Managing Shutdown Decisions in Merchant Commodity and Energy Production: A Social Commerce Perspective

Alessio Trivella
Department of Management Engineering, Technical University of Denmark
atri@dtu.dk

Selvaprabu Nadarajah*
College of Business Administration, University of Illinois at Chicago
selvan@uic.edu

Stein-Erik Fleten
Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology
stein-erik.fleten@iot.ntnu.no

Denis Mazieres
Department of Economics, Mathematics and Statistics, Birkbeck University of London
denismazieres@gmail.com

David Pisinger
Department of Management Engineering, Technical University of Denmark
dapi@dtu.dk

Initial version: September, 2017; This version: May, 2019

Problem definition: Merchant commodity and energy production assets operate in markets with volatile prices and exchange rates. Plant closures adversely affect societal entities beyond the specific plant being shutdown such as the parent company and the local community. Motivated by an aluminum producer, we study if mitigating these hard-to-assess broader impacts of a shutdown is financially viable using the plant's operating flexibility. ***Academic/Practical relevance:*** Our social commerce perspective towards managing shutdown decisions deviates from the commonly used asset value maximization objective in merchant operations. Identifying operating policies that delay or decrease the likelihood of a shutdown without incurring a significant asset value loss supports socially-responsible plant shutdown decisions. ***Methodology:*** We formulate a constrained Markov decision process to manage shutdown decisions and limit the probability of future plant closures. We provide theoretical support for approximating this intractable model using unconstrained stochastic dynamic programs with modified shutdown costs and explore two classes of operating policies. Our first policy leverages anticipated regret theory while the second policy generalizes production-margin heuristics used in practice using machine learning. We compute the former and latter policies using a least squares Monte Carlo method and combining this method with binary classification, respectively. ***Results:*** Anticipated-regret policies possess desirable asymptotic properties absent in classification-based policies. On instances created using real data, anticipated-regret and classification-based policies outperform practice based production-margin strategies. Significant reductions in shutdown probability and delays in plant closures are possible while incurring small asset value losses. ***Managerial implications:*** A plant's operating flexibility provides an effective lever to balance the social objective to reduce closures and the financial goal to maximize asset value. Adhering to both objectives requires combining short-term commitments with external stakeholders to avoid shutdown with longer-term internal efforts to reduce the probability of plant closures.

*corresponding author

1. Introduction

Commodity and energy production assets play a critical role in the supply of ethanol, aluminum, copper, iron, and steel, and power from coal, natural gas, and renewable sources. In 2015, iron ore, copper, and aluminum were respectively \$225, \$130, and \$90 billion US dollars annual industries (IMF 2015), and roughly \$10 trillion US dollars will be invested in new power generation by 2040 (BNEF 2017). Based on work with a major base metal producer, we study the management of permanent shutdown decisions in a merchant commodity and energy production asset.

For illustration, consider an aluminum production facility that takes as inputs bauxite, carbon, and electricity, and produces aluminum to be sold into the wholesale market. Production is economical when the conversion spread between the aluminum price and the input prices is positive after accounting for production costs. If this spread is negative the merchant may temporarily suspend production. If a combination of production and suspension leads to losses it may be economical to permanently shut down the plant. Analogous operational flexibility is common in other commodity production assets, refineries and hydrocarbon crackers in the petro-chemical industry, and power plants (Brennan and Schwartz 1985, Tseng and Barz 2002, Méndez et al. 2006, Cortazar et al. 2008, Adkins and Paxson 2011, Boyabatli et al. 2011, Kazaz and Webster 2011, Boyabatli 2015, Nadarajah et al. 2016, Hekimoğlu et al. 2016, Boyabatli et al. 2017), including renewable energy production, for example in biogas plants (Di Corato and Moretto 2011, Hochloff and Braun 2014).

A merchant producer can maximize the market value of the plant by adapting production, suspension, and shutdown decisions over time to the evolution of uncertain market factors such as prices of commodity and energy sources and exchange rates (Secomandi and Seppi 2014). While this pure asset value perspective is popular, the cost of a permanent plant shutdown is hard to assess as it may impact societal entities outside the specific plant being shut down, which could include the parent company owning the plant and the local community. For example, metal producers (e.g., aluminum and steel) often own both a production plant and a power source, such as a dam, due to the large amounts of power consumed during the production process. It is common for licenses to operate plants and power sources to include social and economic criteria that the producer must meet (e.g., maintaining local jobs). Abandoning a plant can make the producer default on these obligations and the government to deny the renewal of operating licenses for the power source as well as jeopardize licenses for assets in other regions. In addition, such a shutdown can cause loss of employment and adverse publicity and result in political resistance from powerful unions and interest groups (see Kasa 2000 for evidence of institutionalized links between unions and government).

Given the gravity of a plant shutdown on society, companies deviate from a pure asset value perspective: several years of severely challenged profitability are typically needed to justify shutdown to a labor union, but waiting to incur such losses before shutting down may be at odds with maximizing asset value. This focus on "social commerce" is evident, for instance, from the public web pages of major producers of aluminum, coal, copper, oil, and gas, as well as, related stewardship initiatives and ISO 26000 standards adopted by firms (Gazette-Mail 1999, Codelco 2011, ISO 2014, ASI 2017, Hydro 2018, IPIECA 2019; see also Kleindorfer et al. 2005, Kaya 2016, Lee and Tang 2017, and references therein, for other examples of socially responsible operations and reporting requirements). In sum, producers would benefit from models and methods to manage asset value and the broader impact of plant closures in a principled manner under evolving market factors, which include overproduction, unfavorable commodity prices, and volatility in the prices of production inputs/outputs and exchange rates (BI 2016, Reuters 2018).

We define the shutdown probability of a plant's operating policy at a given time as the probability of it closing down by this time. Each policy can thus be associated with a shutdown profile, that is, its shutdown probability at each time period over a finite planning horizon. We formulate a constrained Markov decision process (henceforth, social-commerce MDP or SC-MDP for short) to maximize the plant's market value subject to restrictions on the shutdown profile of its operating policy that capture social preferences. Such restrictions could model agreements between the plant and the local union to avoid shutdown for a few years by requiring the probability of shutdown to equal zero in these years. Constraints on the shutdown profile could also model internal corporate efforts to reduce the likelihood of a shutdown in early parts of the planning horizon and/or reducing the shutdown probability at the end of this horizon. SC-MDP is a constrained version of the well-known risk-neutral MDP formulation (henceforth referred to as the shutdown-neutral MDP) that maximizes the asset value alone (Guthrie 2009). Indeed, optimal policies of the shutdown-neutral MDP may have socially undesirable shutdown profiles. If the loss in asset value due to constraining the shutdown profile in SC-MDP is small, then producers may be able to justify the operations of the plant for a longer period of time on the basis of unaccounted social costs.

Solving a constrained MDP is in general significantly more challenging than its unconstrained counterpart. In the specific case of SC-MDP, constraints limit the probability of shutdowns at future stages, which is an effect that can intuitively be achieved by modifying the fixed shutdown cost in the shutdown-neutral MDP. We thus explore if such a modified shutdown-neutral MDP can recover an optimal policy of SC-MDP. We show that (static) shutdown cost modifications not adapted to the evolution of the uncertain MDP state, while tractable, are not enough to recover an optimal policy of SC-MDP. On the other hand, we establish that state-dependent modifications suffice for fully recovering an optimal policy of SC-MDP (under mild assumptions) and we also characterize situations when only partial recovery is possible. While our analysis lends conceptual support for pursuing shutdown cost modifications, finding state-dependent modifications is computationally challenging because the number of cost parameters that need to be tuned scales with the size of the MDP state space, which is high dimensional when using realistic models for the evolution of uncertain factors.

We thus propose a low-dimensional state-dependent shutdown cost modification of the shutdown-neutral MDP that can be interpreted from a social commerce perspective as trying to mitigate the possibility that shutdown decisions could lead to regrettable social outcomes as future uncertainty unfolds. Such a regrettable outcome would arise if a holding company closes a plant but economic conditions turn out to be better than expected, which makes it clear that keeping the plant open would have been advantageous. We highlight that this notion of regret is consistent with anticipated regret (AR) theory (Bell 1982, Loomes and Sugden 1982, Zeelenberg 1999) in behavioral economics. Formally, this leads to a stochastic dynamic program (SDP), dubbed the anticipated-regret (AR) SDP, where the shutdown fixed cost is inflated by an anticipated-regret term that accounts for the expected loss on future scenarios where a non-shutdown decision provides higher utility than shutting down. We establish that the optimal policy of this SDP is consistent with SC-MDP in asymptotic regimes where the social preferences are either shutdown-neutral or extremely shutdown-averse (i.e., require zero shutdowns). We adapt a least squares Monte Carlo (LSM) heuristic (Longstaff and Schwartz 2001, Glasserman and Yu 2004, Nadarajah et al. 2017) to overcome the high-dimensionality of the AR SDP state space and compute policies.

In practice, production margin-based policies are popular to determine when to decide between production and shutdown. Such policies choose to shut down when a sum of current and expected-future production margins are below a threshold, which facilitates showing to external stakeholders the challenged profitability leading to a potential shutdown decision. We formalize production margin policies used in industry and then combine approximate dynamic programming and machine learning to improve them by incorporating social preferences toward plant closures. Specifically, we create new policies using binary classification (Bishop 2006, chapter 4) where thresholds are determined on training data (i.e., operating decisions) generated by simulating a shutdown-neutral policy obtained using LSM. These thresholds are subsequently modified to account for shutdown profile preferences. We refer to the resulting policy as the classification-based margin (CM) policy.

We perform a numerical study involving the operations of a real aluminum producer over a forty-year time horizon. Our case study uses operational data from this producer and market data from the Nord Pool, London Metal Exchange (LME), and FOREX markets. We calibrate an eight-factor stochastic model to capture the evolution of uncertainty, which includes electricity and aluminum prices and exchange rates (Farkas et al. 2017). We use an LSM method to obtain a near optimal Monte Carlo simulation estimate of the maximum asset value benchmarked against a dual upper bound estimate (Brown et al. 2010). We also employ LSM to obtain a collection of AR and CM policies for different social preferences including commitements to avoid shutdown for a few years and reductions in the probability of shutdowns. We find that both policies can substantially improve the shutdown profile compared to that of the shutdown-neutral policy for small asset value losses. For instance, AR and CM policies can decrease the shutdown probability by 25% and, in addition, delay shutdown decisions by an average of 4 years for a 4% loss in asset value. The CM policies dominate the practice-based margin policies for managing the trade-off between shutdowns and asset value, even though the latter rules exhibit good performance from a value maximization standpoint.

Insights from our numerical study inform the socially responsible management of shutdown decisions in merchant production assets beyond aluminum smelting. First, socially responsible plant closure decisions can also be financially responsible, that is, significant improvements in the shutdown profile of a plant may be possible for small losses in asset value that are outweighed by unaccounted social costs. Identifying such win-win situations requires producers to proactively manage the shutdown profile of the plant's operating policy. Second, both deterministic and probabilistic social commerce objectives are useful when managing plant closures. While a (deterministic) guarantee to avoid plant shutdowns is easier to communicate to external stakeholders, such a commitment may only be feasible in the short-term. Intuitively, longer-term production guarantees may lead to significant financial loss as it translates to production in financial unfavorable states, which are more likely in the long run. Thus, constraining the shutdown profile of an operating policy is a necessary internal social objective to promote the long term viability of a plant. Finally, AR and CM policies indicate that judiciously modifying the shutdown-cost of risk-neutral MDP and the switching-thresholds of margin policies, respectively, are effective strategies to incorporate social preferences into otherwise purely asset-value driven plant operating decisions.

Our work on managing the trade-off between shutdown probability and asset value adds to the literature on merchant commodity and energy operations (Geman 2005, Lai et al. 2010, Berling and Martínez-de-Albéniz 2011, Devalkar et al. 2011, Wu et al. 2012, Nadarajah et al. 2015, Secomandi 2015, Thompson 2016, Devalkar et al. 2005, Nadarajah et al. 2017), and socially responsible operations (Kleindorfer et al. 2005, Lee and Tang 2017). Models of production assets as switching options have been considered, for example, by Kulatilaka and Trigeorgis (2001), Tseng and Barz (2002), and Adkins and Paxson (2011). These papers model a temporary shutdown option. Brennan and Schwartz (1985), Cortazar et al. (2008), Nadarajah and Secomandi (2017), and Yang et al. (2017) consider switching options with permanent shutdown but focus on risk-neutral valuation and operations. The growing literature on socially responsible operations incorporates social considerations into operations models but our social commerce view on shutdown decisions in merchant energy production is new (see Lee and Tang 2017 and references therein). We also complement research in this literature and merchant operations by introducing methods that are grounded in anticipated regret theory and practitioner heuristics, as well as applying them to a real aluminum case study using a state-of-the-art multi-commodity and multi-factor stochastic process (Farkas et al. 2017) .

Our analysis and methodology for approximating SC-MDP by modifying the cost of an (unconstrained) MDP contributes to the constrained MDP literature, where practical methods to solve high-dimensional problems are limited (Dufour and Prieto-Rumeau 2013, 2014) and, in addition, standard approaches for tackling low-dimensional problems, such as a lagrangian techniques (Altman 1999), are not easily interpretable from a social commerce perspective. The use of AR theory to obtain a tractable shutdown cost modification is thus novel and is consistent with experimental research showing that AR is large for irreversible decisions and negligible for reversible decisions (see, e.g., Tsiros and Mittal 2000). We note that our use of AR theory is also judicious as its alternatives such as prospect theory, rank-dependent expected utility, and induced preferences (Tversky and Kahneman 1992, Eeckhoudt et al. 2005, Chapter 13, Bleichrodt and Wakker 2015) are not tailored for comparing actions at a given state and thus entail more choices for implementation. For instance, using prospect theory requires specifying weighting functions, value functions, and reference points (Tversky and Kahneman 1992, Bleichrodt and Wakker 2015).

The combination of LSM and machine learning classification methods in this paper adds to the extant research that obtains heuristic policies to high-dimensional real option problems (see, e.g., Glasserman 2004, Chapter 8, Cortazar et al. 2008, Carmona and Ludkovski 2010, Mazières and Boogert 2013, Nadarajah et al. 2015, 2017) and the literature on decision rule approximations (Ben-Tal et al. 2004, Georghiou et al. 2018). The LSM approach has regress-now (Longstaff and Schwartz 2001) and regress-later (Glasserman and Yu 2004, Nadarajah et al. 2017) variants that approximate the SDP continuation and value functions, respectively. The computation of risk-neutral policies in a commodity production application with a permanent shutdown decision has been approached by Cortazar et al. (2008) using regress-now LSM and by Nadarajah and Secomandi (2017) using regress-later LSM. The use of LSM for dual bounding (Haugh and Kogan 2004, Brown et al. 2010) in energy production has been explored by the latter paper and Yang et al. (2017). Unlike these papers, we adapt LSM to account for the shutdown profile of operating policies. Our combination of regress-later LSM with binary classification to learn margin-based policies, that is the CM policy, and its modification to satisfy policy constraints are both new and add to the decision rule approximations literature. Prior work in this literature has considered affine and piece-wise affine decision rules for approximating stochastic optimization problems (see, e.g., Georghiou et al. 2018).

The rest of this paper is organized as follows. We formulate a model of a merchant production plant and formalize the notion of shutdown profile in §2. We introduce polices based on shutdown cost modifications in §3. We describe production margin-based policies in §4. We discuss our numerical study and findings in §5 followed by conclusions in §6. Additional details regarding proofs, model calibration, and algorithms are provided in an Online Supplement.

2. Merchant commodity and energy production operations

In this section we discuss the merchant operations of commodity/energy production assets. We present in §2.1 a model for the operations of such assets that embed a permanent shutdown option. To identify a specific operating policy, we describe the known asset value maximization perspective in §2.2 and then introduce a social commerce perspective in §2.3.

2.1 Operating model

We consider a commodity and energy production facility operating over I time periods (stages). This plant produces output and sells it into the wholesale market when prices are favorable, that is, it operates in a merchant fashion. We model the plant's operating problem as an MDP, where at each stage $i \in \mathcal{I} := \{0, \dots, I-1\}$ the plant could be open or permanently shutdown. Examples of open states include in production, ramping up/down to a higher/lower production rate, and temporary suspension. We denote the finite set of open states at stage i by $\mathcal{O}_i$ and assume that this set includes at least the state corresponding to full production labeled O. The permanent shutdown state is represented by C. The stage i operating status belongs to set $\mathcal{X}_i := \mathcal{O}_i \cup \{\mathrm{C}\}$.

The full state of the production asset is composed of the operating (endogenous) state component described above as well as a market information (exogenous) state component that affects the operating cash flows. Let w_i denote the vector of stochastic market factors at stage i driving the evolution of market information (e.g., commodity/energy prices and exchange rates). We assume the vector w_i evolves in a Markovian manner and denote its support by $\mathcal{W}_i$.

We assume that the plant is initially in an open state, that is, the known stage-0 state represented by (x_0, w_0) satisfies $x_0 \in \mathcal{O}_0$. A decision a_i from the feasible action set $\mathcal{A}_i(x_i)$ executed at stage $i \in \mathcal{I}$ and state $(x_i, w_i) \in \mathcal{X}_i \times \mathcal{W}_i$ results in an immediate reward $r_i(x_i, w_i, a_i)$ and a transition to a stage

$i+1$ operating state $f_i(x_i, a_i)$ in set $\mathcal{X}_{i+1}$. We define a terminal stage I where no action is allowed and a terminal reward $r_I(x_I, w_I)$ is received; the pair (x_I, w_I) belongs to the terminal-stage state space $\mathcal{X}_I \times \mathcal{W}_I$. Since $f_i(x_i, a_i)$ is an element of $\mathcal{X}_{i+1}$ we use the next stage operating state as action labels, that is, taking action $a_i = x_{i+1}$ at stage i and operating state x_i results in a transition to the stage $i+1$ operating status $f_i(x_i, x_{i+1}) = x_{i+1}$. To account for the permanent nature of a shutdown decision, we impose the following conditions on the action set and transition function: $\mathcal{A}_i(\mathrm{C}) = \{\mathrm{C}\}$ and $f_i(\mathrm{C}, \mathrm{C}) = \mathrm{C}$ for all $i \in \mathcal{I}$. A shutdown decision at stage $i \in \mathcal{I}$ incurs a one time fixed cost, which we model by requiring $r_i(x_i, w_i, \mathrm{C}) = -K^{(x_i, \mathrm{C})}$ for all $(x_i, w_i) \in \mathcal{O}_i \times \mathcal{W}_i$ and $r_i(\mathrm{C}, w_i, \mathrm{C}) = 0$, where $K^{(x_i, \mathrm{C})}$ is the strictly positive fixed cost of shutting down at an open state $x_i \in \mathcal{O}_i$.

This operating model can be specialized to capture a fairly broad class of production assets including coal and natural gas power plants, ethanol and biogas plants, and metal smelters. An example application of this model can be found in §5, where we use it to describe the merchant operations of a real aluminum plant as part of our numerical study.

