

Callegaro, Giorgia; Ceci, Claudia; Ferrari, Giorgio

Working Paper

Optimal reduction of public debt under partial observation of the economic growth

Center for Mathematical Economics Working Papers, No. 608

Provided in Cooperation with:

Center for Mathematical Economics (IMW), Bielefeld University

Suggested Citation: Callegaro, Giorgia; Ceci, Claudia; Ferrari, Giorgio (2019) : Optimal reduction of public debt under partial observation of the economic growth, Center for Mathematical Economics Working Papers, No. 608, Bielefeld University, Center for Mathematical Economics (IMW), Bielefeld, <https://nbn-resolving.de/urn:nbn:de:0070-pub-29333605>

This Version is available at:

<https://hdl.handle.net/10419/201633>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

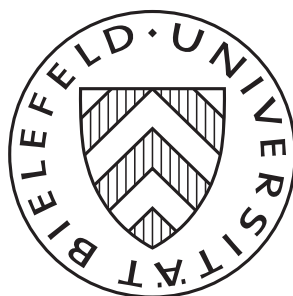
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

January 2019

Optimal Reduction of Public Debt under Partial Observation of the Economic Growth

Giorgia Callegaro, Claudia Ceci and Giorgio Ferrari



OPTIMAL REDUCTION OF PUBLIC DEBT UNDER PARTIAL OBSERVATION OF THE ECONOMIC GROWTH

GIORGIA CALLEGARO, CLAUDIA CECI, AND GIORGIO FERRARI

ABSTRACT. We consider a government that aims at reducing the debt-to-gross domestic product (GDP) ratio of a country. The government observes the level of the debt-to-GDP ratio and an indicator of the state of the economy, but does not directly observe the development of the underlying macroeconomic conditions. The government's criterion is to minimize the sum of the total expected costs of holding debt and of debt's reduction policies. We model this problem as a singular stochastic control problem under partial observation. The contribution of the paper is twofold. Firstly, we provide a general formulation of the model in which the level of debt-to-GDP ratio and the value of the macroeconomic indicator evolve as a diffusion and a jump-diffusion, respectively, with coefficients depending on the regimes of the economy. These are described through a finite-state continuous-time Markov chain. We reduce via filtering techniques the original problem to an equivalent one with full information (the so-called separated problem), and we provide a general verification result in terms of a related optimal stopping problem under full information. Secondly, we specialize to a case study in which the economy faces only two regimes, and the macroeconomic indicator has a suitable diffusive dynamics. In this setting we provide the optimal debt reduction policy. This is given in terms of the continuous free boundary arising in an auxiliary fully two-dimensional optimal stopping problem.

Keywords: singular stochastic control; partial observation; filtering; separated problem; optimal stopping; free boundary; debt-to-GDP ratio.

MSC2010 subject classification: 93E20, 60G35, 93E11, 60G40, 60J60, 91B64.

1. INTRODUCTION

The question of optimally managing debt-to-GDP ratio (also called “debt ratio”) of a country has become particularly important in the latest years. Indeed, concurrently with the financial crisis started in 2007, debt-to-GDP ratio exploded from an average of 53% to circa 80% in developed countries. Clearly, the debt management policy of a government highly depends on the underlying macroeconomic conditions; indeed, these affect, for example, the growth rate of GDP which, in turn, determines the growth rate of the debt-to-GDP ratio of a country. However, in practice it is typically neither possible to measure in real-time the growth rate of GDP, nor one can directly observe the underlying business cycles. On August 24, 2018, Jerome H. Powell, Chairman of the Federal Reserve, said:¹

*...In conventional models of the economy, major economic quantities such as inflation, unemployment, and the growth rate of gross domestic product fluctuate around values that are considered “normal”, or “natural” or “desired”. The FMO (Federal Open Market Committee) has chosen a 2 percent inflation objective as one of these desired values. **The other values are not directly observed, nor can they be chosen by anyone...***

Following an idea that dates back to [27], in this paper we suppose that the GDP growth rate of a country is modulated by a continuous-time Markov chain that is not directly observable. The Markov chain has $Q \geq 2$ states modeling the different business cycles of the economy, so that a shift in the macroeconomic conditions induces a change in the value of the growth rate of GDP. The government can observe only the current levels of the debt-to-GDP ratio and of a macroeconomic indicator. The latter might be, e.g., one of the so-called “Big Four”², which are usually considered proxies of the industrial production index, hence of the business conditions.

