficient for practical purposes, it is of great theoretical interest to construct a continuous-time model counterpart. The overall structure of a switching game as a sequence of stopping games straightforwardly carries over to continuous-time. However, description of correlated stopping equilibria in continuous time has not been attempted so far. In fact, the only reference dealing with randomized continuous-time stopping games is [41] (see also [28] for the latest results on general continuous timing games). Note that in continuous-time one must work with Nash ϵ-equilibria since all stopping strategies are defined only in the almost-sure sense. Second, to ensure the representation of V_i as iterative stopping games through $V_i^{n,m}$, it is necessary to *a priori* show that each player makes finitely many regime switches. At this point we are not able to state any conditions to guarantee this, except requiring mandatory "cool-off" periods between each emission regime switch.

In our Markovian setting, solutions of continuous-time single-player switching problems have representations in terms of reflected backward stochastic differential equations (BSDE) [22]. This representation should continue to hold in a game setting. Similar representations were obtained for stochastic differential game analogues of our setup, whence u_i is continuous, see [23, 24]. Note that the BSDE reflection mechanism for the game value is highly non-trivial and depends on the chosen communication device, as well as all eight possible continuation (payoff) functions $Z_i^{jk}(t)$. This direction will be explored in a separate paper.

Acknowledgment. I am deeply grateful to René Carmona for many useful discussions. I also thank David Ramsey, Krzysztof Szajowski, Jianfeng Zhang and participants at the MSRI Workshop on Economic Games and Mechanisms to Address Climate Change (May 2009) and the IPAM New Directions in Financial Mathematics Workshop (January 2010) for their valuable feedback on earlier versions of this paper.

References

[1] L. H. R. Alvarez. A class of solvable stopping games. *Appl. Math. Optim.*, 58(3):291–314, 2008.

[2] R. J. Aumann. Correlated equilibrium as an expression of Bayesian rationality. *Econometrica*, 55(1):1–18, 1987.

[3] S. Basar and G. J. Olsder. *Dynamic noncooperative Game Theory.* SIAM Classics in Applied Mathematics, 1999.

[4] M. Brennan and E. Schwartz. Evaluating natural resource investments. *J. Business*, 58:135–157, 1985.

[5] A. Calvó-Armengol. The set of correlated equilibria of 2 × 2 games. Technical report, Universitat Autonoma de Barcelona, 2006.

[6] R. Carmona, M. Fehr, and J. Hinz. Optimal stochastic control and carbon price formation. *SIAM J. Control Optim.*, 48(4):2168–2190, 2009.

[7] R. Carmona, M. Fehr, J. Hinz, and A. Porchet. Market design for emission trading schemes. *SIAM Review*, to appear, 2009.

[8] R. Carmona and M. Ludkovski. Pricing asset scheduling flexibility using optimal switching. *Appl. Math. Finance*, 15(4):405–447, 2008.

[9] J. F. Carrière. Valuation of the early-exercise price for options using simulations and nonparametric regression. *Insurance: Math. Econom.*, 19:19–30, 1996.

[10] M. Chesney and L. Taschini. The endogenous price dynamics of emission allowances: An application to CO_2 option pricing. Technical Report Research Paper No. 08-02, Swiss Finance Institute, 2009. Available at SSRN: http://ssrn.com/abstract=1090150.

[11] J. Cvitanić and I. Karatzas. Backward stochastic differential equations with reflection and Dynkin games. *Ann. Probab.*, 24(4):2024–2056, 1996.

[12] S. Dayanik and M. Egami. A direct solution method for optimal switching problems of one-dimensional diffusions, 2005. working paper.

[13] A. Dixit and R. S. Pindyck. *Investment Under Uncertainty.* Princeton University Press, 1994.

[14] E. J. Dockner, S. Jørgensen, N. V. Long, and G. Sorger. *Differential games in economics and management science.* Cambridge University Press, Cambridge, 2000.

[15] E. Dynkin. A game-theoretic version of an optimal stopping problem. *Soviet Math. Dokl.*, 10:270–274, 1969.

[16] D. Egloff. Monte Carlo algorithms for optimal stopping and statistical learning. *Ann. Appl. Probab.*, 15(2):1396–1432, 2005.

[17] E. Ekström and G. Peskir. Optimal stopping games for Markov processes. *SIAM J. Control Optim.*, 47(2):684–702, 2008.