2.2 Asset value maximization perspective

To make decisions in the operating model of §2.1, the plant manager requires a policy, which is a collection of stage-specific decision rules. At a given stage i, a decision rule A_i^π specifies a feasible operating decision in $\mathcal{A}_i(x_i)$ for each state $(x_i, w_i) \in \mathcal{X}_i \times \mathcal{W}_i$. A policy π is then the collection $\{A_i^\pi, i \in \mathcal{I}\}$, and the set of all feasible policies is denoted by Π. The value of the production asset when using a policy $\pi \in \Pi$ is the discounted sum of expected cash flows from using this policy over the finite problem horizon. The policy that maximizes the asset value thus solves

$$\max_{\pi \in \Pi} \mathbb{E}_0 \left[\sum_{i \in \mathcal{I}} \delta^i r_i \left(x_i^\pi, w_i, A_i^\pi(x_i^\pi, w_i) \right) + \delta^I r_I \left(x_I^\pi, w_I \right) \right], \tag{1}$$

where $\mathbb{E}_0$ denotes expectation with respect to a Markovian stochastic process describing the distribution of w_i given the stage 0 information w_0 (see (8) for an example of such a process), $\delta \in (0, 1]$ is the discount factor, and x_i^π the random endogenous state reached in stage i by following policy π. We refer to the operating policy solving (1) as the *shutdown-neutral* optimal policy and name it π^{SN}. It is well known that this policy is the solution of the following stochastic dynamic program (SDP):

$$V_I^{\mathrm{SN}}(x_I, w_I) = r_I(x_I, w_I), \ \forall (x_I, w_I) \in \mathcal{X}_I \times \mathcal{W}_I, \tag{2a}$$

$$V_i^{\mathrm{SN},\mathcal{O}}(x_i, w_i) = \max_{a_i \in \mathcal{A}_i(x_i) \setminus \{\mathrm{C}\}} \left\{ r_i(x_i, w_i, a_i) + \delta \mathbb{E}_i \left[V_{i+1}^{\mathrm{SN}} \left(f_i(x_i, a_i), w_{i+1} \right) \right] \right\}, \quad \forall (i, x_i, w_i) \in \mathcal{I} \times \mathcal{X}_i \times \mathcal{W}_i, \tag{2b}$$

$$V_i^{\mathrm{SN}}(x_i, w_i) = \max \left\{ V_i^{\mathrm{SN},\mathcal{O}}(x_i, w_i), -K^{(x_i, \mathrm{C})} \right\}, \ \forall (i, x_i, w_i) \in \mathcal{I} \times \mathcal{X}_i \times \mathcal{W}_i. \tag{2c}$$

Here $V_i^{\mathrm{SN}}(\cdot,\cdot)$ denotes the shutdown-neutral value function at stages i in set $\mathcal{I} \cup \{I\}$. The value function at the terminal stage is specified in (2a) using the terminal reward $r_I(\cdot,\cdot)$. In (2b), we define an intermediate value function $V_i^{\mathrm{SN},\mathcal{O}}(x_i, w_i)$ that excludes the shutdown action at stage i. Here $\mathbb{E}_i$ denotes expectation given stage i information w_i. We define $V_i^{\mathrm{SN}}(x_i, w_i)$ as the maximum of $V_i^{\mathrm{SN},\mathcal{O}}(x_i, w_i)$ and the shutdown cost in (2c), that is, this equation models the choice between the shutdown action and the best open action selected in (2b).

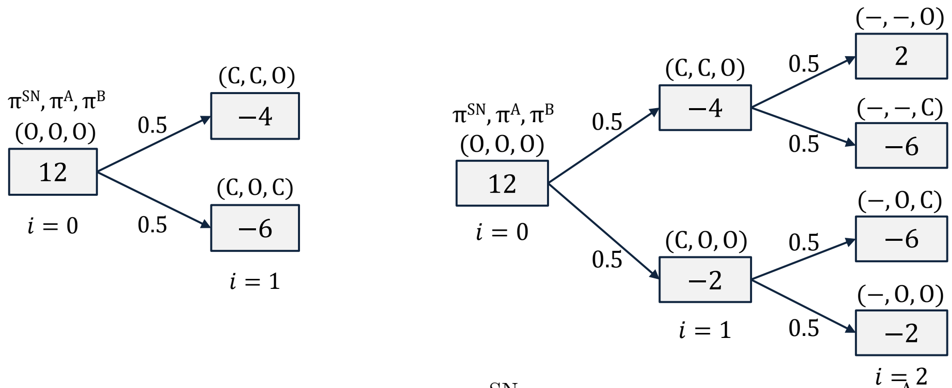
2.3 Social commerce perspective

As discussed in §1, we model social preferences by associating a shutdown profile with each policy $\pi \in \Pi$, defined as the collection of shutdown probabilities $\{\Pr(x_i^\pi = \mathrm{C} \,|\, x_0, w_0),\, i \in \mathcal{I} \cup \{I\}\}$. Note that the shutdown probability increases with the stage index as plant closures are irreversible. Social preferences include agreements with stakeholders (e.g., unions) as well as internal efforts to curb shutdowns. A common deterministic commitment with stakeholders is to avoid shutting down the plant for the immediate T future years, which translates to requiring $\Pr(x_i^\pi = \mathrm{C} \,|\, x_0, w_0) = 0$ for $i \in \{1, \ldots, T\}$. Internally, firms strive to delay and reduce the likelihood of shutdowns over the planning horizon. Reducing the shutdown probability entails finding an operating policy π where the probability at the end of the planning horizon $\Pr(\mathrm{C}; \pi) := \Pr(x_I^\pi = \mathrm{C} \,|\, x_0, w_0)$ is smaller than the analogous probability $\Pr(\mathrm{C}; \pi^{\mathrm{SN}})$ under the shutdown-neural policy. Delaying shutdowns on the other hand requires emphasizing shutdown probability decreases at early stages greater than at later stages. Shutdown delays could be measured by computing the expected time to shutdown along sample paths of uncertainty where a shutdown decision is chosen. Since $\Pr(x_i^\pi \neq \mathrm{C},\, x_{i+1}^\pi = \mathrm{C} \,|\, x_0, w_0)$ represents the probability that π will shut down exactly at stage $i \in \mathcal{I}$, the expected time to shutdown of policy π can be formally defined as $\sum_{i \in \mathcal{I}} i \cdot \Pr(x_i^\pi \neq \mathrm{C},\, x_{i+1}^\pi = \mathrm{C} \,|\, x_0, w_0) / \Pr(\mathrm{C}; \pi)$. Examples 1 and 2 illustrate the notions of shutdown probability reduction and delay, also highlighting that achieving a balance between asset value and shutdown profile can be non-trivial. Specifically, Example 1 shows that multiple policies can result in the same shutdown probability but different asset value, while Example 2 presents a case where multiple policies have the same combination of asset value and shutdown probability but delay shutdown differently.

Example 1 (Shutdown probability reduction). *Consider the operations of a stylized plant that can either produce or shut down over two periods $i \in \mathcal{I} = \{0, 1\}$. The cost of shutdown is equal to 2 at each stage. Figure 1(a) displays the reward (inside the rectangular box) from producing at time* 0 *(now) as well as the two equally likely random rewards from producing at period* 1 *(future). We ignore discounting for simplicity. The coordinates of the triples in Figure 1(a) contain the*

Figure 1: Production rewards, probabilities, and policies for examples 1 and 2.

(a) Example 1. (b) Example 2.



decisions taken by the shutdown-neutral policy π^{SN} *and two heuristic policies* π^{A} *and* π^{B}*. The optimal shutdown-neutral policy* π^{SN} *produces in period* 0 *but shuts down in both states in period 1. The value of this policy is* 10 $(= 12 - 2 \cdot 0.5 - 2 \cdot 0.5)$ *and its shutdown probability is 100%. Next, consider the policies* π^{A} *and* π^{B}*. Both policies have a shutdown probability of* 50% *and the values of* π^{A} *and* π^{B} *are* 8 $(= 12-2\cdot 0.5-6\cdot 0.5)$ *and* 9 $(= 12-4\cdot 0.5-2\cdot 0.5)$*, respectively. Thus, it is possible to reduce shutdown probability by* 50% *for a* 10% *decrease in asset value using* π^{B}*, but choosing* π^{A} *instead entails a* 20% *decrease in asset value to achieve the same shutdown probability reduction.*

Example 2 (**Shutdown delay).** *Consider a plant operating for three-periods with production rewards and probabilities summarized in Figure 1(b). The shutdown cost at each stage equals 2. The coordinates of the triples in Figure 1(b) contain the decisions taken by the shutdown-neutral policy* π^{SN} *and two heuristic policies* π^{A} *and* π^{B}*. The value of* π^{SN} *is* 10 *and its shutdown probability is 100%. For both* π^{A} *and* π^{B}*, the shutdown probability and asset value are 50% and 8, respectively, but these policies shutdown at different periods. The expected time to shutdown on sample paths where* π^{A} *chooses to shutdown is equal to* 1 $(= [0.5 \cdot 1]/0.5)$*. The analogous measure for* π^{B} *evaluates to* 2 $(= [0.25 \cdot 2 + 0.25 \cdot 2]/0.5)$*. Therefore, for a* 20% *decrease in asset value, both* π^{A} *and* π^{B} *reduce shutdown probability by* 50% *but the latter policy delays shutdown by one more period on average.*

Finding operating policies with improved shutdown profiles that do not result in losing significant asset value is critical to the financial viability of partaking in social commerce. Conceptually, social preferences on the shutdown profile of a policy can be modeled by adding bounds on the shutdown probabilities at each stage to MDP (1):

$$\max_{\pi \in \Pi} \mathbb{E}_0 \left[\sum_{i \in \mathcal{I}} \delta^i r_i \left(x_i^\pi, w_i, A_i^\pi(x_i^\pi, w_i) \right) + \delta^I r_I \left(x_I^\pi, w_I \right) \right] \tag{3a}$$

$$\text{s.t. } \Pr(x_i^\pi = \mathrm{C} \mid x_0, w_0) \leqslant U_i, \quad \forall i \in \mathcal{I} \cup \{I\}. \tag{3b}$$

Here, upper bounds U_i on the shutdown probabilities are used to control the preferences to reduce and delay shutdown decisions. If all these bounds are equal to 1, then MDP (3) reduces to the shutdown-neutral MDP (1) which only focuses on maximizing asset value. If the first $T+1$ bounds are equal to zero and the remaining ones equal to one (i.e., $U_i = 0$ for $i = 0, \ldots, T$ and $U_i = 1$ for $i = T+1, \ldots, I$), then MDP (3) corresponds to the shutdown-neutral MDP (1) in which the shutdown action C is removed from T future stages and allowed afterwards. This case corresponds to a deterministic commitment to avoid shutdown for the first T years. The shutdown-averse extreme of this trade-off includes policies with zero shutdowns. Among them, the one maximizing asset value is obtained by setting $U_i = 0$ for each i in (3), that is, completely removing the shutdown action C. This policy, labeled $\pi^{\mathrm{SN}\setminus\{\mathrm{C}\}}$, can be determined by solving the following modified version of SDP (2):

$$V_i^{\mathrm{SN}\setminus\{\mathrm{C}\}}(x_i, w_i) = \max_{a_i \in \mathcal{A}_i(x_i)\setminus\{\mathrm{C}\}} \Big\{ r_i(x_i, w_i, a_i) + \delta \mathbb{E}_i \big[V_{i+1}^{\mathrm{SN}\setminus\{\mathrm{C}\}}\left(f_i(x_i, a_i), w_{i+1}\right) \big] \Big\},$$

$$\forall (i, x_i, w_i) \in \mathcal{I} \times \mathcal{X}_i \times \mathcal{W}_i. \tag{4}$$

We exclude boundary conditions here and later SDPs as they are analogous to (2a).

Although the shutdown-neutral and zero-shutdown policies can be characterized, what is needed in practice is the ability to define a collection of policies that offer different trade-offs between shutdown profile and asset value so that managers can choose one that is acceptable, possibly taking into account the non-financial costs of shutdown and other strategic considerations. To obtain these policies, ideally we would solve our constrained MDP (3) for different choices of the bounds U_i. However, the solution of constrained MDPs is in general considerably more challenging than MDPs (Dufour and Prieto-Rumeau 2013, 2014). The challenge of solving our constrained MDP is more acute as it is also high-dimensional under realistic models for the evolution of the uncertainty such as the one used in our numerical study and described in §5.2.

3. Operating policies based on modified shutdown costs

SC-MDP strives to compute policies with smaller future shutdown probabilities than under π^{SN}. Intuitively, analogous improvements in the shutdown profile of π^{SN} would result if plant closures incurred higher fixed costs. Motivated by this observation, we analyze in §3.1 if an optimal policy of SC-MDP can be recovered by solving a variant of the shutdown-neutral MDP with modified shutdown costs. In §3.2, we formulate a shutdown-averse SDP based on a tractable shutdown cost modification scheme that can be interpreted from a social commerce perspective. In §3.3, we describe an LSM approach to tackle this SDP and compute policies.

3.1 Analysis of shutdown cost modifications

Suppose the shutdown neutral optimal policy π^{SN} does not satsify constraints (3b). Gradually inflating the shutdown cost $K^{(x_i,\mathrm{C})}$ in the shutdown-neutral MDP will lead to fewer plant closures and result in these constraints being satisfied. While a feasible policy to SC-MDP can be constructed in this manner, it is unclear if the resulting policy coincides with an optimal policy of SC-MDP.

We begin by analyzing the optimal policy of a variant of the shutdown-neutral SDP (2) where the shutdown cost $K^{(x_i,\mathrm{C})}$ is replaced by stage-dependent parameters $\{\bar{K}_i^{(x_i,\mathrm{C})} \in \mathbb{R}, (i, x_i) \in \mathcal{I}\backslash\{I\} \times \mathcal{X}_i\}$. Example 3 shows that stage-dependent modifications may not recover an optimal policy of SC-MDP.

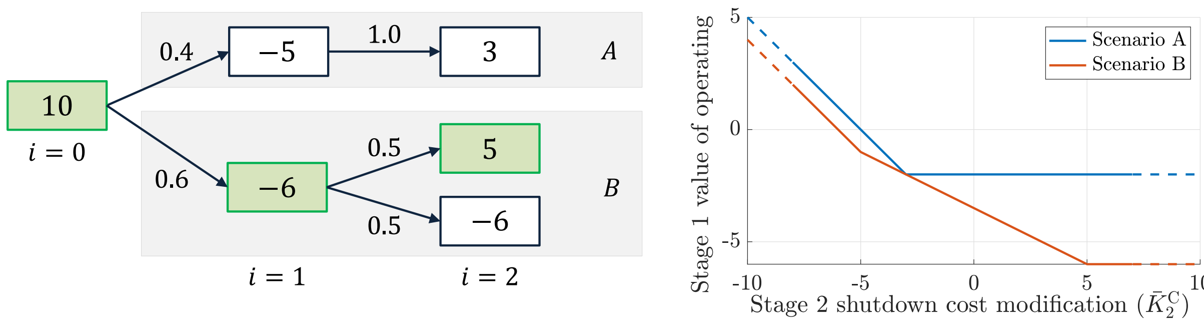
Example 3. *Consider the three-stage instance with a shutdown cost of $K = 1$ and the shutdown probability bound vector $U = (U_i)_i = (0, 0.5, 0.8)$. At stage 1, two scenarios A and B are possible with deterministic rewards from producing denoted by r_1^{A} and r_1^{B}, respectively. Conditioned on these stage-1 scenarios, we have random stage-2 rewards $\tilde{r}_2^{\mathrm{A}}$ and $\tilde{r}_2^{\mathrm{B}}$. Figure 2 illustrates the production rewards and scenario probabilities. The optimal shutdown-neutral policy produces at stage 0 and shuts down in both stage-1 scenarios. This policy has a value of 9 $(= 10 - 1)$ and a shutdown profile given by the vector $(0, 1.0, 1.0)$. The optimal SC-MDP policy produces in the shaded nodes in Figure 2(a), has a value of 7.2 $(= 10 - 0.4 \cdot 1 + 0.6 \cdot [-6 - 0.5 \cdot 1 + 0.5 \cdot 5])$, and the shutdown profile $(0, 0.4, 0.7)$. We highlight that due to the probability bound U, this policy shuts down at stage 1 in scenario A but not in scenario B. This behavior cannot be replicated using a stage dependent cost modification as explained next. Suppose we modify the stage 2 costs using the parameter $\bar{K}_2^{\mathrm{C}}$. The stage 1 expected value of operating in scenario A $(= -5 + \max\{3, -\bar{K}_2^{\mathrm{C}}\})$ is greater than or equal to the analogous value in scenario B $(= -6 + 0.5 \cdot \max\{5, -\bar{K}_2^{\mathrm{C}}\} + 0.5 \cdot \max\{-6, -\bar{K}_2^{\mathrm{C}}\})$ for any stage-dependent shutdown cost modification that can be selected at stage 2 as shown in Figure 2(b). Therefore, no stage-dependent shutdown cost modification can alter this preference order, that is, swap the stage-1 decision from shut down to operate in scenario B but not in scenario A. Thus, there does not exist a stage-dependent cost modification that recovers an optimal SC-MDP policy.*

Example 3 highlights value in adapting the shutdown cost modifications to the exogenous state w_i as this additional flexibility may facilitate recovering the SC-MDP optimal policy. Proposition 1 establishes that this is indeed the case. We denote an SC-MDP optimal policy by π^* and by π^{D} an optimal policy obtained by solving the shutdown-neutral SDP (2) with modified shutdown fixed costs. Further, we use $\bar{K}_i^{(x_i,w_i,\mathrm{C})}$ to represent the state-dependent parameter that we use to replace $K^{(x_i,\mathrm{C})}$ in this SDP at stage i and state (x_i, w_i).

Proposition 1. *The following properties hold:*

Figure 2: Information on Example 3.

(a) Production rewards and probabilities. (b) Value of operating at stage 1.



(a) There exist state-dependent shutdown cost inflations $\{\bar{K}_i^{(x_i,w_i,\mathrm{C})} \geqslant K^{(x_i,\mathrm{C})}, (i,x_i,w_i) \in \mathcal{I} \times \mathcal{X}_i \times \mathcal{W}_i\}$ *such that the* <u>*shutdown decisions*</u> *taken by* π^{D} *and* π^* *coincide.*

(b) Suppose for each $(i,x_i) \in \mathcal{I} \times \mathcal{X}_i$ *the set of exogenous states where* π^* *shuts down has strictly positive probability. Then there exist state-dependent shutdown cost modifications* $\{\bar{K}_i^{(x_i,w_i,\mathrm{C})} \in \mathbb{R}, (i,x_i,w_i) \in \mathcal{I} \times \mathcal{X}_i \times \mathcal{W}_i\}$ *such that* π^{D} *and* π^* *are equivalent.*

Part (a) establishes that inflating the shutdown cost in a state-dependent manner is sufficient to ensure that π^{D} takes exactly the same shutdown decisions as π^*, which implies that π^{D} and π^* have the same shutdown profile. When $\mathcal{X}_i = \{\mathrm{O}, \mathrm{C}\}$, that is production is the only non-shutdown state, this result implies that $\pi^{\mathrm{D}} \equiv \pi^*$. However, if the set of non-shutdown states is richer (i.e., $\{\mathrm{O}, \mathrm{C}\} \subset \mathcal{X}_i$), part (a) guarantees only partial recovery of π^* and it is possible to construct counter examples where there is no state-dependent shutdown cost modification such that π^{D} and π^* employ the same shutdown and non-shutdown decisions. In this case, Part (b) provides a full recovery result but requires non-zero likelihoods for shutdown at each future stage and allows shutdown cost modifications to either inflate or deflate the shutdown-neutral SDP plant closure costs. Our findings indicate that shutdown cost modifications provide a conceptual strategy to solve SC-MDP but identifying a good shutdown cost modification is not simple, as it entails computing a function that has the same dimension as the MDP value function.

3.2 Anticipated-regret SDP

We provide a low-dimensional and state-dependent shutdown cost modification that also has a social commerce motivation. The idea is to inflate the shutdown fixed cost by a larger amount in stages and states where there is a higher likelihood of this decision being suboptimal as future uncertainty is revealed. Such an inflation would discourage a shutdown decision being regretted by internal and external stakeholders as time progresses and is consistent with anticipated regret theory in behavioral psychology (Loomes and Sugden 1982, Zeelenberg 1999). According to this theory, when choosing among multiple decisions, one accounts for the anticipated regret from the chosen decision being worse than an alternative under realizations of future uncertainty. This regret has also been shown to be significantly higher for irreversible decisions (Tsiros and Mittal 2000), which supports our focus on shutdowns in our merchant production setting.

To formalize the above ideas, let $X(w_{i+1})$ denote a random variable defined at stage i as a function of the random stage $i+1$ information state w_{i+1}. Given a $\kappa \in \mathbb{R}_+$, we define anticipated regret of $X(w_{i+1})$ as $\mathbb{E}_i\left[\max\{X(w_{i+1}) + \kappa, 0\}\right]$. To build some intuition, suppose that κ represents the shutdown cost and $X(w_{i+1})$ is the random utility from not shutting down the plant at a stage i operating state. The term $X(w_{i+1}) + \kappa$ is then the excess of this utility over the shutdown cost κ, which is the positive regret from shutting down on this specific realization of uncertainty. The maximum of this term and zero captures the realizations of w_{i+1} where regret is strictly positive. Thus, anticipated regret is the expected regret from shutting down along sample paths where regret is positive.