Date: January 24, 2019.

¹Speech at “Changing Market Structure and Implications for Monetary Policy”, a symposium sponsored by the Federal Reserve Bank of Kansas City in Jackson Hole, Wyoming.

²These indicators constitute the Conference Board's Index of Coincident Indicators; they are employment in non agricultural businesses, industrial production, real personal income less transfers, and real manufacturing and trade

The government may intervene in order to decrease the level of the debt ratio, e.g. through fiscal policies or imposing austerity policies in the form of spending cuts. We assume that the debt ratio is instantaneously affected by any such policy. Debt reductions must not necessarily be performed at rates, but also lump sum actions are allowed, and the cumulative amount of debt ratio's decrease is the government's control variable. Any decrease of the debt ratio results in proportional costs, and the government aims at choosing a debt-reduction policy that minimizes the total expected costs of holding debt, plus the total expected costs of interventions on the debt ratio. In line with recent papers on stochastic control methods for optimal debt management (see [5], [6], [22] and [23]), we model the previous problem as a singular stochastic control problem. However, differently to all the previous works, our problem is formulated in a partial observation setting, thus leading to a completely different mathematical analysis. In our model, the observations consist of the debt ratio and of the macroeconomic indicator. The debt ratio is a linearly controlled geometric Brownian motion, and its drift is given in terms of the GDP growth rate, which is modulated by the unobservable continuous-time Markov chain Z . The macroeconomic indicator is a real-valued jump-diffusion which is correlated to the debt ratio process, and which has drift, and both intensity and jump sizes, depending on Z .

Our Contributions. Our study of the optimal debt reduction problem is performed through three main steps.

First of all, via advanced filtering techniques with mixed-type observations, we reduce the original problem to an equivalent problem under full information, the so-called separated problem. The filtering problem consists in characterizing the conditional distribution of the unobservable Markov chain Z , at any time t , given observations up to time t . The case of diffusion observations has been widely studied in literature and textbook treatments can be found in [19], [31], and [36]. There are also known results for pure-jump observations (see, e.g., [4], [8], [9], [34], and references therein). More recently, filtering problems with mixed-type information, which involve pure-jump processes and diffusions, have been studied in [11], [12], and [24]. Due to the structure of our observations' dynamics we cannot apply the probability reference method (see [12] and [45]), and for this reason we choose an alternative route based on the innovation approach, which leads to the Kushner-Stratonovich equation. Moreover, differently to [11] and [24], in our framework the innovation process is two-dimensional and, therefore, the innovation method employed in these papers must be suitably adapted to our context. By showing that the Kushner-Stratonovich equation admits a unique strong solution, we are then able to prove that the original problem under partial observation and the separated problem are equivalent in the sense that they share the same value and the same optimal control.

Secondly, we exploit the convex structure of the separated problem, and we provide a general probabilistic verification theorem. This result - which is in line with findings in [2], [15] and [22], among others - relates the optimal control process to the solution to an auxiliary optimal stopping problem. Moreover, it proves that the value function of the separated problem is the integral - with respect to the controlled state variable - of the value function of the optimal stopping problem. The stopping problem thus gives the optimal timing at which one additional unit of debt should be reduced.

Finally, by specifying a setting in which the continuous-time Markov chain faces only two regimes (a fast growth or slow growth phase) and the macroeconomic indicator is a suitable diffusion process, we are able to characterize the optimal debt reduction policy. In this framework, the filter process is a two-dimensional process $(\pi_t, 1 - \pi_t)_{t \geq 0}$, where π_t is the conditional probability at time t that the economy enjoys the fast growth phase. We prove that the optimal control prescribes to keep at any time the debt ratio below an endogenously determined curve that is a function of government's belief about the current state of the economy. Such a debt ceiling is the free boundary of the fully two-dimensional optimal stopping problem that is related to the separated problem in the sense of the previously discussed verification theorem. By using almost exclusively probabilistic means, we are able to show that the value function of the auxiliary optimal stopping problem is a C^1 -function of its arguments, and thus enjoys the so-called smooth-fit property. Moreover, the free boundary is a

sales. We refer to, e.g., [43], where the authors present a wide range of economic indicators and examine the forecasting performance of various of them in the recession of 2001.

continuous, bounded, and increasing function of the filter process. This last monotonicity property has also a clear economic interpretation: the more the government believes that the economy enjoys a regime of fast growth, the less strict the optimal debt reduction policy should be.