[18] A. Eydeland and K. Wolyniec. *Energy and Power Risk Management: New Developments in Modeling, Pricing and Hedging.* John Wiley&Sons, Hoboken, NJ, 2003.

[19] M. Fehr and J. Hinz. A quantitative approach to carbon price risk modeling. Technical report, Institute for Operations Research, ETH Zurich, 2008.

[20] E. Z. Ferenstein. Randomized stopping games and Markov market games. *Math. Methods Oper. Res.*, 66(3):531–544, 2007.

[21] D. Fudenberg and J. Tirole. *Game Theory.* MIT Press, 1991.

[22] S. Hamadène and M. Jeanblanc. On the starting and stopping problem: application in reversible investments. *Math. Oper. Res.*, 32(1):182–192, 2007.

[23] S. Hamadène, J.P.Lepeltier, and S.Peng. BSDE with continuous coefficients and application to Markovian non-zero sum stochastic differential games. In N. N. El-Karoui and L. Mazliak, editors, *Pitman Research Notes in Mathematics Series*, volume 364, pages 115–128. Longman, Harlow, 1997.

[24] S. Hamadène and J. Lepeltier. Reflected BSDEs and mixed game problem. *Stochastic Process. Appl.*, 85(2):177–188, 2000.

[25] S. Hamadéne and J. Zhang. The continuous time nonzero-sum Dynkin game problem and application in game options, 2009. preprint, available at http://www-rcf.usc.edu/ jianfenz/Papers/HZ2.pdf.

[26] H. J. Kushner. Numerical approximations for nonzero-sum stochastic differential games. *SIAM J. Control Optim.*, 46(6):1942–1971, 2007.

[27] H. J. Kushner and P. Dupuis. *Numerical methods for stochastic control problems in continuous time*, volume 24 of *Applications of Mathematics*. Springer-Verlag, New York, second edition, 2001.

[28] R. Laraki, E. Solan, and N. Vieille. Continuous-time games of timing. *J. Econom. Theory*, 120(2):206–238, 2005.

[29] F. Longstaff and E. Schwartz. Valuing American options by simulations: a simple least squares approach. *Rev. Finan. Studies*, 14:113–148, 2001.

[30] M. Ludkovski. Financial hedging of operational flexibility. *Int. J. Theor. Appl. Finance*, 11 (8):799–839, 2008.

[31] R. Myerson. Multistage games with communication. *Econometrica*, 54(2):323–358, 1986.

[32] A. Nowak. On a new class of nonzero-sum discounted stochastic games having stationary Nash equilibrium points. *International Journal of Game Theory*, 32 (1):121–132, 2003.

[33] A. S. Nowak and T. E. S. Raghavan. Existence of stationary correlated equilibria with symmetric information for discounted stochastic games. *Mathematics of Operations Research*, 17 (3):519–526, 1992.

[34] Y. Ohtsubo. A nonzero-sum extension of Dynkin's stopping problem. *Math. Oper. Res.*, 12(2):277–296, 1987.

[35] Y. Ohtsubo. On a discrete-time non-zero-sum Dynkin problem with monotonicity. *J. Appl. Probab.*, 28(2):466–472, 1991.

[36] H. Pham and V. Ly Vath. Explicit solution to an optimal switching problem in the two-regime case. *SIAM J. Control Optim.*, 46(2):395–426, 2007.

[37] D. M. Ramsey and K. Szajowski. Selection of a correlated equilibrium in Markov stopping games. *European J. Oper. Res.*, 184(1):185–206, 2008.

[38] D. Rosenberg, E. Solan, and N. Vieille. Stopping games with randomized strategies. *Probab. Theory Related Fields*, 119(3):433–451, 2001.

[39] E. Shmaya and E. Solan. Two-player nonzero-sum stopping games in discrete time. *Ann. Probab.*, 32(3B):2733–2764, 2004.

[40] E. Solan and R. V. Vohra. Correlated equilibrium in quitting games. *Math. Oper. Res.*, 26 (3):601–610, 2001.

[41] N. Touzi and N. Vieille. Continuous-time Dynkin games with mixed strategies. *SIAM J. Control Optim.*, 41(4):1073–1088, 2002.

(M. Ludkovski) Department of Statistics and Applied Probability, University of California, Santa Barbara, CA 93106-3110, USA

E-mail address: ludkovski@pstat.ucsb.edu, www.pstat.ucsb.edu/faculty/ludkovski