Example 4. *Consider the same instance as in Example 3. At stage 1, let* $\kappa = 2$ *and define* $X(w_2)$ *as the sum of the known stage 1 and random stage 2 rewards. Then anticipated regret is zero in scenario A* $(= \max\{-5 + 3 + 2, 0\} = 0)$*, while it is positive in scenario B* $(= 0.5 \cdot \max\{-6 + 5 + 2, 0\} + 0.5 \cdot \max\{-6 - 6 + 2, 0\} = 0.5)$*. A decision maker that heavily weights anticipated regret would switch to a non-shutdown decision in scenario B but never switch in scenario A. Thus, inflating the shutdown cost using an anticipated regret term in the shutdown-neutral SDP will allow us to obtain the optimal policy of SC-MDP in this example.*

We use an anticipated regret term to modify the shutdown cost in the shutdown-neutral SDP (2). This term depends on the triple $\{\lambda, \xi, \eta\}$, which we abbreviate by Θ, and require that it satisfies $\lambda \geqslant 0$, $\xi \in (0, 1]$, and $\eta \geqslant 1$. We assume these parameters to be stage-independent but they can be made stage-dependent at the expense of having to tune more parameters. The AR SDP is

$$V^{\mathrm{A},\mathcal{O}}_{\Theta,i}(x_i, w_i) = \max_{a_i \in \mathcal{A}_i(x_i)\setminus\{\mathrm{C}\}} \Big\{ r_i(x_i, w_i, a_i) + \delta \mathbb{E}_i\big[V^{\mathrm{A}}_{\Theta,i+1}\left(f_i(x_i, a_i), w_{i+1}\right) \big] \Big\}, \quad \forall (i, x_i, w_i) \in \mathcal{I} \times \mathcal{X}_i \times \mathcal{W}_i, \tag{5a}$$

$$\mathrm{AR}_{\Theta,i}(x_i, w_i) = \max_{a_i \in \mathcal{A}_i(x_i)\setminus\{\mathrm{C}\}} \mathbb{E}_i\big[\max\{ r_i(x_i, w_i, a_i) + \delta V^{\mathrm{A}}_{\Theta,i+1}(f_i(x_i, a_i), w_{i+1}) + \eta \cdot K^{(x_i,\mathrm{C})}, 0\} \big], \tag{5b}$$

$$V^{\mathrm{A}}_{\Theta,i}(x_i, w_i) = \max\Big\{ V^{\mathrm{A},\mathcal{O}}_{\Theta,i}(x_i, w_i), -K^{(x_i,\mathrm{C})} - \lambda \xi^i \mathrm{AR}_{\Theta,i}(x_i, w_i) \Big\}, \ \forall (i, x_i, w_i) \in \mathcal{I} \times \mathcal{X}_i \times \mathcal{W}_i. \tag{5c}$$

The structure of the optimization used to compute $V^{\mathrm{A},\mathcal{O}}_{\Theta,i}(x_i, w_i)$ from $V^{\mathrm{A}}_{\Theta,i+1}$ in (5a) is identical to the one used in the shutdown-neutral setting, that is, (2b). However, comparing (5c) and (2c) shows that the shutdown cost is increased by $\lambda \xi^i \mathrm{AR}_{\Theta,i}(x_i, w_i)$ in the former case, where the term $\mathrm{AR}_{\Theta,i}(x_i, w_i)$ is based on a maximization of the anticipated regret across non-shutdown decisions.

Specifically, the anticipated regret of a non-shutdown action $a_i \in \mathcal{A}_i(x_i)\backslash\{\mathrm{C}\}$ is computed with respect to the sum of the immediate reward and the random discounted next stage value function, $r_i(x_i, w_i, a_i) + \delta V^{\mathrm{A}}_{\Theta,i+1}(f_i(x_i, a_i), w_{i+1})$, and the shutdown cost $K^{(x_i,\mathrm{C})}$ scaled by η. The role of $\mathrm{AR}_{\Theta,i}(x_i, w_i)$ can be interpreted as a way to dynamically increase the shutdown cost in a state-dependent fashion using the value function.

Parameters λ and ξ model the preferences to reduce and delay shutdowns, respectively. Increasing λ leads to a larger inflation of the shutdown cost due to the anticipated-regret term $\mathrm{AR}_{\Theta,i}$. Reducing ξ results in discounting $\mathrm{AR}_{\Theta,i}$ more heavily at later stages than at earlier stages, which amounts to modeling the preference to delay shutdown. The constant η controls the set of states where $\mathrm{AR}_{\Theta,i}(x_i, w_i)$ is positive and can thus cause a reversal of a shutdown decision. As η $(\geqslant 1)$ increases, the term $\mathrm{AR}_{\Theta,i}(x_i, w_i)$ will become positive at states where it was previously zero due to the shutdown cost $K^{(x_i,C)}$ in its definition being multiplied by η.

Proposition 2 establishes properties of the AR value function $V^{\mathrm{A}}_{\Theta,i}(x_i, w_i)$ and policy $\pi^{\mathrm{A}}_{\Theta}$.

Proposition 2. *(a) For any $\Theta = \{\lambda, \xi, \eta\}$ such that $\lambda \geqslant 0$, $\xi \in (0, 1]$, and $\eta \geqslant 1$, we have $V_i^{\mathrm{SN}\backslash\{\mathrm{C}\}}(x_i, w_i) \leqslant V^{\mathrm{A}}_{\Theta,i}(x_i, w_i) \leqslant V_i^{\mathrm{SN}}(x_i, w_i)$, $\forall (i, x_i, w_i) \in \mathcal{I} \times \mathcal{X}_i \times \mathcal{W}_i$;*

(b) If $\mathcal{W}_i$ is compact for each stage $i \in \mathcal{I} \cup \{I\}$, then for each $\xi \in (0, 1]$, it holds that $\lim_{(\lambda,\eta)\to\infty} \Pr(\mathrm{C}; \pi^{\mathrm{A}}_{\Theta}) = 0$ and $\lim_{(\lambda,\eta)\to\infty} V^{\mathrm{A}}_{\Theta,i}(x_i, w_i) = V_i^{\mathrm{SN}\backslash\{\mathrm{C}\}}(x_i, w_i)$.

Part (a) of Proposition 2 shows that the AR value function is bounded below and above, respectively, by the shutdown-neutral value functions that incorporate and exclude the shutdown action. Part (b) establishes that the shutdown probability associated with policy $\pi^{\mathrm{A}}_{\Theta}$ becomes zero as η and λ are increased to sufficiently large values. In addition, it shows that the value of the latter policy converges to the value of the policy $\pi^{\mathrm{SN}\backslash\{\mathrm{C}\}}$, which is optimal among the set of all policies with zero shutdown probability. Therefore, the anticipated-regret policy exhibits desirable asymptotic behavior with respect to both shutdown probability and asset value. (We assume a compact set of information states $\mathcal{W}_i$, $i \in \mathcal{I} \cup \{I\}$, in part (b) of Proposition 2 to avoid the technicalities that arise when dealing with unbounded distributions.) The asymptotic behavior of the AR policy with respect to ξ is intuitive. For $\xi = 1$, the policy is neutral to shutdown delay. As ξ decreases, the shutdown cost is inflated lesser at later stages compared to earlier stages. When $\xi \to 0$, there is no preference to delay shutdowns at stages other than stage 0, which translates to inflating only the shutdown cost at this initial stage.

3.3 Computing anticipated-regret policies

Optimal policies of the AR SDP (5) for different triples Θ provide a family of shutdown-averse policies. However, solving this SDP is challenging as it suffers from the curses of dimensionality typically associated with the shutdown-neutral SDP, which are a high-dimensional state space and potentially hard to compute expectations (Powell 2011). In addition, it embeds the anticipated regret term that contains more involved expectations. These curses of dimensionality indeed arise when considering the aluminum production application described in §5.1 as the exogenous state that we model is eight dimensional (see §5.2 for details on the price model). We consider the regress-later least squares Monte Carlo (LSML) method for computing heuristic policies based on low-dimensional approximations of the SDP value function (Glasserman and Yu 2004, Nadarajah et al. 2017). Specifically, we adapt an LSML approach to compute shutdown-averse policies by approximating the AR value function of SDP (5).

LSML computes a value function approximation (VFA) at each stage $i \in \{1, \dots, I\}$ that is a linear combination of a given set of B_i basis functions $\Phi_i := \{\Phi_{i,b}(w_i), b \in \mathcal{B}_i\}$, where $\mathcal{B}_i := \{1, \dots, B_i\}$. The VFA at stage i and state (x_i, w_i) is $\sum_{b\in\mathcal{B}_i} \beta_{i,x_i,b}\, \Phi_{i,b}(w_i)$, where $\beta_{i,x_i,b}$ denotes the b-th weight of this linear combination. Possible choices for basis functions include polynomials, radial basis functions, and Laguerre polynomials of the information state (Longstaff and Schwartz 2001, Mazières and Boogert 2013). At a high level, the LSML approach generates samples of the information state in Monte Carlo simulation and then combines backward recursion with regression to compute the VFA weights. Hard-to-compute expectations are replaced by sample average approximations in this procedure. For a given Θ, the output of LSML is the VFA weight vectors $\beta^{\mathrm{A}}_{i,x_i} := (\beta^{\mathrm{A}}_{i,x_i,1}, \dots, \beta^{\mathrm{A}}_{i,x_i,B_i})$ for each stage $i \in \{1, \dots, I\}$ and operating state $x_i \in \mathcal{X}_i$, which approximate $V^{\mathrm{A}}_{\Theta,i}(x_i, w_i)$ by $\sum_{b\in\mathcal{B}_i} \beta^{\mathrm{A}}_{i,x_i,b}\, \Phi_{i,b}(w_i)$. An AR policy decision can be computed at stage i and state (x_i, w_i) by first substituting the aforementioned VFA for $V^{\mathrm{A}}_{\Theta,i}(x_i, w_i)$ in the right hand side of the AR SDP (5) and then solving the maximization over actions. Online Supplement B contains the details of the LSML algorithm and describes its use for approximating the AR SDP (5) as well as the shutdown-neutral SDPs (2) and (4).

By varying our choices of λ and ξ in the triple Θ, LSML can be used to obtain a family of polices offering different trade-offs between asset value and shutdown profile in a tractable manner. We set a value for the parameter η in Θ based on Proposition 2(b), that is, we remove η as a policy parameter by fixing it before computing any AR policy. Specifically, we employ the following steps to choose a large enough value for this parameter so that increasing λ eventually leads to a zero-shutdown policy: (i) we compute the shutdown-neutral policy π^{SN}, (ii) we simulate the inflation term $\mathrm{AR}_{\Theta,i}$

in (5b) for different η values, and measure at each stage the percentage of states where this term is positive, and (iii) we select the smallest η value so that the coverage is 100% at each stage. Under this choice, the AR policy converges very closely to a zero-shutdown policy as λ is increased.

4. Operating policies based on production margins

In this section, we describe policies that determine operating decisions based on production margins. We first formalize a version of such policies used in practice and then improve on them by combining approximate dynamic programming and machine learning classification methods.

Stakeholders often consider poor production margins as a factor when making shutdown decisions. Production margin-based policies are appealing as they choose to shut down when production margins are below a certain threshold. For example, a myopic margin-based policy would (i) shut down the plant when the immediate production margin is less than the shutdown cost and (ii) continue to produce otherwise. More formally, assuming that the plant is producing at stage i and exogenous state w_i, a shut down decision is selected if $r_i(\mathrm{O}, w_i, \mathrm{O}) < -K^{(\mathrm{O},\mathrm{C})}$. The short-sighted nature of this policy can be overcome by modifying it to consider the sum of current and discounted future production margins. A forward-looking policy that considers T expected-future margins chooses to shut down at stage i and state (O, w_i) if

$$\hat{r}_i^T(w_i) := r_i(\mathrm{O}, w_i, \mathrm{O}) + \sum_{j=1}^{\min\{T, I-i\}} \delta^j \mathbb{E}_i \left[r_{i+j}(\mathrm{O}, w_{i+j}, \mathrm{O}) \right] < -K^{(\mathrm{O},\mathrm{C})}, \tag{6}$$

and continues to produce otherwise.

The aforementioned simple production margin policies use the fixed cost $K^{(\mathrm{O},\mathrm{C})}$ as the threshold to switch between full production and shutdown. This threshold choice is somewhat ad hoc and, in addition, does not incorporate the effect of shutdown profile constraints in (3b). Moreover, it is unclear how these policies can be extended to account for the full flexibility of the plant, that is, incorporate operating states other than full production. We address these issues next by presenting a more principled approach that first learns thresholds in an attempt to "mimic" the shutdown-neutral policy and then modifies them to account for the shutdown profile constraints.

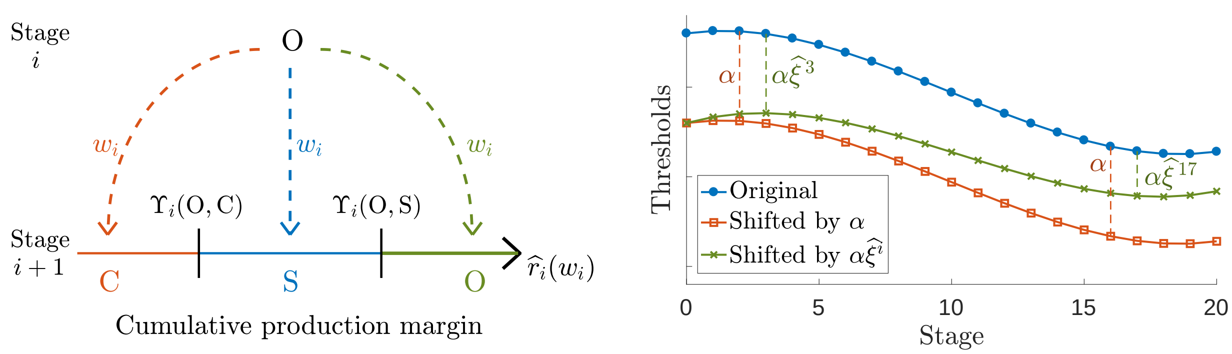
Suppose momentarily that the only operating state is full production O. For each stage $i \in \mathcal{I}$, we define a threshold $\Upsilon_i \in \mathbb{R}$ on the cumulative margin $\hat{r}_i^T(w_i)$ below which the decision switches from full production to shutdown. This threshold is computed as follows. First, we simulate the shutdown-neutral policy π^{SN} defined in §2.2 along sample paths $\{w_i^p, (i,p) \in \mathcal{I} \cup \{I\} \times \mathcal{P}\}$ of the uncertain information state generated in Monte Carlo simulation (π^{SN} can be approximated using LSML as discussed in §3.3 and Online Appendix B). Second, at each stage $i \in \mathcal{I}$, we compute for each of the samples $p \in \mathcal{P}$ the cumulative margin $\widehat{r}_i^T(w_i^p)$ and an action label from the shutdown-neutral policy π^{SN}, which is either to continue full production (i.e. O) or shutdown (i.e. C). Finally, based on these two labels, we partition the sample paths at each stage into two classes. The value of the threshold Υ_i is chosen as the cumulative margin value that best discriminates between (or separates) these two classes. Hence, identifying the shutdown threshold is equivalent to solving a binary classification problem where $\widehat{r}_i^T(w_i) \in \mathbb{R}$ is the explanatory variable and $x_{i+1} \in \{\mathrm{O}, \mathrm{C}\}$ is the outcome class (Bishop 2006). In general, it may not be possible to find a threshold that perfectly separates the "produce" and "shutdown" classes, in which case, it is standard to minimize misclassification error to compute the threshold, where error is measured using a loss function. Common choices in machine learning include hinge, squared, and logarithmic loss functions (Rosasco et al. 2004).

The above classification procedure can be extended to handle multiple operating states. To ease exposition, we focus on the case of two operating states, that is, $\mathcal{O} := \{\mathrm{O}, \mathrm{S}\}$, where S denotes a temporary suspension of production. Suppose the plant is open at stage i ($x_i = \mathrm{O}$) and can transition to state $x_{i+1} \in \{\mathrm{O}, \mathrm{S}, \mathrm{C}\}$. We simulate the shutdown-neutral policy π^{SN} at stage i over multiple sample paths of the uncertain information state $p \in \mathcal{P}$ and compute the cumulative margin $\widehat{r}_i^T(w_i^p)$ as before, but divide the sample paths using three action labels corresponding to the shutdown-neutral policy choosing to continue full production O, suspend production S, and shutdown C. Implementing a production margin policy entails finding two thresholds $\Upsilon_i(\mathrm{O}, \mathrm{S})$ and $\Upsilon_i(\mathrm{O}, \mathrm{C})$ to determine when the plant switches from full production to suspension and shutdown, respectively. We proceed to define these thresholds by converting the multi-class classification problem over the three action labels in set $\{\mathrm{O}, \mathrm{S}, \mathrm{C}\}$ into several binary classification problems between pairs of such classes, which are $\{\mathrm{O}, \mathrm{S}\}$, $\{\mathrm{S}, \mathrm{C}\}$, and $\{\mathrm{O}, \mathrm{C}\}$. This is a well-known strategy in machine learning (see, e.g., Bishop 2006, ch. 4). Applying a binary classification procedure to each pair would result in three different thresholds, that is, one more than the required number of thresholds. We resolve this issue by assuming an adjacency structure over the cumulative margins associated with action labels as illustrated in Figure 3(a). Specifically, cumulative margins associated with the action labels O and S as well as S and C are assumed adjacent, while those associated with O and C are non-adjacent. In other words, the action labels have a weak ordering with respect to the cumulative margin. Then applying binary classification to $\{\mathrm{O}, \mathrm{S}\}$ and $\{\mathrm{S}, \mathrm{C}\}$ gives us the thresholds $\Upsilon_i(\mathrm{O}, \mathrm{S})$ and $\Upsilon_i(\mathrm{O}, \mathrm{C})$, respectively.

Finally, we modify the shutdown thresholds $\{\Upsilon_i(\mathrm{O}, \mathrm{C}), \Upsilon_i(\mathrm{S}, \mathrm{C}), i \in \mathcal{I}\}$, which mimic the shutdown-neutral policy, to account for shutdown profile preferences as illustrated in Figure 3(b). Decreasing the value of these thresholds (blue curve) by a constant at each stage results in a set of downward-

Figure 3: Illustration of threshold computation using binary classification with states in set $\{\mathrm{O}, \mathrm{S}, \mathrm{C}\}$ and subsequent modification of thresholds to model shutdown profile preferences.

(a) Classification using cumulative margin thresholds. (b) Threshold modification using α and $\widehat{\xi}$.



shifted thresholds (orange curve), which will result in a policy with a smaller shutdown probability. In other words, the threshold levels at each stage are proportional to the shutdown probability at that stage when using a production margin policy, thus providing a lever to shape the shutdown profile. As a consequence, preferences to delay shutdown decisions can be emphasized by decreasing thresholds by a larger value at earlier stages in the horizon and less aggressively at later periods, which involves modifying the original shutdown probability thresholds (blue curve) and obtaining a shifted and tilted set of thresholds (green curve). Formally, given a "shift" parameter $\alpha \geqslant 0$ and a "tilt" parameter $\widehat{\xi} \in (0, 1]$, we define the modified shutdown thresholds as follows:

$$\Upsilon_i^{\alpha,\widehat{\xi}}(\mathrm{O}, \mathrm{C}) := \Upsilon_i(\mathrm{O}, \mathrm{C}) - \alpha\widehat{\xi}^i, \quad \forall i \in \mathcal{I},$$
$$\Upsilon_i^{\alpha,\widehat{\xi}}(\mathrm{S}, \mathrm{C}) := \Upsilon_i(\mathrm{S}, \mathrm{C}) - \alpha\widehat{\xi}^i, \quad \forall i \in \mathcal{I},$$

and leave unchanged the non-shutdown thresholds, e.g. $\Upsilon_i(\mathrm{O}, \mathrm{S})$. The parameter α represents the magnitude of a downward-shift in the thresholds related to shutdown actions, while $\widehat{\xi}$ tilts the threshold curve to account for shutdown delay. We vary the values of α and $\widehat{\xi}$ in Monte Carlo simulation to obtain a family of policies with different shutdown profile preferences. We first fix $\widehat{\xi}$ equal to a constant, and then perform a line search on α to identify its smallest value for which the shutdown probability of the corresponding production margin policy satisfies the bound U_I in SC-MDP (3b). This modification of $\widehat{\xi}$ implies values for the U_i bounds at intermediate stages. Modifications to this procedure are possible, for instance, one could do a grid search over both α and $\widehat{\xi}$, if needed.