As a remarkable byproduct of the regularity of the value function of the optimal stopping problem, we also obtain that the value function of the singular stochastic control problem is a classical solution to its associated Hamilton-Jacobi-Bellman (HJB) equation. The latter takes the form of a variational inequality involving an elliptic second-order partial differential equation (PDE). It is worth noticing that the C^2 regularity of the value function implies the validity of a second-order principle of smooth fit, usually observed in one-dimensional problems.

We believe that the study of the auxiliary fully two-dimensional optimal stopping problem is a valuable contribution to the literature on its own. Indeed, if the literature on one-dimensional optimal stopping problems is very rich, the problem of characterizing the optimal stopping rule in multi-dimensional settings has been so far rarely explored in the literature (see the recent [13], [15] and [30], among the very few papers dealing with multi-dimensional stopping problems). This discrepancy is due to the fact that a standard guess-and-verify approach, based on the construction of an explicit solution to the variational inequality arising in the considered optimal stopping problem, is not anymore applicable in multi-dimensional settings where the variational inequality involves a PDE rather than an ordinary differential equation.

Related Literature. As already noticed above, our paper is placed among those recent works addressing the problem of optimal debt management via continuous-time stochastic control techniques. In particular, [5] and [6] model an optimal debt reduction problem as a one-dimensional control problem with singular and bounded-velocity controls, respectively. In [23] the government is allowed to increase and decrease the current level of debt ratio, and the interest rate on debt is modulated by a continuous-time observable Markov chain. The mathematical formulation leads to a one-dimensional bounded-variation stochastic control problem with regime switching. In [22], when optimally reducing the debt ratio, the government takes into consideration the evolution of the inflation rate of the country. The latter evolves as an uncontrolled diffusion process and affects the growth rate of the debt ratio, which is a process of bounded variation. In this setting, the debt reduction problem is formulated as a two-dimensional singular stochastic control problem whose HJB equation involves a second-order linear parabolic partial differential equation. All the previous papers are formulated in a full information setting, while ours is under partial observation.

The literature on singular stochastic control problems under partial observation is also still quite limited. Theoretical results on the PDE characterization of the value function of a two-dimensional optimal correction problem under partial observation are obtained in [37], whereas a general maximum principle for a not necessarily Markovian singular stochastic control problem under partial information has been more recently derived in [38]. We also refer to [14] and [18], where it is provided a thorough study of the optimal dividend strategy in models in which the surplus process evolves as drifted Brownian motion with unknown drift that can take only two constant values, with given probability.

Outline of the Paper. The rest of the paper is organized as follows. In Section 2 we introduce the setting and formulate the problem. The reduction of the problem under partial observation to the separated problem is performed in Section 3; in particular, the filtering results are presented in Section 3.1. The probabilistic verification theorem connecting the separated problem to one of optimal stopping is then proved in Section 3.3. In Section 4 we then consider a case study in which the economy faces only two regimes. Its solution, presented in Sections 4.2 and 4.3, hinges on the study of a two-dimensional optimal stopping problem that is performed in Section 4.1. Finally, Appendix A collects the proofs of some technical filtering results.

2. SETTING AND PROBLEM FORMULATION

2.1. The Setting. Consider the complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$, capturing all the uncertainty of our setting. Here, $\mathbb{F} := (\mathcal{F}_t)_{t \geq 0}$ denotes the *full information filtration*. We suppose that such a filtration satisfies the usual hypotheses of completeness and right-continuity.

We denote by Z a continuous-time finite-state Markov chain describing the different states of the economy. For $Q \geq 2$, let $S := \{1, 2, \dots, Q\}$ be the state space of Z and $\{\lambda_{ij}\}_{1 \leq i, j \leq Q}$ its generator

matrix. Here, λ_{ij} , $i \neq j$, gives the intensity of a transition from state i to state j , and it is such that $\lambda_{ij} \geq 0$, for $i \neq j$, and $\sum_{j=1, j \neq i}^Q \lambda_{ij} = -\lambda_{ii}$. For any time $t \geq 0$, Z_t is $\mathcal{F}_t$ -measurable.