Overall, we have generalized practice-based production margin policies to use thresholds computed in a principled manner while accounting for shutdown profile preferences and maintaining their intuitive margin structure. The ideas discussed above can be extended to multiple operating states by (i) assuming a weak ordering on operating states, and (ii) computing thresholds between adjacent states alone.

5. Numerical study

We next numerically evaluate the performance of our methods. In §5.1, we introduce a case study of a real aluminum producer, which serves as our application for this evaluation. In §5.2, we describe the stochastic process used for modeling the evolution of uncertainty. In §5.3, we define the aluminum production instances used for our experiments and the associated computational setup. In §§5.4-5.6, we discuss our findings related to methodological performance, the role of different social commerce objectives, and the robustness of insights to parameter changes.

5.1 Aluminum production case study

Our case study is based on a real aluminum producer. Shutting down an aluminum production plant (often referred to as a smelter) or temporarily suspending its production are strategic decisions that are re-evaluated on an annual basis. Assessing the shutdown profile thus requires planning over a long time horizon. Our instance considers a forty-year horizon (i.e., $I = 40$) where each stage i corresponds to a year and decisions are made from $i = 0$ (present) to $I - 1 = 39$.

Aluminum production relies on an energy intensive electrolysis process that takes as inputs alumina from bauxite, carbon, and electricity, and produces aluminum as its main output. The aluminum producer is vertically integrated and owns bauxite mines and carbon plants. Therefore, capturing the uncertainty in alumina and carbon prices is not critical, whereas modeling the volatile electricity and aluminum prices is important. In addition to price risk, the producer faces exchange rate risk because aluminum is sold globally in US dollars (USD), electricity is purchased in Euro (EUR) from the Nord Pool electricity market, and the operating costs are incurred in Norwegian Krone (NOK), that is, the local currency. Assuming cash flows are measured in USD, the production spread is exposed to volatile EUR-USD and NOK-USD exchange rates. Thus, we model four sources of uncertainty in the set $\{P_i^{\text{AL}}, P_i^{\text{EL}}, P_i^{\text{EUR-USD}}, P_i^{\text{NOK-USD}}\}$ denoting, respectively, aluminum price, electricity price, and the two exchange rates. We capture the dynamics of prices and exchange rates using eight stochastic factors w_i discussed in §5.2.

The specific aluminum plant we model has a shutdown state (C) and open states corresponding to (i) production at full load (O), which we normalize to an output of 1 metric tonne of aluminum; and (ii) temporary suspension. We denote by S_m the m-th consecutive year of suspension, where m is restricted to at most 3, that is, $m \in \{1, 2, 3\}$. Intermediate production capacity options are not modeled in this case since the electrolysis cells within a smelter operate in a near continuous fashion, that is, these cells can only be turned off for a few hours. A prolonged shut off results in cell damage and expensive repairs before a restart (see, e.g., Øye and Sørlie 2011). Thus, $\mathcal{O}_i = \{\text{O}, \text{S}_1, \text{S}_2, \text{S}_3\}$

for each $i \in \mathcal{I}\backslash\{0\}$. We assume that the plant must be either in production or shutdown at the end of the planning horizon.

Figure 4: State-action set, reward function, and production state transitions in the aluminum production case study.

State $[x_i]$	Decision $[a_i]$	Reward $[r_i(x_i, w_i, a_i)]$
O	O S_1 C	$r_{\mathrm{O}}(w_i)$ $-K^{(\mathrm{O},\mathrm{S}_1)}$ $-K^{(\mathrm{O},\mathrm{C})}$
$\mathrm{S}_m,\ m \in \{1,2,3\}$	O S_{m+1} (if $m < 3$) C	$r_{\mathrm{O}}(w_i) - K^{(\mathrm{S}_m,\mathrm{O})}$ $-K^{(\mathrm{S}_m,\mathrm{S}_{m+1})}$ $-K^{(\mathrm{S}_m,\mathrm{C})}$
C	C	0

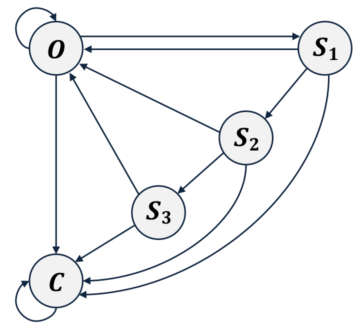


Figure 4 illustrates the set of feasible actions $\mathcal{A}_i(x_i)$ at a given operating state $x_i \in \mathcal{X}_i$, the resulting reward, and a diagram indicating the next stage operating state. Here $K^{(x_i,x_{i+1})}$ represents the fixed operating and/or maintenance cost associated with a state transition from x_i to x_{i+1}. The function $r_{\mathrm{O}}(w_i)$ denotes the profit from producing aluminum and is defined as

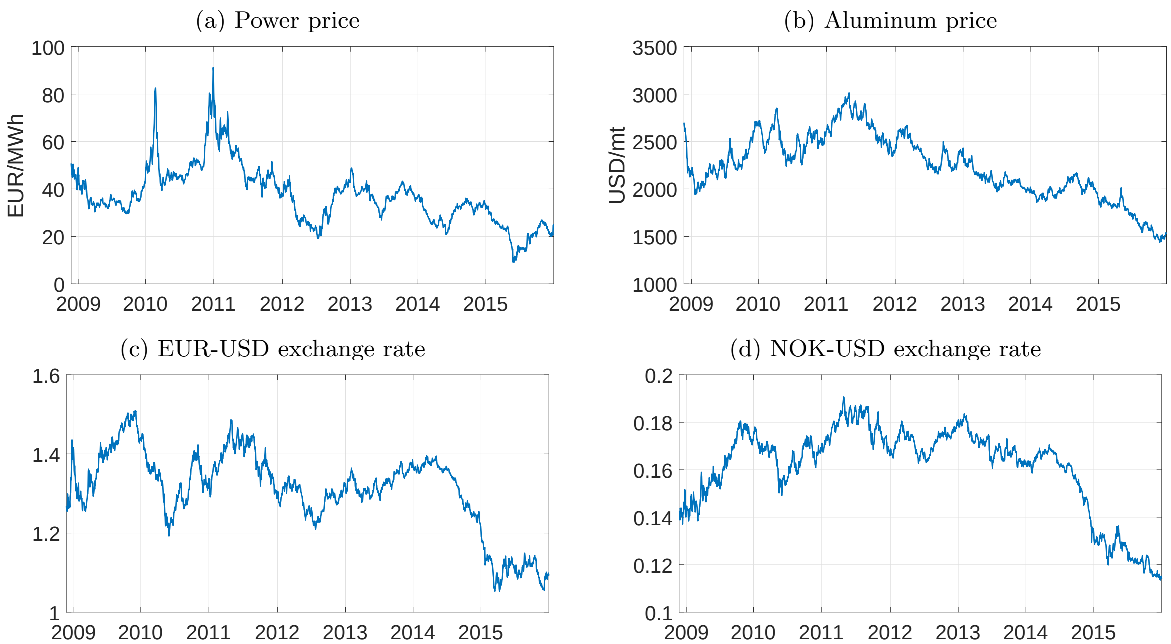
$$r_{\mathrm{O}}(w_i) := (1-\tau)\left[P_i^{\mathrm{AL}}(1+\gamma^{\mathrm{AL}}) - c^{\mathrm{USD}} - c^{\mathrm{NOK}}\, P_i^{\text{NOK-USD}} - \rho\, P_i^{\mathrm{EL}} P_i^{\text{EUR-USD}}\right], \tag{7}$$

where τ is the corporate tax rate; γ^{AL} the aluminum premium, which is the surcharge a buyer pays for taking delivery of aluminum; c^{USD} and c^{NOK} the production fixed costs incurred in USD and NOK, respectively; and ρ the rate of electricity consumption in the aluminum production process. We have fixed costs in both USD and NOK because they include components of the electrolysis and casting costs in each of these currencies.

5.2 Model of price and exchange rate dynamics

As discussed in §5.1, the exogenous market information in our SDP consists of the factors driving four sources of uncertainty: aluminum price, electricity price, and two exchange rates (EUR-USD and NOK-USD). A stochastic process to model the evolution of these sources of uncertainty can be calibrated using information from financially traded contracts. Futures contracts on (primary) aluminum are traded at the LME in US dollars with monthly maturities going out to 10 years. The aluminum contract with a 3-month maturity is the most liquid. Currency futures are traded in the FOREX market with monthly maturities up to 3 years. In contrast to aluminum and currencies, power is traded in regional markets. In the Nord Pool market, which covers Scandinavia and part of northern Europe, power forward contracts are denominated in Euro and extend out to 10 years. Negative prices present in the electricity spot markets are not a feature seen in these longer term contract prices. Moreover, none of the aforementioned contracts have maturities that cover our planning horizon of 40 years.

Figure 5: Front month futures prices for power and aluminum and exchange rates for EUR-USD and NOK-USD between 2009 and 2015.



We capture the dynamics of the four sources of uncertainty using an eight-factor continuous-time price model, where each source is driven by short term and long term factors. All the eight factors are correlated. This choice addresses the lack of traded contracts for the entire planning horizon and captures the clear short and long term trends observed in market data, which we illustrate in Figure 5 using only the front month contracts from the broader data set used for calibration. In particular, we denote the short and long term factors by the vectors $Y(t) = \left[Y_t^{\text{EL}}, Y_t^{\text{AL}}, Y_t^{\text{EUR-USD}}, Y_t^{\text{NOK-USD}}\right]$ and $Z(t) = \left[Z_t^{\text{EL}}, Z_t^{\text{AL}}, Z_t^{\text{EUR-USD}}, Z_t^{\text{NOK-USD}}\right]$, respectively. These factors evolve according to the following stochastic differential equations:

$$d\begin{bmatrix} Y(t) \\ Z(t) \end{bmatrix} = \begin{bmatrix} -K_y(Y(t) - Z(t)) \\ \mu_z - K_z Z(t) \end{bmatrix} dt + \begin{bmatrix} \Sigma_y^{1/2} & 0 \\ \Sigma_{yz}^{1/2} & \Sigma_z^{1/2} \end{bmatrix} d\begin{bmatrix} W_y(t) \\ W_z(t) \end{bmatrix}, \tag{8}$$

where K_y and K_z represent the speeds of mean reversion and are assumed to be diagonal (4×4) matrices, μ_z is the long term drift, $\Sigma_y^{1/2}$, $\Sigma_z^{1/2}$, and $\Sigma_{yz}^{1/2}$ are diffusion matrices, and W_y and W_z are each independent four-dimensional standard Brownian motions. The short term factors in this model revert to the long term factors. Moreover, the short term factors are related to the prices and exchange rates by $P_t^k = \exp\left(Y_t^k - \psi^k(t)\right)$, for $k \in \mathcal{J}$, where $\psi^{(\cdot)}(t)$ is a function that captures seasonality (see Online Supplement C for details). Our model can be seen as a multi-commodity extension of the one in Schwartz and Smith (2000), where we added a cross-commodity correlation structure, and a special case of the process in Farkas et al. (2017). Under this specification, the information state is $w_i = (Y_{i,k}, k \in \mathcal{J}) \cup (Z_{i,k}, k \in \mathcal{J})$, where $Y_{i,k}$ and $Z_{i,k}$, respectively, denote the k-th short and long term factor values at the beginning of stage i. Moreover, $\mathcal{W}_i = \mathbb{R}^8_+$.

Calibrating our multi-asset and multi-factor stochastic process requires several steps (see Online Supplement C for details). We collected futures contract data with multiple maturities for aluminum, power, and exchange rates from the LME, Nord Pool, and FOREX markets, respectively. We used interpolation and spline smoothing to ensure that data across different sources were consistent. We then employed a multi-stage Kalman filter process to estimate parameters. Our calibrated model provides a statistically sound representation of market data and allowed us to generate sample paths of the sources of uncertainty in Monte Carlo simulation needed to implement our algorithms as well as estimate the asset value and shutdown profile of a policy.

5.3 Instances and computational setup

We define our reference instance using the stochastic process calibration described in §5.2 and real operational data from an aluminum producer, which includes parameter values of the reward function (7), the operations and maintenance costs, and the discount factor. We summarize this information in Table 1. We define the terminal reward function at a production state by extending the problem horizon beyond the actual planning horizon, which is standard. Specifically, we choose $r_I(\mathrm{O}, w_I)$ to represent the value of a plant that operates for 20 years starting from stage I and state (O, w_I) using an LSM approach. We employ the same terminal condition across all methods for consistency.

Table 1: Parameters defining the reference aluminum production instance.

Name	Value	Unit	Name	Value	Unit
τ	25%	-	$K^{(\mathrm{O},\mathrm{S}_1)}$	600	USD/mt
γ^{AL}	5%	-	$K^{(\mathrm{O},\mathrm{C})}$	1200	USD/mt
c^{USD}	520	USD/mt	$K^{(\mathrm{S}_m,\mathrm{C})}$	600	USD/mt
c^{NOK}	6110	NOK/mt	$K^{(\mathrm{S}_m,\mathrm{O})}$	600	USD/mt
ρ	14	MWh/mt	$K^{(\mathrm{S}_m,\mathrm{S}_{m+1})}$	0	USD/mt
δ	0.971	-			

By modifying the reference instance described above, we created an extended instance set to study the performance of our methods when operational and market parameters change. We increased/decreased the costs c^{USD} and c^{NOK}, which are the only two parameters in the reward function (7) that are directly related to plant operations. We changed these costs by $\pm 15\%$ to obtain instances OP1-OP4 in Table 2 because these perturbations resulted in asset value changes of $\pm 40\%$, which seem reasonable upper bounds on the changes that can be expected in practice. We also considered significantly changing the volatility estimates of both the short and long term factors for aluminum and power by $\pm 30\%$ to obtain instances MP1-MP4 in Table 2. Power price volatilities vary by region while LME is the primary market for assessing aluminum volatility. Our aluminum price volatility change is consistent with historic data: Brunetti and Gilbert (1995) examined the monthly volatility of LME-traded metals, including aluminum, over a 24-year period and $\pm 30\%$ is representative of the maximum variations they observe.

We next describe our computational setup. To compute AR (i.e., anticipated-regret) policies, we used LSML as described in §3. We also used LSML to compute heuristic versions of the shutdown-neutral policy π^{SN} and the zero-shutdown policy $\pi^{\mathrm{SN}\setminus\{\mathrm{C}\}}$ which include and exclude the shutdown option, respectively. We find the switching thresholds of the classification based margin policies (i.e., CM policies) while computing the LSML VFA for π^{SN}. We chose the LSML basis functions to be the following set of third-degree polynomials of the eight stochastic factors discussed in §5.2:

$$\{\Phi_{i,b}(w_i)\}_{b=1}^{B_i} = \{1,\, Y_{i,k},\, Z_{i,k},\, Y_{i,k_1}Y_{i,k_2},\, Z_{i,k_1}Z_{i,k_2},\, Y_{i,k}Z_{i,k},\\ Y_{i,k_1}Y_{i,k_2}Y_{i,k_3},\, Z_{i,k_1}Z_{i,k_2}Z_{i,k_3},\, Y_{i,k}Z_{i,k}^2,\, Y_{i,k}^2 Z_{i,k} \,\big|\, k,\, k_1,\, k_2,\, k_3 \in \mathcal{J}\}.$$

Based on experimentation, we chose the number of regression samples and inner samples in LSML (see Algorithm 1 in Online Supplement B) equal to 20,000 and 200, respectively. We estimated the asset value under each policy in Monte Carlo simulation using 20,000 sample paths. For consistency, we ensured that the same sample paths were used across policy evaluations. We also estimated (dual) upper bounds in Monte Carlo simulation using the LSML VFA (see Online Supplement D for details). We implemented the binary classification scheme needed to obtain CM policies with both hinge and squared loss functions, and found them to perform almost identically. We thus only report results related to the squared loss case.

We implemented all the algorithms using Matlab R2016b and executed them on a server equipped with two Intel Xeon E5-2660v3 processors with 10 cores each and a shared memory of 128 GB RAM. Obtaining families of policies offering different asset value and shutdown profile trade-

Table 2: Changes in operational and market parameters of the reference instance to obtain the extended instance set.

Label	Instance description	Label	Instance description
OP1	Cost c^{USD} is raised by 15%	MP1	Power volatility is raised by 30%
OP2	Cost c^{USD} is reduced by 15%	MP2	Power volatility is reduced by 30%
OP3	Cost c^{NOK} is raised by 15%	MP3	Aluminum volatility is raised by 30%
OP4	Cost c^{NOK} is reduced by 15%	MP4	Aluminum volatility is reduced by 30%

offs required significant computation. We distributed the computational load across the 20 cores in our server to significantly reduce the run time to obtain trade-off curves. Specifically, when using 20,000 sample paths on a single core, estimating the LSML VFA weights consumed about 18 minutes for each of the two shutdown-neutral policies and 22 minutes on average for the AR policy for each value of λ and ξ. Computing the asset value estimate of one of these policies required roughly the same running time as estimating the VFA, and estimating a dual bound took about 16 minutes on average. Computing the expected margins in the production margin-based heuristics used in practice and our extensions thereof required 4 minutes. Fitting the CM thresholds (including the LSML run) as discussed in §4 entailed an additional 18 minutes of CPU time. Updating these thresholds for a given U_I and $\widehat{\xi}$ took less than 10 seconds on average.

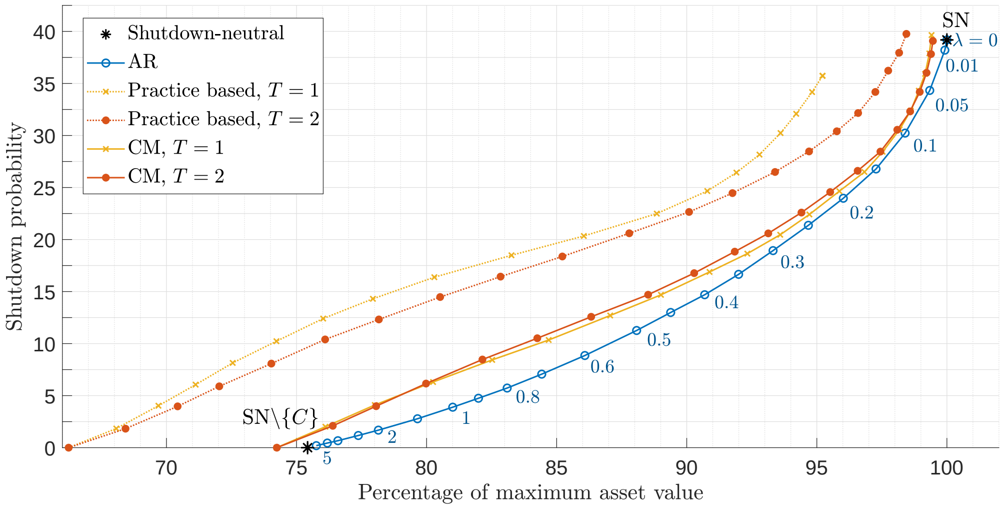
5.4 Shutdown probability and asset value trade-off

We consider the reference instance described in §5.3. The asset value estimates for the policies π^{SN} and $\pi^{\mathrm{SN}\setminus\{\mathrm{C}\}}$ are within 2.0% and 2.6% of their respective dual bound estimates (standard errors of asset value and dual bound estimates are at most 1.23% and 0.10%, respectively). That is, our LSML shutdown-neutral policies are near optimal. The asset value estimate of π^{SN} equals 4902 USD/mt and will henceforth be referred to as maximum asset value. The π^{SN} shutdown probability is 39.2%.