In absence of any intervention by the government, we assume that the (uncontrolled) debt-to-GDP ratio evolves as

$$(2.1) \quad dX_t^0 = (r - g(Z_t))X_t^0 dt + \sigma X_t^0 dW_t, \quad X_0^0 = x \in (0, \infty),$$

where W is a standard $\mathbb{F}$ -Brownian motion on $(\Omega, \mathcal{F})$ independent of Z , $r \geq 0$ and $\sigma > 0$ are constants, and $g : S \rightarrow \mathbb{R}$. The constant r is the real interest rate on debt, σ is the debt's volatility, and $g(i) \in \mathbb{R}$ is the rate of the GDP's growth when the economy is in state $i \in S$.

It is clear that equation (2.1) admits a unique strong solution, and, when needed, for any $x > 0$ we shall denote it by $X^{x,0}$. The current level of the debt-to-GDP ratio is known to the government at any time t , and $X^{x,0}$ is therefore the first component of the so-called *observation process*.

The government also observes a macroeconomic stochastic indicator η , e.g. one of the so-called “Big Four”, which we interpret as a proxy of the business conditions. We assume that η is a jump-diffusion process solving the stochastic differential equation

$$(2.2) \quad d\eta_t = b_1(\eta_t, Z_t)dt + \sigma_1(\eta_t)dW_t + \sigma_2(\eta_t)dB_t + c(\eta_{t-}, Z_{t-})dN_t, \quad \eta_0 = q \in \mathcal{I},$$

where b_1 , c , $\sigma_1 > 0$, and $\sigma_2 > 0$ are measurable functions of their arguments, and $\mathcal{I} \subseteq \mathbb{R}$ is the state space of η . Here, B is an $\mathbb{F}$ -standard Brownian motion, independent of W and Z . Moreover, N is an $\mathbb{F}$ -adapted point process, without common jump times with Z , independent of W and B . The predictable intensity of N is denoted by $\{\lambda^N(Z_{t-})\}_{t \geq 0}$ and depends on the current state of the economy, with $\lambda^N(\cdot) > 0$ being a measurable function. From now on, we assume the following assumptions that ensure strong existence and uniqueness of the solution to equation (2.2) (see [25], among others).

Assumption 2.1. *The functions $b_1 : \mathcal{I} \times S \rightarrow \mathbb{R}$, $\sigma_1 : \mathcal{I} \rightarrow (0, \infty)$, $\sigma_2 : \mathcal{I} \rightarrow (0, \infty)$, and $c : \mathcal{I} \times S \rightarrow \mathbb{R}$ are such that for any $i \in S$:*

- (i) *(Continuity) $b_1(\cdot, i)$, $\sigma_1(\cdot)$, $\sigma_2(\cdot)$ and $c(\cdot, i)$ are continuous;*
- (ii) *(Local Lipschitz conditions) for any $R > 0$, there exists a constant $L_R > 0$ such that if $|q| < R, |q'| < R$, $q, q' \in \mathcal{I}$, then*

$$|b_1(q, i) - b_1(q', i)| + |\sigma_1(q) - \sigma_1(q')| + |\sigma_2(q) - \sigma_2(q')| + |c(q, i) - c(q', i)| \leq L_R |q - q'|;$$

- (iii) *(Growth conditions) there exists a constant $C > 0$ such that*

$$|b_1(q, i)|^2 + |\sigma_1(q)|^2 + |\sigma_2(q)|^2 + |c(q, i)|^2 \leq C(1 + |q|^2).$$

The dynamics proposed in equation (2.2) is of a jump-diffusive type, and it allows for a size and intensity of the jumps affected by the state of the economy. It is therefore flexible enough to describe a large class of stochastic factors which may exhibit jumps.

The *observation filtration* $\mathbb{H} = (\mathcal{H}_t)_{t \geq 0}$ is defined as

$$(2.3) \quad \mathbb{H} := \mathbb{F}^{X^0} \vee \mathbb{F}^\eta,$$

where $\mathbb{F}^{X^0}$ and $\mathbb{F}^\eta$ denote the natural filtrations generated by X^0 and η , respectively, as usual augmented by $\mathbb{P}$ -null sets. Clearly, (X^0, η) is adapted to both $\mathbb{H}$ and $\mathbb{F}$, and

$$\mathbb{H} \subset \mathbb{F}.$$

The above inclusion means that the government cannot directly observe the state of the economy Z , but this has to be inferred through the observation of (X^0, η) . We are therefore working in a *partial information setting*.