We assess the performance of different policies along the shutdown probability dimension, that is, we temporarily neglect the preference to delay shutdown by fixing $\xi = 1$ in LSML and $\widehat{\xi} = 1$ in the threshold shifting procedure at the end of §4. Figure 6 displays the trade-offs between asset value and shutdown probability for the policies discussed in §§3-4. Each asset value estimate is expressed as a percentage of the maximum asset value (standard errors of the asset value and shutdown probability are at most 1.24% and 0.35%, respectively), which is the π^{SN} asset value denoted by an asterisk in Figure 6 and labeled SN. The AR policies were computed by varying λ from 0 to 10 to obtain policies that range between the shutdown-neutral policy and the zero-shutdown policy.

We find that the AR approach performs well and achieves substantial shutdown probability reductions for small asset value losses. For asset value losses of 2% and 4%, the shutdown probability is reduced by 26% and 40%, respectively, which are significant reductions considering that π^{SN} closes the plant on roughly 39% of the scenarios. Consistent with Proposition 2, the AR policies converge to $\pi^{\mathrm{SN}\setminus\{\mathrm{C}\}}$, which is the optimal zero-shutdown asset value, when λ is large enough and η is chosen as discussed at the end of §3.3. Indeed, avoiding shutdowns altogether leads to significant asset value loss. Specifically, the asset value estimate of $\pi^{\mathrm{SN}\setminus\{\mathrm{C}\}}$ is only 75.4% of the maximum asset value (i.e., the asterisk labeled SN\{C} in this figure). For a typical plant with capacity of 200,000 mt, a 10% loss in asset value translates to roughly 100 million USD. Thus, shutdown flexibility

Figure 6: Trade-offs in asset value and shutdown probability.



has substantial value but its use can be limited without significant value loss, which bodes well for social commerce, that is, unaccounted social costs amounting to a small fraction of the maximum asset value can justify using operating policies that reduce the likelihood of future shutdowns.

The production margin-based policies were tested using one or two expected-future margins, in addition to the current margin, which entails choosing $T \in \{1, 2\}$. The shutdown probabilities U_i, $i \in \mathcal{I} \cup \{I\}$, were set to the same value U which was varied from 40% to 0%. The best asset value achieved by the practice-based policies (dotted lines) when T equals 2 is only around 1.3% less than the maximum asset value, but it is 4.8% less when T equals 1. Both practice-based policies decrease shutdown probability in an inefficient manner. We instead find that the CM method outperforms the practice-based methods and is more robust, that is, the asset value and shutdown probability trade-off curves produced by CM for T equal to one and two are almost identical. The performance of CM is comparable to AR for small asset value losses but slightly worse for larger losses, which suggests that the favorable asymptotic properties of AR policies translate into good performance of these policies in non-asymptotic settings as well. For example, the shutdown probability is reduced by 25% in the CM and AR polices for asset value losses of 2.2% and 1.8%, respectively. We focus on the AR and CM policies in the rest of the numerical analysis as their performance is similar for small asset value losses but outperform the practice-based heuristics.

5.5 Evaluation of social commerce objectives

We begin by comparing the social commerce preferences to reduce shutdown probability and delay shutdown decisions. One may expect that delaying shutdown decisions also reduces the shutdown probability at the end of the horizon. However, this intuition may not be true because delaying/reducing shutdowns also affects asset value. Therefore, in Figure 7 we display the trade-off between shutdown probability reduction and shutdown delay for fixed asset value losses. Shutdown probability reduction (x-axis), expected shutdown delay (y-axis), and asset value loss (legend) are measured relative to the respective values of 39.2%, 26 years, and 4902 \$/mt, respectively, of π^{SN}. To obtain this figure, we considered $\xi \in \{1, 0.99, 0.98, 0.97, 0.95\}$ for AR policies and $\widehat{\xi} \in \{1, 0.98, 0.95, 0.93, 0.9\}$ for CM policies with T equal to 2. For each ξ and $\widehat{\xi}$, we computed profiles akin to the ones in Figure 6 so that we could choose AR and CM policies that matched a target asset value within 0.3%.

Figure 7: Trade-offs in shutdown probability decrease and shutdown delay for AR and CM policies.

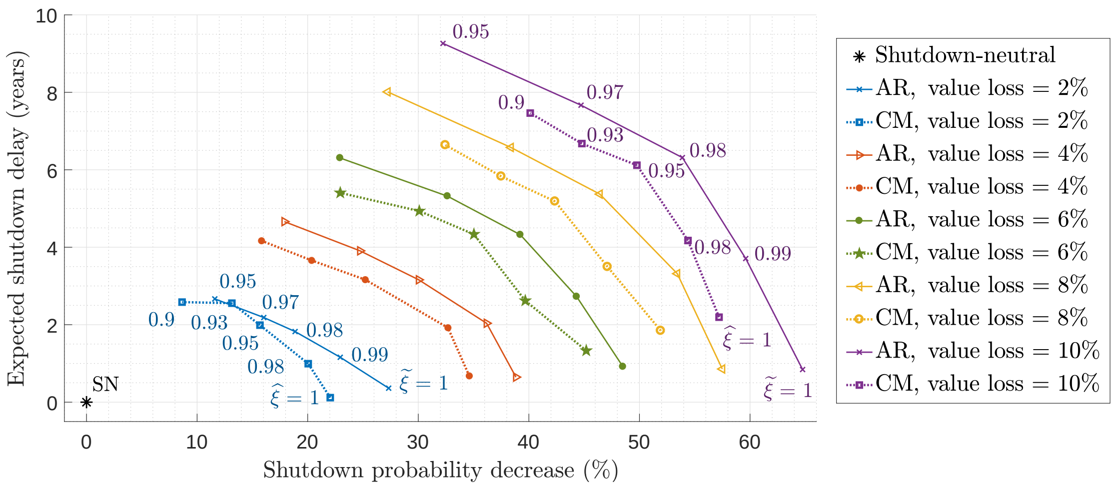


Figure 7 suggests that shutdown delay and shutdown reduction are substitute social commerce preferences competing for a fixed asset value loss and thus cannot be improved simultaneously. For example, for an asset value loss of 6%, AR could delay shutdowns by 6 years on average and decrease the shutdown probability of 23%, or delay shutdowns by an average of only 3 years and capture a larger shutdown probability decrease of 44%. To understand this behavior, recall that requiring larger values for either expected shutdown delay or shutdown probability reduction entails an asset value loss relative to π^{SN}. Thus, for a given budget of asset value loss, increasing expected shutdown delay amounts to allocating a larger fraction of this loss budget to the shutdown delay preference, which then makes the remaining budget smaller and only allows for a lower shutdown probability reduction. Furthermore, the concavity of most trade-off curves in Figure 7 suggests diminishing returns: higher average shutdown delays imply a higher marginal increment in shutdown probability, and vice versa. The trade-offs originating from AR slightly dominate the CM ones, that is, for the same asset value and shutdown probability, AR policies delay shutdown by roughly one additional year on average relative to the CM policies. As one would expect, it appears harder to manage the preference to reduce and to delay shutdown decisions substantially for smaller asset value losses.

Both social commerce preferences discussed above are based on modifying shutdown probabilities. An alternative, as discussed in §2.3, would be to provide a deterministic commitment (DC) to not shutdown the plant for a predefined number of years. Next, we investigate when DCs are practical. Policies encoding DCs can be computed analogously to π^{SN} using LSML but removing the shutdown action for the first T years. We consider values for T in set $\{2, 5, 6, 8, 10\}$. We compare a DC policy for a given T with an AR policy that has the same expected time to shutdown, that is, these policies differ only in asset value and shutdown probability. Indeed, under favorable price conditions, DC and AR policies are similar because the likelihood to shutdown is small in the short-term. This is the case for our baseline instance where the starting price for aluminum is 1800 USD/mt.

LME aluminum prices, however, have been below 1400 USD/mt many times in the last 20 years (e.g., 1291 USD/mt in 2003 and 1330 USD/mt in 2009). For an initial aluminum price of 1400 USD/mt, shutdown in the first few years is more likely and establishing agreements not to shut down is especially relevant. Table 3 compares DC and AR policies for an initial aluminum price of 1400 USD/mt. The shutdown-neutral asset value in this instance is equal to 559 USD/mt. DC policies lose significantly more asset value than AR policies to achieve the same shutdown delay, especially when the commitment exceeds 5 years. AR policies also outperform DC policies along the shutdown probability dimension. This difference is due to DC policies having to avoid shutdown in all possible scenarios in the first T periods, including those where the plant would incur large losses, while AR policies have the flexibility to shutdown in such financially adverse scenarios. Figure 8 visually illustrates this effect by displaying the asset value distribution histograms for DC and AR policies restricted to sample paths where at least one of these policies shuts down. Clearly, the AR asset value distribution has considerable mass to the right of the DC asset value distribution. In sum, DCs of a few years are practical as they provide certainty regarding no shutdowns but extending such guarantees for several years can lead to major asset value losses. Thus, a combination of DCs and bounds on the shutdown probability appears necessary to balance social and financial objectives of the plant.

Table 3: Comparison of DC and AR policies with the same shutdown delay.

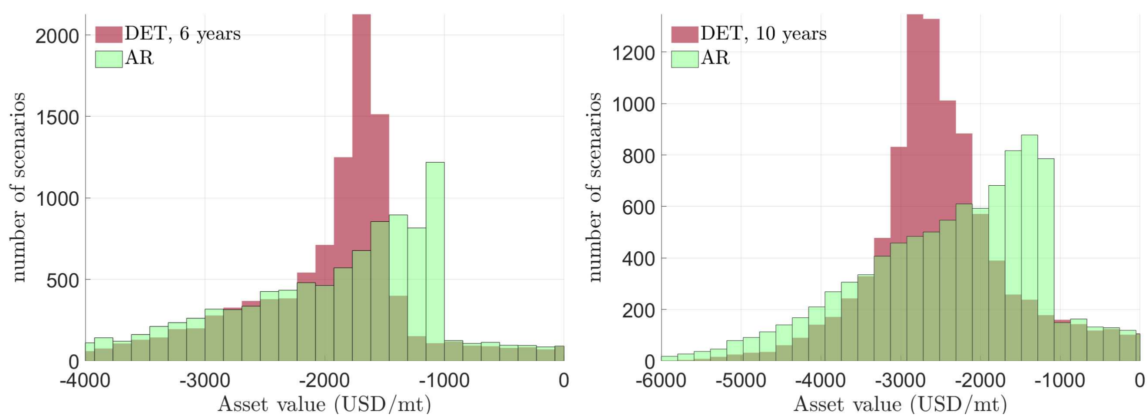
		Commitment of T years				
	Policy	2	5	6	8	10
% asset value	DC	100.0	98.5	94.9	83.0	70.1
	AR	100.0	99.7	97.6	86.0	78.6
Shutdown %	DC	53.0	53.5	51.8	50.1	48.6
	AR	53.0	51.9	50.4	46.1	44.3

5.6 Comparative statics and robustness

We finally consider the performance of AR and CM on the extended instance set, that is, OP1-OP4 and MP1-MP4 defined in Table 2 by varying operational costs and market volatilities of the

Figure 8: Asset value distributions associated with DC and AR policies for T equal to six and ten.

(a) Six year deterministic commitment. (b) Ten year deterministic commitment.



reference instance (labeled REF).

Table 4 reports results that examine the asset value and shutdown probability trade-off for policies with $\xi = \widehat{\xi} = 1$ similar to Figure 6. Specifically, for decreases in shutdown probability equal to 10%, 20%, 30%, and 40%, this table shows the asset value loss expressed as a percentage of the maximum asset value. We find that AR and CM incur similar asset value losses to achieve a target shutdown probability reduction on the extended instance set, with the former method performing slightly better than the latter method. For example, both methods reduce shutdown probability by 10% and 20% in most instances for asset value losses, respectively, below 1% and 2%. A larger shutdown probability reduction of 40% entails average losses of 4.7% and 5.7% when using AR and CM, respectively. Changing the fixed cost c^{NOK} leads to larger fluctuation than c^{USD} in the asset value loss incurred to achieve a shutdown probability reduction, and the largest across the extended instance set. For example, reducing the shutdown probability by 40% requires

Table 4: Asset value loss as a percentage of the maximum asset value for 10%, 20%, 30%, and 40% decreases in shutdown probability when using AR and CM with T equal to 2.

	Shutdown probability decrease							
	10%		20%		30%		40%	
Instance	AR	CM	AR	CM	AR	CM	AR	CM
REF	0.7	0.9	1.1	1.6	2.7	3.0	4.0	5.0
OP1	1.0	1.3	1.6	2.8	4.4	5.3	6.4	8.3
OP2	0.5	0.5	0.8	1.0	1.9	2.1	2.7	3.2
OP3	1.2	1.7	3.0	3.9	5.3	7.4	9.3	12.3
OP4	0.2	0.4	0.6	0.6	1.2	1.2	1.8	2.0
MP1	0.7	1.0	1.8	1.9	3.1	3.9	5.2	6.2
MP2	0.3	0.7	1.0	1.4	2.0	2.5	3.7	4.2
MP3	0.2	0.7	0.8	1.2	1.7	2.0	2.6	3.2
MP4	0.8	1.2	1.9	2.2	3.3	4.0	6.2	7.3
Average	0.6	0.9	1.4	1.9	2.8	3.5	4.7	5.7

a roughly 9–12% asset value loss in OP3 while the analogous loss is less than 2% for OP4. This difference in impact can be attributed to c^{NOK} appearing in the reward function multiplying the exchange rate $P_i^{\text{NOK-USD}}$, which is random, while c^{USD} features on its own. The higher impact of the cost c^{NOK} is consistent with the industry focus on local costs. Changing the power (an input) price and aluminum (an output) price volatilities has opposite effects. Specifically, the asset value loss to achieve a target shutdown-probability reduction becomes smaller as power and aluminum volatilities, respectively, decrease (compare MP2 with MP1) and increase (compare MP3 with MP4). The observed behavior is consistent with the distribution of aluminum and power prices becoming more skewed as volatilities increases such that larger price values are more likely. Since aluminum is an output, such skewness results in more profit, while the same effect on power prices increases cost and reduces profit. This relationship between skewness and volatility is known for stochastic processes with log-normally distributed factors.

Table 5 analyzes both shutdown reduction and shutdown delay on the extended instance set. For fixed asset value losses, we determined trade-off curves for AR and CM analogous to those displayed in Figure 7. In particular, we consider the two extreme ends of each trade-off curve, which correspond to policies computed with (i) $\xi = \widehat{\xi} = 1$, and (ii) $\xi = 0.95$ for AR and $\widehat{\xi} = 0.90$ for CM (as in Figure 7). The shutdown probability reduction is expressed as a percentage of the shutdown-neutral probability computed for the same instance and the average shutdown delay is reported in number of years and displayed within parenthesis. Both AR and CM emerge as robust methods to balance the asset value and the shutdown profile. Moreover, for a fixed asset value, shutdown probability reduction and shutdown delay are substitutes on all instances in the extended instance set. The AR policies manage this substitution slightly better than CM policies on average. The sensitivity of results to changes in fixed costs and market volatilities are analogous to those observed in Table 4. For asset value losses of 3% and 6%, shutdown decisions can be delayed by 4–5 years on average across the instances by incurring a larger shutdown probability later in the horizon.

While shutdown probability reduction and shutdown delay are indeed sensitive to operational and market parameters, our findings underscore that shutdown decisions can be significantly delayed or made less likely for small asset value losses. These findings are also robust when the plant has intermediate production options, a feature not needed in the aluminum setting as discussed in §5.1. Specifically, the main social commerce trade-off and related insights already discussed continue to hold when we artificially added intermediate production options to the aluminum instances.

Table 5: Shutdown probability reduction (%) and average shutdown delay (years) for fixed asset value losses when using AR and CM policies with delay-neutral and delay-averse preferences.

	Asset value loss							
	3%				6%			
	AR		CM		AR		CM	
Instance	delay-neutral	delay-averse	delay-neutral	delay-averse	delay-neutral	delay-averse	delay-neutral	delay-averse
REF	31.6 (0.5)	16.7 (4.2)	30.0 (0.4)	13.2 (3.7)	48.5 (0.9)	22.9 (6.3)	45.2 (1.3)	22.9 (5.4)
OP1	27.3 (0.9)	13.0 (3.2)	20.0 (0.4)	7.3 (2.6)	39.0 (1.2)	19.7 (5.4)	32.5 (1.2)	13.8 (4.6)
OP2	42.9 (0.5)	19.8 (4.9)	40.0 (0.4)	20.5 (4.0)	60.6 (0.4)	31.2 (8.0)	57.5 (1.2)	39.7 (6.3)
OP3	20.4 (1.0)	12.5 (3.2)	15.9 (0.6)	6.0 (2.4)	32.2 (1.3)	15.8 (4.3)	26.0 (1.1)	9.8 (3.8)
OP4	54.6 (0.2)	25.3 (5.7)	49.2 (0.2)	32.6 (4.3)	76.6 (-0.5)	51.2 (8.4)	71.1 (2.3)	60.6 (6.0)
MP5	30.3 (0.6)	14.7 (3.7)	25.6 (0.3)	10.0 (3.2)	43.8 (0.9)	21.3 (6.1)	40.4 (1.1)	19.1 (5.2)
MP6	36.5 (0.6)	17.3 (4.1)	31.9 (0.5)	14.9 (3.7)	53.3 (0.8)	25.1 (7.0)	47.4 (1.4)	29.2 (5.9)
MP7	42.6 (1.1)	21.2 (4.9)	36.7 (1.2)	19.5 (4.5)	62.3 (1.4)	31.7 (8.5)	55.1 (2.3)	39.5 (7.1)
MP8	28.8 (0.4)	14.0 (3.5)	24.6 (0.2)	10.5 (3.3)	41.4 (0.7)	19.1 (5.3)	36.1 (0.7)	16.0 (4.6)
Average	35.0 (0.6)	17.2 (4.2)	30.4 (0.5)	14.9 (3.5)	50.8 (0.8)	26.4 (6.6)	45.7 (1.4)	27.8 (5.4)

6. Conclusions

Motivated by an aluminum producer, we studied the management of permanent shutdown decisions in merchant commodity and energy production assets from a social commerce perspective, which deviates from the popular asset value maximization approach. We formulated a constrained MDP to maximize the asset value subject to constraints on the shutdown profile, that is, we imposed bounds on the shutdown probability of an operating policy at each period to capture the preferences to delay and reduce the likelihood of a plant shutdown. We established that modifying the shutdown cost of the shutdown-neutral SDP in a state-dependent manner provides a conceptual strategy to tackle the constrained MDP but computing such cost modifications is challenging. We thus propose a low-dimensional cost modification based on anticipated regret theory, as well as, an LSM approach to compute policies. We also extended production margin-based heuristics used in practice by combining LSM and binary classification methods in machine learning. We found on realistic aluminum production instances that both policies can improve the shutdown profile substantially for small asset value losses. Production margin heuristics used by practitioners instead incur larger asset value losses to achieve these improvements. Our results highlight that social commerce appears financially viable, that is, shutdown decisions can be significantly delayed or avoided when unaccounted social costs amount to a few percent of the plant's maximum asset value. Moreover, adapting a plant's operating flexibility emerges as an effective lever to achieve socially-responsible shutdown decisions.

References

Adkins, R., D. Paxson. 2011. Reciprocal energy-switching options. *The Journal of Energy Markets* **4**(1) 91–120.

Altman, Eitan. 1999. *Constrained Markov decision processes*, vol. 7. CRC Press.

ASI. 2017. Aluminium stewardship initiative standards. https://aluminium-stewardship.org/asi-standards/. Accessed on April 17, 2018.

Bell, David E. 1982. Regret in decision making under uncertainty. *Operations research* **30**(5) 961–981.

Ben-Tal, A., A. Goryashko, E. Guslitzer, A. Nemirovski. 2004. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming* **99**(2) 351–376.

Berling, P., V. Martínez-de-Albéniz. 2011. Optimal inventory policies when purchase price and demand are stochastic. *Operations Research* **59**(1) 109–124.

BI. 2016. Business Insider. Alcoa plans to close largest U.S. aluminum smelter amid tumbling prices. Accessed on August 17, 2017.

Bishop, C. 2006. *Pattern Recognition and Machine Learning*. Springer, New York, NY, USA.

Bleichrodt, Han, Peter P Wakker. 2015. Regret theory: A bold alternative to the alternatives. *The Economic Journal* **125**(583) 493–532.

BNEF. 2017. New energy outlook 2017. Tech. rep., Bloomberg New Energy Finance (BNEF).

Boyabatli, O. 2015. Supply management in multiproduct firms with fixed proportions technology. *Management Science* **61**(12) 3013–3031.

Boyabatli, O., P. R. Kleindorfer, S. R. Koontz. 2011. Integrating long-term and short-term contracting in beef supply chains. *Management Science* **57**(10) 1771–1787.

Boyabatli, O., Q. D. Nguyen, T. Wang. 2017. Capacity management in agricultural commodity processing and its application in the palm industry. *Manufacturing & Service Operations Management* **19**(4) 551–567.

Brennan, M. J., E. S. Schwartz. 1985. Evaluating natural resource investments. *The Journal of Business* **58**(2) 135–157.

Brown, D. B., J. E. Smith, P. Sun. 2010. Information relaxations and duality in stochastic dynamic programs. *Operations Research* **58**(4) 785–801.

Brunetti, C., C. L. Gilbert. 1995. Metals price volatility, 1972–95. *Resources Policy* **21**(4) 237–254.

Carmona, R., M. Ludkovski. 2010. Valuation of energy storage: An optimal switching approach. *Quantitative Finance* **10**(4) 359–374.

Codelco. 2011. Community development. https://www.codelco.com/community-development/prontus_codelco/2011-07-08/152316.html. Accessed on March 30, 2019.

Cortazar, G., M. Gravet, J. Urzua. 2008. The valuation of multidimensional American real options using the LSM simulation method. *Computers & Operations Research* **35**(1) 113–129.

Devalkar, S. K., R. Anupindi, A. Sinha. 2005. Dynamic risk management of commodity operations: Model and analysis. *Manufacturing & Service Operations Management* **20**(2) 317–332.

Devalkar, S. K., R. Anupindi, A. Sinha. 2011. Integrated optimization of procurement, processing, and trade of commodities. *Operations Research* **59**(6) 1369–1381.

Di Corato, L., M. Moretto. 2011. Investing in biogas: Timing, technological choice and the value of flexibility from input mix. *Energy Economics* **33**(6) 1186–1193.

Dufour, F., T. Prieto-Rumeau. 2013. Finite linear programming approximations of constrained discounted markov decision processes. *SIAM Journal on Control and Optimization* **51**(2) 1298–1324.

Dufour, F., T. Prieto-Rumeau. 2014. Stochastic approximations of constrained discounted markov decision processes. *Journal of Mathematical Analysis and Applications* **413**(2) 856–879.

Eeckhoudt, Louis, Christian Gollier, Harris Schlesinger. 2005. *Economic and Financial Decisions Under Risk*. Princeton University Press.

Farkas, W., E. Gourier, R. Huitema, C. Necula. 2017. A two-factor cointegrated commodity price model with an application to spread option pricing. *Journal of Banking and Finance* **77**(1) 249–268.

Gazette-Mail. 1999. Arch Coal to close Dal-Tex by August. https://www.wvgazettemail.com/news/special_reports/arch-coal-to-close-dal-tex-by-august/article_f40f7c2b-c14a-5242-9032-a1dbf525ee77.html. Accessed on March 30, 2019.

Geman, H. 2005. *Commodities and Commodity Derivatives: Modeling and Pricing for Agriculturals, Metals and Energy*. John Wiley & Sons, West Sussex, UK.

Georghiou, Angelos, Daniel Kuhn, Wolfram Wiesemann. 2018. The decision rule approach to optimization under uncertainty: methodology and applications. *Computational Management Science* .

Glasserman, P. 2004. *Monte Carlo Methods in Financial Engineering*. Springer, New York, NY, USA.

Glasserman, P., B. Yu. 2004. Simulation for American options: Regression now or regression later? H. Niederreiter, ed., *Monte Carlo and Quasi-Monte Carlo Methods 2002*. Springer-Verlag, Berlin, Germany, 213–226.

Guthrie, G. A. 2009. *Real Options in Theory and Practice*. Oxford University Press, NY, USA.

Haugh, M. B., L. Kogan. 2004. Pricing American options: A duality approach. *Operations Research* **52**(2) 258–270.

Hekimoğlu, M. H., B. Kazaz, S. Webster. 2016. Wine analytics: Fine wine pricing and selection under weather and market uncertainty. *Manufacturing & Service Operations Management* **19**(2) 202–215.

Hochloff, P., M. Braun. 2014. Optimizing biogas plants with excess power unit and storage capacity in electricity and control reserve markets. *Biomass and Bioenergy* **65**(1) 125–135.

Hydro. 2018. Restructuring. https://www.hydro.com/en/our-future/Society/Restructuring/. Accessed on March 14, 2018.

IMF. 2015. World Economic Outlook: Commodity Special Feature. Tech. rep., International Monetary Fund.

IPIECA. 2019. Social responsibility. http://www.ipieca.org/our-work/social/. Accessed on March 30, 2019.

ISO. 2014. ISO 26000 - Social responsibility. https://www.iso.org/iso-26000-social-responsibility.html. Accessed on March 30, 2019.

Kasa, S. 2000. Policy networks as barriers to green tax reform: The case of CO2-taxes in Norway. *Environmental Politics* **9**(4) 104–122.

Kaya, I. 2016. The mandatory social and environmental reporting: Evidence from france. *Procedia-Social and Behavioral Sciences* **229** 206–213.

Kazaz, B., S. Webster. 2011. The impact of yield-dependent trading costs on pricing and production planning under supply uncertainty. *Manufacturing & Service Operations Management* **13**(3) 404–417.

Kleindorfer, P. R., K. Singhal, L. N. Wassenhove. 2005. Sustainable operations management. *Production and Operations Management* **14**(4) 482–492.

Kulatilaka, N., L. Trigeorgis. 2001. The general flexibility to switch: Real options revisited. E. S. Schwartz, L. Trigeorgis, eds., *Real Options and Investment Under Uncertainty: Classical Readings and Recent Contributions*. MIT press, Boston, MA, 179–198.

Lai, G., F. Margot, N. Secomandi. 2010. An approximate dynamic programming approach to benchmark practice-based heuristics for natural gas storage valuation. *Operations Research* **58**(3) 564–582.

Lee, H. L., C. S. Tang. 2017. Socially and environmentally responsible value chain innovations: New operations management opportunities. *Management Science* **64**(3) 983–996.

Longstaff, F. A., E. S. Schwartz. 2001. Valuing American options by simulation: A least-squares approach. *Review of Financial Studies* **14**(1) 113–147.

Loomes, G., R. Sugden. 1982. Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal* **92**(368) 805–824.

Mazières, D., A. Boogert. 2013. A radial basis function approach to gas storage valuation. *Journal of Energy Markets* **6**(2) 19–50.

Méndez, C. A., I. E. Grossmann, Harjunkoski I., P. Kaboré. 2006. A simultaneous optimization approach for off-line blending and scheduling of oil-refinery operations. *Computers & Chemical Engineering* **30**(4) 614–634.

Nadarajah, S., F. Margot, N. Secomandi. 2015. Relaxations of approximate linear programs for the real option management of commodity storage. *Management Science* **61**(12) 3054–3076.

Nadarajah, S., F. Margot, N. Secomandi. 2017. Comparison of least squares Monte Carlo methods with applications to energy real options. *European Journal of Operational Research* **256**(1) 196–204.

Nadarajah, S., N. Secomandi. 2017. Relationship between least squares Monte Carlo and approximate linear programming. *Operations Research Letters* **45**(5) 409–414.

Nadarajah, S., N. Secomandi, G. Sowers, J. M. Wassick. 2016. *Real Option Management of Hydrocarbon Cracking Operations*, chap. Chapter 5. World Scientific/NOW Publishers, 173–202.

Øye, H. A., M Sørlie. 2011. Power failure, restart and repair. *Aluminium International Today* .

Powell, W. B. 2011. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. 2nd ed. John Wiley & Sons, Hoboken, NJ, USA.

Reuters. 2018. Green Plains shuts plant, faces ethanol downturn. https://www.reuters.com/article/us-green-plains-ethanol-operations/green-plains-shuts-plant-faces-ethanol-downturn-idUSKCN1NK2ON. Accessed on March 30, 2019.

Rosasco, L., E. D. Vito, A. Caponnetto, M. Piana, A. Verri. 2004. Are loss functions all the same? *Neural Computation* **16**(5) 1063–1076.

Schwartz, E. S., J. E. Smith. 2000. Short-term variations and long-term dynamics in commodity prices. *Management Science* **46**(7) 893–911.

Secomandi, N. 2015. Merchant commodity storage practice revisited. *Operations Research* **63**(5) 1131–1143.

Secomandi, N., D. Seppi. 2014. Real Options and Merchant Operations of Energy and Other Commodities. *Foundations and Trends in Technology, Information and Operations Management* **6**(3-4) 161–331.

Thompson, M. 2016. Natural gas storage valuation, optimization, market and credit risk management. *Journal of Commodity Markets* **2**(1) 26–44.

Tseng, C., G. Barz. 2002. Short-term generation asset valuation: A real options approach. *Operations Research* **50**(2) 297–310.

Tsiros, M., V. Mittal. 2000. Regret: A model of its antecedents and consequences in consumer decision making. *Journal of Consumer Research* **26**(4) 401–417.

Tversky, Amos, Daniel Kahneman. 1992. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty* **5**(4) 297–323.

Wu, O. Q., D. D. Wang, Z. Qin. 2012. Seasonal energy storage operations with limited flexibility: The price-adjusted rolling intrinsic policy. *Manufacturing & Service Operations Management* **14**(3) 455–471.

Yang, B., S. Nadarajah, N. Secomandi. 2017. Approximate dynamic programming for merchant energy production. Working paper, Carnegie Mellon Univ. and Univ. of Illinois at Chicago.

Zeelenberg, M. 1999. Anticipated regret, expected feedback and behavioral decision making. *Journal of Behavioral Decision Making* **12**(2) 93.

Managing Shutdown Decisions in Merchant Commodity and Energy Production: A Social Commerce Perspective (Online Supplement)

A. Proofs

Lemma 1 is used to prove Proposition 1.

Lemma 1. *There exists an optimal SC-MDP policy π^* such that for each $(i, x_i, w_i) \in \mathcal{I} \times \mathcal{X}_i \times \mathcal{W}_i$, if $\pi_i^{\mathrm{SN}}(x_i, w_i) = a_i \neq \mathrm{C}$, then $\pi_i^*(x_i, w_i) = a_i$.*

Proof. We prove that the optimal shutdown-neutral and SC-MDP policies choose to operate at the same states. We then argue that the specific operating decisions taken by these policies is the same.

The proof proceeds by contradiction, that is, if the statement was false, we will exhibit a new feasible SC-MDP policy π^{**} with higher expected value than π^*. Suppose there exists some $(i, x_i, w_i) \in \mathcal{I} \times \mathcal{X}_i \times \mathcal{W}_i$ in which $\pi_i^{\mathrm{SN}}(x_i, w_i) \neq \mathrm{C}$ but $\pi_i^*(x_i, w_i) = \mathrm{C}$. In other words, there exist stage i states $\mathcal{O}_i^1 \subseteq \mathcal{X}_i \times \mathcal{W}_i$ in which the shutdown neutral policy operates but π^* instead shuts down. Without loss of generality, we assume that $\mathcal{O}_i^1$ has positive probability, i.e., $\Pr(\mathcal{O}_i^1) > 0$. We denote by $\mathcal{O}_i^2 := (\mathcal{X}_i \times \mathcal{W}_i) \backslash \mathcal{O}_i^1$ and by $(x_j, w_j)|\mathcal{O}_i^1$ the states reached at stage $j \geqslant i$ conditional on the stage i state belonging to $\mathcal{O}_i^1$. We define a new policy π^{**} by combining π^{SN} and π^* as follows: $\pi_j^{**}(x_j, w_j) := \pi_j^{\mathrm{SN}}(x_j, w_j)$ for $j \geqslant i$ and states $(x_j, w_j)|\mathcal{O}_i^1$, and $\pi_j^{**}(x_j, w_j) := \pi_j^*(x_j, w_j)$ otherwise. The MDP bounds U_j, $j \in \mathcal{I}$, are satisfied by π^{**}. In fact, for $j < i$, π^* and π^{**} coincide, thus the probability bound U_j is respected since π^* is feasible to SC-MDP. For $j \geqslant i$, we can decompose the shutdown probability of π^{**} as follows (we omit the conditioning on (x_0, w_0)):

$$
\begin{aligned}
\Pr(x_j^{\pi^{**}} = \mathrm{C}) &= \Pr(x_j^{\pi^{**}} = \mathrm{C}|\mathcal{O}_i^1) \cdot \Pr(\mathcal{O}_i^1) + \Pr(x_j^{\pi^{**}} = \mathrm{C}|\mathcal{O}_i^2) \cdot \Pr(\mathcal{O}_i^2) && \text{(11a)}\\
&= \Pr(x_j^{\pi^{**}} = \mathrm{C}|\mathcal{O}_i^1) \cdot \Pr(\mathcal{O}_i^1) + \Pr(x_j^{\pi^{*}} = \mathrm{C}|\mathcal{O}_i^2) \cdot \Pr(\mathcal{O}_i^2) && \text{(11b)}\\
&\leqslant \Pr(x_j^{\pi^{*}} = \mathrm{C}|\mathcal{O}_i^1) \cdot \Pr(\mathcal{O}_i^1) + \Pr(x_j^{\pi^{*}} = \mathrm{C}|\mathcal{O}_i^2) \cdot \Pr(\mathcal{O}_i^2) && \text{(11c)}\\
&= \Pr(x_j^{\pi^{*}} = \mathrm{C}) && \text{(11d)}\\
&\leqslant U_j, && \text{(11e)}
\end{aligned}
$$

where (11b) holds because π^* and π^{**} coincide in $\mathcal{O}_i^2$, (11c) is a consequence of $\Pr(x_j^{\pi^*} = \mathrm{C}|\mathcal{O}_i^1) = 1$ by definition of $\mathcal{O}_i^1$, i.e., in states $(x_j, w_j)|\mathcal{O}_i^1$ the policy π^* shuts down at stage i, and (11e) holds as π^* is feasible. Moreover, π^{**} having strictly greater expected value than π^* follows directly from $V_i^{\mathrm{SN},\mathcal{O}}(x_i, w_i) > -K$ on a set with positive mass $(x_i, w_i) \in \mathcal{O}_i^1$ (since π^{SN} operates in states $\mathcal{O}_i^1$), which contradicts π^* being optimal SC-MDP policy.

In addition, the specific operating (i.e., non-shutdown) actions taken by π^{SN} and π^* coincide. If this wasn't true, we could define a new feasible policy π^{**} with higher value than π^* where π^{**} coincides with π^{SN} on the non-shutdown states in which π^* deviates from π^{SN}. The higher value of such a policy follows from the fact that the shutdown-neutral policy π^{SN} maximizes value. □

Proof of Proposition 1. **Part (a)**. The SDP with state-dependent shutdown cost modification is:

$$V_I^{\mathrm{D}}(x_I, w_I) = r_I(x_I, w_I), \ \forall (x_I, w_I) \in \mathcal{X}_I \times \mathcal{W}_I, \tag{12a}$$

$$V_i^{\mathrm{D},\mathcal{O}}(x_i, w_i) = \max_{a_i \in \mathcal{A}_i(x_i) \backslash \{\mathrm{C}\}} \left\{ r_i(x_i, w_i, a_i) + \delta \mathbb{E}_i\left[V_{i+1}^{\mathrm{D}}\left(f_i(x_i, a_i), w_{i+1}\right)\right]\right\}, \quad \forall (i, x_i, w_i) \in \mathcal{I} \times \mathcal{X}_i \times \mathcal{W}_i, \tag{12b}$$

$$V_i^{\mathrm{D}}(x_i, w_i) = \max\left\{V_i^{\mathrm{D},\mathcal{O}}(x_i, w_i), -\bar{K}_i^{(x_i, w_i, \mathrm{C})}\right\}, \ \forall (i, x_i, w_i) \in \mathcal{I} \times \mathcal{X}_i \times \mathcal{W}_i. \tag{12c}$$

Since this cost modification only inflates the shutdown cost, i.e., $\bar{K}_i^{(x_i,w_i,\mathrm{C})} \geqslant K^{(x_i,\mathrm{C})}$, the value functions in SDP (12) can only decrease compared to the shutdown neutral SDP, i.e. $V_i^{\mathrm{D},\mathcal{O}}(x_i,w_i) \leqslant V_i^{\mathrm{SN},\mathcal{O}}(x_i,w_i)$, $\forall (i,x_i,w_i) \in \mathcal{I} \times \mathcal{X}_i \times \mathcal{W}_i$. At stage i, we define three sets of states:

1. $\mathcal{O}_i^1 := \{(x_i,w_i) \in \mathcal{X}_i \times \mathcal{W}_i : \pi_i^{\mathrm{SN}}(x_i,w_i) = \pi_i^*(x_i,w_i) = a_i \neq \mathrm{C}\}$, the *good* states;
2. $\mathcal{O}_i^2 := \{(x_i,w_i) \in \mathcal{X}_i \times \mathcal{W}_i : \pi_i^{\mathrm{SN}}(x_i,w_i) = \mathrm{C},\ \pi_i^*(x_i,w_i) \neq \mathrm{C}\}$, the *swap* states;
3. $\mathcal{O}_i^3 := \{(x_i,w_i) \in \mathcal{X}_i \times \mathcal{W}_i : \pi_i^{\mathrm{SN}}(x_i,w_i) = \pi_i^*(x_i,w_i) = \mathrm{C}\}$, the *bad* states.

Notice that Lemma 1 ensures $\mathcal{X}_i \times \mathcal{W}_i = \mathcal{O}_i^1 \cup \mathcal{O}_i^2 \cup \mathcal{O}_i^3$, for every $i \in \mathcal{I}$, that is, $\pi_i^{\mathrm{SN}}(x_i,w_i) \neq \mathrm{C}$ and $\pi_i^*(x_i,w_i) = \mathrm{C}$ cannot happen. Proceeding backward for $i = I-1,\ldots,0$, for each $w_i \in \mathcal{W}_i$, we define the modified shutdown costs $\bar{K}_i^{(x_i,w_i,\mathrm{C})}$ at stage i and state $(x_i,w_i) \in \mathcal{X}_i \times \mathcal{W}_i$ as:

$$\bar{K}_i^{(x_i,w_i,\mathrm{C})} := \begin{cases} K^{(x_i,\mathrm{C})} + V_i^{\mathrm{SN},\mathcal{O}}(x_i,w_i) - V_i^{\mathrm{D},\mathcal{O}}(x_i,w_i) & \text{if } (x_i,w_i) \in \mathcal{O}_i^1; \\ -V_i^{\mathrm{D},\mathcal{O}}(x_i,w_i) + 1 & \text{if } (x_i,w_i) \in \mathcal{O}_i^2; \\ K^{(x_i,\mathrm{C})} & \text{if } (x_i,w_i) \in \mathcal{O}_i^3. \end{cases} \tag{13}$$

By defining the shutdown cost as above, the shutdown actions taken by an optimal policy π^{D} of SDP (12) at every stage and state will always match those taken by π^*. In fact:

1. if $(x_i,w_i) \in \mathcal{O}_i^1$, then we know that $V_i^{\mathrm{SN},\mathcal{O}}(x_i,w_i) \geqslant -K^{(x_i,\mathrm{C})}$, but the same is not necessarily true for $V_i^{\mathrm{D},\mathcal{O}}(x_i,w_i)$ since $V_i^{\mathrm{D},\mathcal{O}}(x_i,w_i) \leqslant V_i^{\mathrm{SN},\mathcal{O}}(x_i,w_i)$. The cost inflation in (13) ensures that $V_i^{\mathrm{D},\mathcal{O}}(x_i,w_i) \geqslant -\bar{K}_i^{(x_i,w_i,\mathrm{C})}$ for all states $(x_i,w_i) \in \mathcal{O}_i^1$, i.e., π^{D} chooses to operate.
2. if $(x_i,w_i) \in \mathcal{O}_i^2$, $V_i^{\mathrm{D},\mathcal{O}}(x_i,w_i) \leqslant V_i^{\mathrm{SN},\mathcal{O}}(x_i,w_i) < -K^{(x_i,\mathrm{C})}$ for $(x_i,w_i) \in \mathcal{O}_i^2$ but we want the policy π^{D} to swap the π^{SN} decision from shutdown to operate, which requires the shutdown cost to be $\bar{K}_i^{(x_i,w_i,\mathrm{C})} > -V_i^{\mathrm{D},\mathcal{O}}(x_i,w_i)$. The cost inflation in (13) ensures this inequality holds.
3. if $(x_i,w_i) \in \mathcal{O}_i^3$, since $V_i^{\mathrm{D},\mathcal{O}}(x_i,w_i) \leqslant V_i^{\mathrm{SN},\mathcal{O}}(x_i,w_i) < -K^{(x_i,\mathrm{C})}$ and both π^{SN} and π^* shut down, there is no need to inflate the shutdown cost and we keep $\bar{K}_i^{(x_i,w_i,\mathrm{C})} = K^{(x_i,\mathrm{C})}$.