2.2. The Optimal Debt Reduction Problem. The government can reduce the level of the debt-to-GDP ratio by intervening on the primary budget balance (i.e. the overall difference between government revenues and spending), for example through austerity policies in the form of spending cuts. By doing so the debt ratio dynamics modifies as

$$(2.4) \quad dX_t^\nu = (r - g(Z_t))X_t^\nu dt + \sigma X_t^\nu dW_t - d\nu_t, \quad X_{0-}^\nu = x > 0.$$

The process ν is the control that the government chooses based on the information at its disposal. Precisely, ν_t defines the cumulative reduction of the debt-to-GDP ratio made by the government up to time t , and ν is therefore a nondecreasing process belonging to the set

$$\mathcal{M}(x, \underline{y}, q) := \left\{ \nu : \Omega \times \mathbb{R}_+ \rightarrow \mathbb{R}_+ : (\nu_t(\omega) := \nu(\omega, t))_{t \geq 0} \text{ is nondecreasing, right-continuous,} \right.$$

$$\left. \mathbb{H} - \text{adapted, such that } X_t^\nu \geq 0 \text{ for every } t \geq 0, X_{0-}^\nu = x, \mathbb{P}(Z_0 = i) = y_i, i \in S, \eta_0 = q \text{ a.s.} \right\},$$

for any given and fixed $x \in (0, \infty)$ initial value of X^ν , $q \in \mathcal{I}$ initial value of η , and $\underline{y} \in \mathcal{Y}$. Here

$$\mathcal{Y} := \left\{ \underline{y} = (y_1, \dots, y_Q) : y_i \in [0, 1], i = 1, \dots, Q, \sum_{i=1}^Q y_i = 1 \right\},$$

is the probability simplex on $\mathbb{R}^Q$, representing the space of initial distributions of the process Z . From now on, we set $\nu_{0-} = 0$ a.s. for any $\nu \in \mathcal{M}(x, \underline{y}, q)$.

Remark 2.2. Notice that in the definition of the set $\mathcal{M}$ above, as well as in (2.6) and in (P1) below, we have stressed the dependency on the initial data $(x, \underline{y}, q)$ just for notational convenience, and not to stress any Markovian nature of the considered problem, which is in fact not such.

For any $(x, \underline{y}, q) \in (0, \infty) \times \mathcal{Y} \times \mathcal{I}$ and $\nu \in \mathcal{M}(x, \underline{y}, q)$, there exists a unique solution to (2.4), denoted by $X_t^{x, \nu}$, that is given by

$$(2.5) \quad X_t^{x, \nu} = X_t^{1,0} \left[x - \int_0^t \frac{d\nu_s}{X_s^{1,0}} \right], \quad t \geq 0, \quad X_{0-}^{x, \nu} = x,$$

where

$$X_t^{1,0} = e^{\int_0^t (r - g(Z_s)) ds - \frac{1}{2} \sigma^2 t + \sigma W_t}, \quad t \geq 0.$$

Here, and in the rest of this paper, we shall use the notation $\int_0^t (\cdot) d\nu_s = \int_{[0,t]} (\cdot) d\nu_s$ for the Lebesgue-Stieltjes integral with respect to the random measure $d\nu$ induced by the nondecreasing process ν on $[0, \infty)$.

Remark 2.3. The dynamics (2.4) might be justified in the following way. Suppose that the public debt (in real terms), D , and the GDP, Y , follow the classical dynamics

$$\begin{cases} dD_t = rD_t dt - d\xi_t, & D_{0-} = d > 0, \\ dY_t = g(Z_t)Y_t dt + \sigma Y_t dW_t, & Y_0 = y > 0, \end{cases}$$

where ξ_t is the cumulative real budget balance up to time t . An easy application of Itô's formula then gives that the ratio $X := D/Y$ evolves as in (2.4), upon setting $d\nu := d\xi/Y$.

The government aims at reducing the level of the debt ratio. Having a level of debt ratio $X_t = x$ at time $t \geq 0$ when the state of the economy is $Z_t = i$, the government incurs an instantaneous cost $h(x, i)$. This cost may be interpreted as a measure of the resulting losses for the country due to the debt, as, e.g., a tendency to suffer low subsequent growth (see [21] and [44], among others). The cost function $h : \mathbb{R} \times S \mapsto \mathbb{R}_+$ fulfills the following requirements (see also [5] and [22])

Assumption 2.4.