Part (b). In the following, we construct a shutdown cost modification (not necessarily a cost inflation) such that the shutdown actions from SDP (12) are the same as π^* as in Part (a) and, in addition, the non-shutdown actions from the two policies also coincide. Specifically, we provide a cost modification that respects the π^* shutdown decisions and such that

$$\mathbb{E}_i\big[V_{i+1}^{\mathrm{D}}(x_{i+1},w_{i+1})\big] = \mathbb{E}_i\big[V_{i+1}^{\mathrm{SN}}(x_{i+1},w_{i+1})\big] \tag{14}$$

for each $(i,w_i) \in \mathcal{I} \times \mathcal{W}_i$ and $x_{i+1} \in \mathcal{X}_{i+1}$. This policy would imply the desired result since

$$V_i^{\mathrm{D},\mathcal{O}}(x_i,w_i) = \max_{a_i \in \mathcal{A}_i(x_i)\setminus\{\mathrm{C}\}} \Big\{ r_i(x_i,w_i,a_i) + \delta \mathbb{E}_i\big[V_{i+1}^{\mathrm{D}}(f_i(x_i,a_i),w_{i+1})\big] \Big\} \tag{15a}$$

$$= \max_{a_i \in \mathcal{A}_i(x_i)\setminus\{\mathrm{C}\}} \Big\{ r_i(x_i,w_i,a_i) + \delta \mathbb{E}_i\big[V_{i+1}^{\mathrm{SN}}(f_i(x_i,a_i),w_{i+1})\big] \Big\} = V_i^{\mathrm{SN},\mathcal{O}}(x_i,w_i), \tag{15b}$$

for each $(i,x_i,w_i) \in \mathcal{I} \times \mathcal{X}_i \times \mathcal{W}_i$, meaning that the non-shutdown actions from SDP (12) coincide with those from the shutdown-neutral SDP (2), which in turn coincide with those taken by π^* due Lemma 1. We prove this statement by induction. At the last decision stage $I-1$, (14) is trivial to verify since for each $(x_I,w_I) \in \mathcal{X}_I \times \mathcal{W}_I$ we have $V_I^{\mathrm{D}}(x_I,w_I) = r_I(x_I,w_I) = V_I^{\mathrm{SN}}(x_I,w_I)$, thus it holds that $\mathbb{E}_{I-1}[V_I^{\mathrm{D}}(x_I,w_I)] = \mathbb{E}_{I-1}[V_I^{\mathrm{SN}}(x_I,w_I)]$ (there is no cost modification involved at this stage). We now assume the statement is true from stage i to I, and prove it for $i-1$.

Step 1. To ensure the shutdown actions of π^* are preserved, we inflate the shutdown cost $\bar{K}_i^{(x_i,w_i,\mathrm{C})}$ for states $(x_i,w_i) \in \mathcal{O}_i^1 \cup \mathcal{O}_i^2$ as in Part (a) of the proof, i.e., using equation (13).

Step 2. For each state $x_i \in \mathcal{X}_i$, we define the two sets $\mathcal{O}_i^{\mathrm{A}}(x_i) := \{w_i \in \mathcal{W}_i : (x_i, w_i) \in \mathcal{O}_i^1 \cup \mathcal{O}_i^2\}$ and $\mathcal{O}_i^{\mathrm{B}}(x_i) := \{w_i \in \mathcal{W}_i : (x_i, w_i) \in \mathcal{O}_i^3\}$, representing slices of $\mathcal{O}_i^1 \cup \mathcal{O}_i^2$ and $\mathcal{O}_i^3$, respectively, for fixed x_i.

Step 3. Consider $x_i \in \mathcal{X}_i$ and $w_i \in \mathcal{O}_i^{\mathrm{A}}(x_i)$. After executing the cost modification in Step 1, we have $V_i^{\mathrm{D}}(x_i, w_i) \leqslant V_i^{\mathrm{SN}}(x_i, w_i)$ because: (i) $V_i^{\mathrm{D},O}(x_i, w_i) = V_i^{\mathrm{SN},O}(x_i, w_i)$ by the induction hypothesis, and (ii) the stage-i shutdown cost for exogenous states in $\mathcal{O}_i^{A}(x_i)$ has been inflated. Therefore, $\mathbb{E}_{i-1}[V_i^{\mathrm{D}}(x_i, w_i)|\mathcal{O}_i^{\mathrm{A}}(x_i)] \leqslant \mathbb{E}_{i-1}[V_i^{\mathrm{SN}}(x_i, w_i)|\mathcal{O}_i^{\mathrm{A}}(x_i)]$. The idea is to balance this reduction in expected value using the states $\mathcal{O}_i^{\mathrm{B}}(x_i)$, where we have $V_i^{\mathrm{SN},\mathcal{O}}(x_i, w_i) \leqslant -K^{(x_i,\mathrm{C})}$, i.e., the shutdown neutral policy shuts down and $-K^{(x_i,\mathrm{C})}$ determines the value function. By deflating the shutdown cost when the exogenous state belongs to $\mathcal{O}_i^{\mathrm{B}}(x_i)$, the optimal action of SDP (2) would not be altered but the value function would increase. For $w_i \in \mathcal{O}_i^{\mathrm{B}}(x_i)$, we define $\bar{K}_i^{(x_i,w_i,\mathrm{C})}$ as follows:

$$\bar{K}_i^{(x_i,w_i,\mathrm{C})} := K^{(x_i,\mathrm{C})} - \alpha_i(x_i) \tag{16a}$$

$$\alpha_i(x_i) := \left(\mathbb{E}\left[V_i^{\mathrm{SN}}(x_i, w_i)\,|\mathcal{O}_i^{\mathrm{A}}(x_i)\right] - \mathbb{E}\left[V_i^{\mathrm{D}}(x_i, w_i)\,|\mathcal{O}_i^{\mathrm{A}}(x_i)\right]\right) \frac{\Pr(\mathcal{O}_i^{\mathrm{A}}(x_i))}{\Pr(\mathcal{O}_i^{\mathrm{B}}(x_i))} \tag{16b}$$

where $\alpha_i(x_i)$ is non-negative and well-defined due to $\Pr(\mathcal{O}_i^{\mathrm{B}}(x_i)) > 0$ by assumption. Finally, (14) holds at $i-1$ and $(x_i, w_{i-1}) \in \mathcal{W}_{i-1} \times \mathcal{X}_i$ using (16) because

$$\begin{aligned}
\mathbb{E}_{i-1}\left[V_i^{\mathrm{D}}(x_i, w_i)\right] &= \mathbb{E}_{i-1}\left[V_i^{\mathrm{D}}(x_i, w_i)|\mathcal{O}_i^{\mathrm{A}}(x_i)\right]\Pr(\mathcal{O}_i^{\mathrm{A}}(x_i)) - \bar{K}_i^{(x_i,w_i,\mathrm{C})}\Pr(\mathcal{O}_i^{\mathrm{B}}(x_i)) &(17a)\\
&= \mathbb{E}_{i-1}\left[V_i^{\mathrm{SN}}(x_i, w_i)|\mathcal{O}_i^{\mathrm{A}}(x_i)\right]\Pr(\mathcal{O}_i^{\mathrm{A}}(x_i)) - K^{(x_i,\mathrm{C})}\Pr(\mathcal{O}_i^{\mathrm{B}}(x_i)) &(17b)\\
&= \mathbb{E}_{i-1}\left[V_i^{\mathrm{SN}}(x_i, w_i)\right] &(17c)
\end{aligned}$$

where (17b) is obtained by replacing $\bar{K}_i^{(x_i,w_i,\mathrm{C})}$ with its expression in (16), and simplifying. □

Proof of Proposition 2. **Part (a)**. Our proof is based on backward induction. At the last stage I, the following inequalities trivially hold as equalities because the terminal conditions are the same in each SDP:

$$V_I^{\mathrm{SN}\backslash\{\mathrm{C}\}}(x_I, w_I) = V_{\Theta,I}^{\mathrm{A}}(x_I, w_I) = V_I^{\mathrm{SN}}(x_I, w_I) = r_I(x_I, w_I), \ \ \forall (x_I, w_I) \in \mathcal{X}_I \times \mathcal{W}_I. \tag{18}$$

Assume that analogous inequalities hold from stages $i+1$ through stage I. Consider stage i and the inequalities $V_{i+1}^{\mathrm{SN}\backslash\{\mathrm{C}\}}(x_{i+1}, w_{i+1}) \leqslant V_{\Theta,i+1}^{\mathrm{A}}(x_{i+1}, w_{i+1}) \leqslant V_{i+1}^{\mathrm{SN}}(x_{i+1}, w_{i+1})$ for all $(x_{i+1}, w_{i+1}) \in \mathcal{X}_{i+1} \times \mathcal{W}_{i+1}$, which hold by the induction hypothesis. Using these inequalities, we obtain for each state $(x_i, w_i) \in \mathcal{X}_i \times \mathcal{W}_i$ and non-shutdown action $a_i \in \mathcal{A}_i(x_i)\backslash\{\mathrm{C}\}$, the following relationships:

$$\begin{aligned}
r_i(x_i, w_i, a_i) + \delta\mathbb{E}_i\left[V_{i+1}^{\mathrm{SN}\backslash\{\mathrm{C}\}}(f(x_i, a_i), w_{i+1})\right] &\leqslant r_i(x_i, w_i, a_i) + \delta\mathbb{E}_i\left[V_{\Theta,i+1}^{\mathrm{A}}(f(x_i, a_i), w_{i+1})\right]\\
&\leqslant r_i(x_i, w_i, a_i) + \delta\mathbb{E}_i\left[V_{i+1}^{\mathrm{SN}}(f(x_i, a_i), w_{i+1})\right].
\end{aligned}$$

Taking a maximum over the non shutdown actions preserves this ordering, that is,

$$V_i^{\mathrm{SN}\backslash\{\mathrm{C}\}}(x_i, w_i) \leqslant V_{\Theta,i}^{\mathrm{A},\mathcal{O}}(x_i, w_i) \leqslant V_i^{\mathrm{SN},\mathcal{O}}(x_i, w_i), \ \ \forall (x_i, w_i) \in \mathcal{X}_i \times \mathcal{W}_i.$$

Using $V_i^{\mathrm{SN}\backslash\{\mathrm{C}\}}(x_i, w_i) \leqslant V_{\Theta,i}^{\mathrm{A},\mathcal{O}}(x_i, w_i)$ we get

$$V_i^{\mathrm{SN}\backslash\{\mathrm{C}\}}(x_i, w_i) \leqslant \max\left\{V_{\Theta,i}^{\mathrm{A},\mathcal{O}}(x_i, w_i), -K^{(x_i,\mathrm{C})} - \lambda\xi^i \mathrm{AR}_{\Theta,i}(x_i, w_i)\right\} = V_{\Theta,i}^{\mathrm{A}}(x_i, w_i), \ \ \forall (x_i, w_i) \in \mathcal{X}_i \times \mathcal{W}_i.$$

To prove $V_{\Theta,i}^{\mathrm{A}}(x_i, w_i) \leqslant V_i^{\mathrm{SN}}(x_i, w_i)$ we notice that the penalty term $\mathrm{AR}_{\Theta,i}(x_i, w_i)$ is always positive. This implies that

$$-K^{(x_i,\mathrm{C})} - \lambda\xi^i \mathrm{AR}_{\Theta,i}(x_i, w_i) \leqslant -K^{(x_i,\mathrm{C})}, \ \ \forall (x_i, w_i) \in \mathcal{X}_i \times \mathcal{W}_i. \tag{19}$$

Combining $V^{\mathrm{A},\mathcal{O}}_{\Theta,i}(x_i,w_i) \leqslant V^{\mathrm{SN},\mathcal{O}}_i(x_i,w_i)$ with (19) results in

$$\begin{aligned} V^{\mathrm{A}}_{\Theta,i}(x_i,w_i) &= \max\left\{V^{\mathrm{A},\mathcal{O}}_{\Theta,i}(x_i,w_i), -K^{(x_i,\mathrm{C})} - \lambda\xi^i \mathrm{AR}_{\Theta,i}(x_i,w_i)\right\} \\ &\leqslant \max\left\{V^{\mathrm{SN},\mathcal{O}}_i(x_i,w_i), -K^{(x_i,\mathrm{C})}\right\} = V^{\mathrm{SN}}_i(x_i,w_i), \ \ \forall (x_i,w_i) \in \mathcal{X}_i \times \mathcal{W}_i. \end{aligned}$$

Part (b). We fix $\xi \in (0,1]$. To prove the limiting results, it suffices to show that the shutdown-neutral value function with no shutdown and the anticipated regret value function coincide for any strictly positive lambda λ and large enough η. This result implies that for a given $\lambda > 0$, $\lim_{\eta\to\infty} V^{\mathrm{A}}_{\Theta,i}(x_i,w_i) = V^{\mathrm{SN}\setminus\{\mathrm{C}\}}_I(x_i,w_i)$. Moreover, since the value functions coincide, the associated optimal policies are also equal, which implies that the shutdown probability under the anticipated regret policy is zero, that is, $\lim_{\eta\to\infty} \Pr(\mathrm{C};\pi^{\mathrm{A}}_{\Theta}) = 0$ for any $\lambda > 0$.

We proceed to show that for every $\lambda > 0$, there exists $\eta^*_i \geqslant 1$ (we suppress the dependence of η^* on λ) such that $V^{\mathrm{A}}_{\Theta,i}(x_i,w_i) = V^{\mathrm{SN}\setminus\{\mathrm{C}\}}_i(x_i,w_i)$ for all $(x_i,w_i) \in \mathcal{X}_i \times \mathcal{W}_i$ and $\eta \geqslant \eta^*_i$. The statement is proved by backward induction. At stage I, the equality trivially holds for any η and λ as in (18). Assuming that analogous equalities and values for η exist for stages $i+1$ to I, we establish this property at stage i. The equality $V^{\mathrm{A}}_{\Theta,i+1}(x_{i+1},w_{i+1}) = V^{\mathrm{SN}\setminus\{\mathrm{C}\}}_{i+1}(x_{i+1},w_{i+1})$ for all $(x_{i+1},w_{i+1}) \in \mathcal{X}_{i+1} \times \mathcal{W}_{i+1}$ is true by the induction hypothesis for $\eta \geqslant \eta^*_{i+1}$. Note that $V^{\mathrm{A}}_{\Theta,i+1}(x_{i+1},w_{i+1}) = V^{\mathrm{SN}\setminus\{\mathrm{C}\}}_{i+1}(x_{i+1},w_{i+1})$ implies that $V^{\mathrm{A}}_{\Theta,i+1}(x_{i+1},w_{i+1}) = V^{\mathrm{A},\mathcal{O}}_{\Theta,i+1}(x_{i+1},w_{i+1})$ because the shutdown decision does not determine the value function. Based on these relationships we replace $V^{\mathrm{A}}_{\Theta,i+1}(x_{i+1},w_{i+1})$ in the right hand side of the anticipated regret SDP step (5c) with $V^{\mathrm{SN}\setminus\{\mathrm{C}\}}_i(x_i,w_i)$ and also use the definition of $\mathrm{AR}_{\Theta,i}(x_i,w_i)$ to obtain

$$\begin{aligned} V^{\mathrm{A}}_{\Theta,i}(x_i,w_i) = \max\Big\{V^{\mathrm{SN}\setminus\{\mathrm{C}\}}_i(x_i,w_i), -K^{(x_i,\mathrm{C})} - \lambda\xi^i \max_{a_i\in\mathcal{A}_i(x_i)\setminus\{\mathrm{C}\}} \mathbb{E}_i\big[\max\{r_i(x_i,w_i,a_i) \\ +\delta V^{\mathrm{SN}\setminus\{\mathrm{C}\}}_{i+1}(f_i(x_i,a_i),w_{i+1}) + \eta\cdot K^{(x_i,\mathrm{C})},0\}\big]\Big\}. \end{aligned} \tag{20}$$

Since $\mathcal{W}_i$ is compact at each stage by assumption, there exists an $L \in \mathbb{R}_+$ such that $r_i(x_i,w_i,a_i) \geqslant -L$ for all $(i,x_i,w_i,a_i) \in \mathcal{I} \times \mathcal{X}_i \times \mathcal{W}_i \times \mathcal{A}_i(x_i)$ and $r_I(x_I,w_I) \geqslant -L$ for all $(x_I,w_I) \in \mathcal{X}_I \times \mathcal{W}_I$. An immediate consequence is that the optimal value function is bounded below, that is, $V^{\mathrm{SN}\setminus\{\mathrm{C}\}}_i(x_i,w_i) \geqslant -L(I-i+1)$ for all $(i,x_i,w_i) \in \mathcal{I} \times \mathcal{X}_i \times \mathcal{W}_i$. To prove that $V^{\mathrm{A}}_{\Theta,i}(x_i,w_i) = V^{\mathrm{SN}\setminus\{\mathrm{C}\}}_i(x_i,w_i)$ are equal if η is large enough, we show that shutdown will not be chosen in equation (20) by using the following chain of inequalities for all $(x_i,w_i) \in \mathcal{X}_i \times \mathcal{W}_i$:

$$- K^{(x_i,\mathrm{C})} - \lambda\xi^i \max_{a_i\in\mathcal{A}_i(x_i)\setminus\{\mathrm{C}\}} \mathbb{E}_i\left[\max\left\{r_i(x_i,w_i,a_i) + \delta V^{\mathrm{SN}\setminus\{\mathrm{C}\}}_{i+1}(f_i(x_i,a_i),w_{i+1}) + \eta\cdot K^{(x_i,\mathrm{C})},0\right\}\right]$$

$$\leqslant -\lambda\xi^i \max_{a_i\in\mathcal{A}_i(x_i)\setminus\{\mathrm{C}\}} \mathbb{E}_i\left[r_i(x_i,w_i,a_i) + \delta V^{\mathrm{SN}\setminus\{\mathrm{C}\}}_{i+1}(f_i(x_i,a_i),w_{i+1}) + \eta\cdot K^{(x_i,\mathrm{C})}\right] \tag{21a}$$

$$\leqslant -\lambda\xi^i\left(-L - \delta L(I-i) + \eta\underline{K}\right) \tag{21b}$$

$$< -L - \delta L(I-i) \tag{21c}$$

$$\leqslant \max_{a_i\in\mathcal{A}_i(x_i)\setminus\{C\}} \left\{r_i(x_i,w_i,a_i) + \delta\mathbb{E}_i\left[V^{\mathrm{SN}\setminus\{\mathrm{C}\}}_{i+1}\left(f(x_i,a_i),w_{i+1}\right)\right]\right\} \tag{21d}$$

$$= V^{\mathrm{SN}\setminus\{\mathrm{C}\}}_i(x_i,w_i). \tag{21e}$$

The first inequality (21a) is obtained by dropping the first term $-K^{(x_i,C)} \leqslant 0$ and employing the relation $\mathbb{E}[\max\{X,0\}] \geqslant \mathbb{E}[X]$; the second (21b) by replacing the reward function and the value function terms by their lower bounds based on the compactness of $\mathcal{W}_i$, and the shutdown cost by

$\underline{K} = \min\{K^{(x_i,\mathrm{C})} : x_i \in \mathcal{X}_i, i \in \mathcal{I}\} > 0$; the third (21c) by choosing $\eta > \eta_i^* = (\lambda\xi^i + 1)(L + \delta L(I - i))/(\lambda\xi^i\,\underline{K})$; the fourth (21d) from noticing that $r_i(x_i, w_i, a_i) \geqslant -L$ and $V_{i+1}^{\mathrm{SN}\backslash\{\mathrm{C}\}}(f(x_i, a_i), w_{i+1}) \geqslant -L(I - i)$; and the final equality (21e) by using the definition of $V_i^{\mathrm{SN}\backslash\{\mathrm{C}\}}(x_i, w_i)$. Our claim thus holds at stage i and it is also true at all stages by the principle of mathematical induction. □

B. LSML algorithm

We describe in this section the LSML procedure briefly discussed in §3.2 of the paper. We primarily focus on tackling the anticipated-regret SDP (5) but point out at the end of this section minor changes that can be made to approximate the shutdown-neutral SDPs (2) and (4).

Algorithm 1: LSML

Inputs: Set of information state sample paths $\{w_i^p, (i,p) \in \mathcal{I} \cup \{I\} \times \mathcal{P}\}$, number of sample average approximation samples N, and set of basis function vectors $\{\Phi_i, i \in \{1, \ldots, I\}\}$.

Initialization: For each $x_I \in \mathcal{X}_I$, compute estimates $v_I^{\mathrm{A}}(x_I, w_I^p) := r_I(x_I, w_I^p)$ for $p \in \mathcal{P}$ and perform a least squares regression on these VFA estimates using basis functions Φ_I to determine the vector of VFA weights $\beta_{I,x_I}^{\mathrm{A}}$.

For each $i = I - 1$ to 1 **do**:

For each $x_i \in \mathcal{X}_i$ **do**:

1. **For** each $p \in \mathcal{P}$ **do**:
   (a) Sample N stage $i+1$ information state samples conditional on w_i^p: $\{\bar{w}_{i+1}^{p,n}, n \in \mathcal{N}\}$.
   (b) Compute the VFA estimates
   $$v_i^{\mathrm{A},\mathcal{O}}(x_i, w_i^p) := \max_{a_i \in \mathcal{A}_i(x_i)\backslash\{\mathrm{C}\}} \left\{ r_i(x_i, w_i^p, a_i) + \delta \sum_{b \in \mathcal{B}_{i+1}} \beta_{i+1, f_i(x_i,a_i), b}^{\mathrm{A}} \left[\frac{1}{N} \sum_{n \in \mathcal{N}} \Phi_{i+1,b}(\bar{w}_{i+1}^{p,n}) \right] \right\};$$
   $$v_i^{\mathrm{A}}(x_i, w_i^p) := \max\left\{ v_i^{\mathrm{A},\mathcal{O}}(x_i, w_i^p), -K^{(x_i,\mathrm{C})} - \lambda\xi^i \widehat{\mathrm{AR}}_{\Theta,i}(x_i, w_i^p) \right\}.$$
2. Perform a least squares regression on the VFA estimates in set $\{v_i^{\mathrm{A}}(x_i, w_i^p), p \in \mathcal{P}\}$ using basis functions Φ_i to determine the vector of VFA weights $\beta_{i,x_i}^{\mathrm{A}}$.

Outputs: Vectors of VFA weights $\beta_{i,x_i}^{\mathrm{A}}$ for each $(i, x_i) \in \{1, \ldots, I\} \times \mathcal{X}_i$.

Algorithm 1 summarizes the LSML steps to approximate the anticipated-regret SDP (5). Let $\beta_{i,x_i}^{\mathrm{A}} := (\beta_{i,x_i,1}^{\mathrm{A}}, \ldots, \beta_{i,x_i,B_i}^{\mathrm{A}})$ be the VFA weight vector for stage $i \in \{1, \ldots, I\}$ and operating state $x_i \in \mathcal{X}_i$. To ease notation, we do not subscript terms in LSML by Θ. The inputs to LSML are a set of P information state sample paths $\{w_i^p, (i,p) \in \mathcal{I} \cup \{I\} \times \mathcal{P}\}$ generated in Monte Carlo simulation, where $\mathcal{P} := \{1, \ldots, P\}$; the number of samples N used to construct sample average approximations of expectations; and a set of VFA basis functions. LSML initializes the terminal stage VFA weight vector $\beta_{I,x_I}^{\mathrm{A}}$ by regressing the basis functions Φ_I on evaluations of the terminal reward function for each $x_I \in \mathcal{X}_I$. At each stage $i \in \mathcal{I}$, starting from stage $I - 1$ and moving backward to stage 1, and for each operating state $x_i \in \mathcal{X}_i$ it executes Steps 1 and 2.

- In Step 1(a), for each sample path $p \in \mathcal{P}$, LSML generates N stage $i+1$ information state samples conditioned on w_i^p. We denote the n-th such sample by $\bar{w}_{i+1}^{p,n}$ and the set of these samples by $\{\bar{w}_{i+1}^{p,n}, n \in \mathcal{N}\}$, where $\mathcal{N} := \{1, \ldots, N\}$.
- In Step 1(b), it computes estimates $v_i^{\mathrm{A}}(x_i, w_i^p)$ of the stage i AR value function $V_{\Theta,i}^{\mathrm{A}}(x_i, w_i^p)$ by applying to the right hand sides of (5a) and (5c) the known stage $i+1$ VFA and sample average approximations of expectations based on the samples generated in the previous step.

Specifically, LSML computes $v_i^{A,\mathcal{O}}(x_i, w_i^p)$ by replacing $\mathbb{E}_i[V^A_{\Theta,i+1}(f_i(x_i,a_i), w_{i+1})]$ in the right hand side of (5a) by the sample average approximation

$$\sum_{b\in\mathcal{B}_{i+1}} \beta^A_{i+1,f_i(x_i,a_i),b}\left[\frac{1}{N}\sum_{n\in\mathcal{N}}\Phi_{i+1,b}(\bar{w}^{p,n}_{i+1})\right],$$

and obtains $v_i^A(x_i, w_i^p)$ by substituting for $V^{A,\mathcal{O}}_{\Theta,i}(x_i, w_i)$ and $\mathrm{AR}_{\Theta,i}(x_i, w_i^p)$ in the right hand side of (5c), respectively, with $v_i^{A,\mathcal{O}}(x_i, w_i^p)$ and the sample average approximation

$$\widehat{\mathrm{AR}}_{\Theta,i}(x_i, w_i^p) := \max_{a_i\in\mathcal{A}_i(x_i)\setminus\{\mathrm{C}\}} \frac{1}{N}\sum_{n\in\mathcal{N}} \max\left\{\hat{Q}(x_i, w_i^p, a_i, \bar{w}^{p,n}_{i+1}) + \eta\cdot K^{(x_i,\mathrm{C})}, 0\right\}, \tag{22}$$

where $\hat{Q}(x_i, w_i^p, a_i, \bar{w}^{p,n}_{i+1}) := r_i(x_i, w_i^p, a_i) + \delta\sum_{b\in\mathcal{B}_{i+1}}\beta^A_{i+1,f_i(x_i,a_i),b}\Phi_{i+1,b}(\bar{w}^{p,n}_{i+1})$.

- In Step 2, LSML performs a least squares (2-norm) regression on these estimates to determine the vector of VFA weights β^A_{i,x_i}.

The outputs of LSML are the vectors of VFA weights β^A_{i,x_i} for each stage $i \in \{1,\dots,I\}$ and operating state $x_i \in \mathcal{X}_i$. Given such VFA weights, the action $a_i^A(x_i, w_i)$ taken by the anticipated regret policy at stage i and state (x_i, w_i) is defined as

$$a_i^A(x_i,w_i) := \begin{cases} \mathrm{C}, \text{ if } v_i^{A,\mathcal{O}}(x_i,w_i) < -K^{(x_i,\mathrm{C})} - \lambda\xi^i\widehat{\mathrm{AR}}_{\Theta,i}(x_i,w_i), \\ \underset{a_i\in\mathcal{A}_i(x_i)\setminus\{\mathrm{C}\}}{\operatorname{argmax}}\left\{r_i(x_i,w_i,a_i) + \delta\sum_{b\in\mathcal{B}_{i+1}}\beta^A_{i+1,f_i(x_i,a_i),b}\left[\frac{1}{N}\sum_{n\in\mathcal{N}}\Phi_{i+1,b}(w^n_{i+1})\right]\right\}, \text{otherwise,}\end{cases}$$

where the sample average approximations are constructed using N next stage information state samples $(w^n_{i+1}, n\in\mathcal{N})$ conditioned on w_i.

Algorithm 1 can also be used with minor changes to approximate the shutdown-neutral SDPs (2) and (4). Specifically, computing a sample average approximation of the regret term $\widehat{\mathrm{AR}}_{\Theta,i}(x_i, w_i^p)$ as in (22) is not needed for the shutdown-neutral SDPs, thus, this term is replaced by zero.

C. Calibration of the price and exchange rate dynamics

For all commodities, we considered futures prices for weekly trading dates from November 18th 2008 to December 30th 2015. Aluminum and exchange rate futures contracts have physical/financial delivery at the end of the contract duration (that is, at maturity). We collected aluminum futures prices from Bloomberg for maturities extending out 1, 3, 6, 9, 12, 15, 18, 21, 24, and 36 months. Exchange rate futures for EUR-USD and NOK-USD on Bloomberg were only available for the months of March, June, September, and December. We applied standard linear interpolation to these rates and obtained a term structure curve with maturities that match the ones for aluminum contracts (Guthrie 2009, §12). Nord Pool power contracts were obtained from the information provider Montel (see www.montel.no). The power delivery duration of these contracts were monthly, quarterly, or yearly with trading dates extending out to 6 months, 8 quarters, and 3 years, respectively. In contrast to aluminum and exchange rate futures contracts, power contracts deliver electricity continuously during an interval of time. For example, the first and second nearest quarterly contracts traded in August 2015 delivered power during, respectively, the entire 4th quarter (Oct-Dec) of 2015 and 1st quarter (Jan-Mar) of 2016. Power futures contracts of different lengths can overlap, for instance, the nearest monthly and quarterly contracts. To obtain implied power futures prices of contracts that deliver only at maturity, we used the smoothing approach of Benth et al. (2007) (see Fleten and Lemming 2003 for an alternative smoothing technique), which determines a smooth polynomial spline that replicates the observed market prices for each trading date. This synthetic curve contains power futures contract prices that are consistent with the ones for aluminum and currency exchange rates.

We calibrated the parameters of the stochastic process (8) by applying a Kalman filter (Hamilton 1994, Chapter 13) to match the model-implied log-futures prices with the data by maximizing the log-likelihood function. The transition and measurement equations correspond to Farkas et al. (2017) equations (6)-(8) and (15)-(16), respectively. The integral terms involving matrix exponentials were evaluated numerically using the reductions to new matrix exponentials described in Carbonell et al. (2008). Due to the large number of unknown parameters, we followed a multi-step calibration process where parameters estimated in a given step are kept fixed in future steps, which is common (see, e.g., Farkas et al. 2017). These steps are the following:

1. We estimated electricity price seasonality $\psi^{\mathrm{EL}}(t)$ by regressing the function $\chi_1 \cos(2\pi t) + \chi_2 \sin(2\pi t)$ on the log-futures data (Paschke and Prokopczuk 2009, Farkas et al. 2017). We set the seasonality for aluminum and exchange rates, that is $\psi^{\mathrm{AL}}(t)$, $\psi^{\text{EUR-USD}}(t)$ and $\psi^{\text{NOK-USD}}(t)$, to zero as we observed no empirical evidence of seasonal effects.
2. We calibrated single-commodity 2-factor versions of model (8) for each commodity with a single correlation between short and long term factors. We estimated μ_z, the short and long term risk premia λ_y and λ_z, respectively, the diagonal terms of K_y, K_z and $\rho_{y,z}$, and the variance of the measurement error. The value of μ_z for currencies was not statistically significant and was set to zero. We noticed that the volatility estimates were unrealistically high. Hence σ_y and σ_z were replaced with data estimates directly from historical data.
3. We calibrated the cross-commodity correlation structure, that is, matrices ρ_y, ρ_z and ρ_{yz} keeping all previous estimates fixed. Doing this reduced the number of free variables to 24, which we could handle in the maximum likelihood estimation.

Below we report the parameter estimates where statistical significance at 1%, 5% and 10% are indicated by superscripts ***, ** and at *, respectively. The order of the commodities in the vectors and matrices is power, aluminum, EUR-USD rate and NOK-USD rate.

$$\hat{K}_y = \begin{bmatrix} 1.904^{***} & 0 & 0 & 0 \\ 0 & 0.067^{***} & 0 & 0 \\ 0 & 0 & 0.011^{***} & 0 \\ 0 & 0 & 0 & 0.010 \end{bmatrix}, \quad \hat{K}_z = \begin{bmatrix} 0.055^{***} & 0 & 0 & 0 \\ 0 & 0.184^{***} & 0 & 0 \\ 0 & 0 & 0.001 & 0 \\ 0 & 0 & 0 & 0.005 \end{bmatrix}$$

$$\hat{\mu} = \begin{bmatrix} 0.19^{*} \\ 1.41^{***} \\ 0 \\ 0 \end{bmatrix}, \quad \hat{\sigma}_y = \begin{bmatrix} 0.26 \\ 0.12 \\ 0.08 \\ 0.13 \end{bmatrix}, \quad \hat{\sigma}_z = \begin{bmatrix} 0.11 \\ 0.09 \\ 0.18 \\ 0.15 \end{bmatrix}, \quad \hat{\chi}_1 = \begin{bmatrix} 0.21 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \hat{\chi}_2 = \begin{bmatrix} 0.03 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\hat{\rho}_y = \begin{bmatrix} 1 & 0.79^{***} & 0.63^{***} & 0.65^{***} \\ 0.79^{***} & 1 & 0.72^{***} & 0.79^{***} \\ 0.63^{***} & 0.72^{***} & 1 & 0.83 \\ 0.65^{***} & 0.79^{***} & 0.83 & 1 \end{bmatrix}, \quad \hat{\rho}_z = \begin{bmatrix} 1 & 0.80^{***} & 0.68^{***} & -0.73^{*} \\ 0.80^{***} & 1 & 0.75^{***} & -0.72 \\ 0.68^{***} & 0.75^{***} & 1 & -0.79^{***} \\ -0.73^{*} & -0.72 & -0.79^{***} & 1 \end{bmatrix}$$

$$\hat{\rho}_{yz} = \begin{bmatrix} 0.25^{***} & 0.76^{***} & 0.71^{***} & -0.68^{***} \\ 0.81^{***} & 0.27^{***} & 0.77^{***} & -0.73^{*} \\ 0.61^{***} & 0.73 & 0.02^{***} & -0.55 \\ 0.65^{***} & 0.83^{***} & 0.65^{***} & -0.09^{***} \end{bmatrix} \quad \hat{\lambda}_y = \begin{bmatrix} -0.79^{***} \\ -0.59^{***} \\ -0.28^{***} \\ -0.50^{***} \end{bmatrix}, \quad \hat{\lambda}_z = \begin{bmatrix} 0.74^{***} \\ -0.14 \\ 3.49^{**} \\ -3.24^{*} \end{bmatrix}$$

Most of the estimates above are statistically significant. There is strong positive correlation between the short term factors of all commodities and some negative correlation in the long term factors, in particular between the first three assets and NOK-USD. The estimated covariance matrix Σ was computed as:

$$\Sigma = \begin{bmatrix} diag(\sigma_y) & 0 \\ 0 & diag(\sigma_z) \end{bmatrix} \begin{bmatrix} \rho_y & \rho_{yz} \\ \rho_{yz}^T & \rho_z \end{bmatrix} \begin{bmatrix} diag(\sigma_y) & 0 \\ 0 & diag(\sigma_z) \end{bmatrix}.$$

This matrix is in general not positive semi-definite (PSD), which is a property needed to generate multivariate random draws. Nonetheless, after our calibration Σ was very close to being PSD (the negative eigenvalues are smaller than 10^{-2}). Therefore, to gain the desired PSD property, we computed the nearest PSD matrix to Σ by minimizing the Frobenius norm of the difference (Higham 1988). This can be seen as a small perturbation of Σ. Finally, the initial values for the eight factors used in the analysis are based on representative spot and 1-year futures prices observed during the first quarter of 2017 and are, respectively, $Y_0 = \log\left[35,\ 1800,\ 1.10,\ 0.12\right]^\top$ and $Z_0 = \log\left[33,\ 1900,\ 1.25,\ 0.15\right]^\top$.

D. Dual bound on the asset value

To obtain a dual (upper) bound on the shutdown-neutral asset value, we use the information relaxation and duality framework of Brown et al. (2010). This approach relies on relaxing the non-anticipativity constraints embedded in the SDP, and penalizing knowledge of future information at time $i \in \mathcal{I}$ using a penalty function $q_i\left(f(x_i, a_i), w_i, w_{i+1}\right)$. A feasible dual penalty q_i satisfies $\mathbb{E}[q_i\left(f(x_i, a_i), w_i, w_{i+1}\right) \big| w_i] \leqslant 0$. The dual bound estimation process relies on H Monte Carlo samples of uncertainty $\{w_i^h,\ (i, h) \in \mathcal{I} \cup \{I\} \times \{1, \ldots, H\}\}$. We solve the following deterministic dynamic program on each sample path h:

$$U_I^h(x_I) = r_I(x_I, w_I^h),\ \ \forall x_I \in \mathcal{X}_I,$$
$$U_i^h(x_i) = \max_{a_i \in \mathcal{A}_i(x_i)} \left\{ r_i(x_i, w_i^h, a_i) - q_i\left(f(x_i, a_i), w_i^h, w_{i+1}^h\right) + \delta\, U_{i+1}^h\left(f(x_i, a_i)\right) \right\},\ \forall (i, x_i) \in \mathcal{I} \times \mathcal{X}_i,$$

for all $h \in \{1, \ldots, H\}$, where q_i is a feasible penalty. A dual bound on the option value is then obtained as the sample average $\sum_h U_0^h(x_0)/H$. It is well known (Brown et al. 2010) that given a VFA $\hat{V}_i(\cdot)$, a feasible dual penalty can be defined for $(x_{i+1}, h) \in \mathcal{X}_{i+1} \times \{1, \ldots, H\}$ as follows:

$$q_i\left(x_{i+1}, w_i^h, w_{i+1}^h\right) = \delta\left\{\hat{V}_{i+1}\left(x_{i+1}, w_{i+1}^h\right) - \mathbb{E}\left[\hat{V}_{i+1}\left(x_{i+1}, w_{i+1}\right) | w_i^h\right]\right\}.$$

We employed this penalty function in the computation of the dual bounds.

References

Benth, F. E., S. Koekkebakker, F. Ollmar. 2007. Extracting and applying smooth forward curves from average-based commodity contracts with seasonal variation. *The Journal of Derivatives* **15**(1) 52–66.

Brown, D. B., J. E. Smith, P. Sun. 2010. Information relaxations and duality in stochastic dynamic programs. *Operations Research* **58**(4) 785–801.

Carbonell, F., J. C. Jímenez, L. M. Pedroso. 2008. Computing multiple integrals involving matrix exponentials. *Journal of Computational and Applied Mathematics* **213**(1) 300–305.

Farkas, W., E. Gourier, R. Huitema, C. Necula. 2017. A two-factor cointegrated commodity price model with an application to spread option pricing. *Journal of Banking and Finance* **77**(1) 249–268.

Fleten, S. E., J. Lemming. 2003. Constructing forward price curves in electricity markets. *Energy Economics* **25**(5) 409–424.

Guthrie, G. A. 2009. *Real Options in Theory and Practice*. Oxford University Press, NY, USA.

Hamilton, J. D. 1994. *Time Series Analysis*. Princeton University Press, Princeton, NJ, USA.

Higham, N. J. 1988. Computing a nearest symmetric positive semidefinite matrix. *Linear Algebra and its Applications* **103**(1) 103–118.

Paschke, R., M. Prokopczuk. 2009. Integrating multiple commodities in a model of stochastic price dynamics. *Journal of Energy Markets* **2**(3) 47–82.