- (i) For any $i \in S$, the mapping $x \mapsto h(x, i)$ is strictly convex, continuously differentiable, and it is nondecreasing on $\mathbb{R}_+$. Moreover, $h(0, i) = 0$;
- (ii) For any given $x \in (0, \infty)$ and $i \in S$ one has

$$\mathbb{E} \left[\int_0^\infty e^{-\rho t} h(X_t^{x,0}, i) dt \right] + \mathbb{E} \left[\int_0^\infty e^{-\rho t} X_t^{1,0} h_x(X_t^{x,0}, i) dt \right] < \infty.$$

A quadratic cost function of the form $h(x, i) = \frac{1}{2} \vartheta_i x^2$, $(x, i) \in [0, \infty) \times S$, $\vartheta_i > 0$, clearly satisfies Assumption 2.4 for a suitable $\rho > 0$.

Whenever the government intervenes in order to reduce the debt-to-GDP ratio, it incurs a proportional cost. We assume that the marginal cost of each intervention is normalized to one.

Given an intertemporal discount rate $\rho > 0$, for any given and fixed $(x, \underline{y}, q) \in (0, \infty) \times \mathcal{Y} \times \mathcal{I}$, the government thus aims at minimizing the expected total cost functional

$$(2.6) \quad \mathcal{J}_{x, \underline{y}, q}(\nu) := \mathbb{E}_{(x, \underline{y}, q)} \left[\int_0^\infty e^{-\rho t} h(X_t^{x, \nu}, Z_t) dt + \int_0^\infty e^{-\rho t} d\nu_t \right], \quad \nu \in \mathcal{M}(x, \underline{y}, q).$$

Here $\mathbb{E}_{(x, \underline{y}, q)}$ is the expectation under the condition that $X_{0-}^{x, \nu} = x$, Z has initial distribution $\underline{y}$, and $\eta_0 = q$. The government's problem under partial observation can be therefore defined as

$$(P1) \quad V_{po}(x, \underline{y}, q) := \inf_{\nu \in \mathcal{M}(x, \underline{y}, q)} \mathcal{J}_{x, \underline{y}, q}(\nu), \quad (x, \underline{y}, q) \in (0, \infty) \times \mathcal{Y} \times \mathcal{I}.$$

One has that V_{po} is well defined and finite. Indeed, it is nonnegative, due to the nonnegativity of h ; moreover, since the admissible policy “instantaneously reduce at initial time the debt ratio to 0” is a priori suboptimal and it has cost x , then $V_{po} \leq x$.

We would like to stress once more that any $\nu \in \mathcal{M}(x, \underline{y}, q)$ is $\mathbb{H}$ -adapted, and therefore Problem (P1) is a stochastic optimal control problem under partial observation. In particular, it is a *singular stochastic control problem under partial observation*; that is, an optimal control problem in which the random measures induced by the nondecreasing control processes on $[0, \infty)$ might be singular with respect to the Lebesgue measure, and in which one component of the state variable, Z , is not directly observable by the controller.

In its current formulation, the optimal debt reduction problem is not Markovian and it is therefore not directly solvable by standard means of stochastic control theory. In the next section, by using techniques from filtering theory, we will introduce an equivalent problem under complete information, the so-called *separated problem*. This will enjoy a Markovian structure and its solution will be characterized in Section 3.3 through a Markovian optimal stopping problem.

3. REDUCTION TO AN EQUIVALENT PROBLEM UNDER COMPLETE INFORMATION

In this section we derive the *separated problem*. To this end, we first study the filtering problem arising in our model. As already discussed in the introduction, results on such a filtering problem cannot be directly obtained from existing literature due to the structure of our dynamics.

3.1. The Filtering Problem. The filtering problem consists in finding the best-mean squared estimate of $f(Z_t)$, for any t and any measurable function f , on the basis of the information collected up to time t . In our setting, such an information flow is given by the filtration $\mathbb{H}$. That estimate can be described through the filter process $(\pi_t)_{t \geq 0}$, which provides the conditional distribution of Z_t given $\mathcal{H}_t$, for any time t (see, for instance, [36]). It is defined as the $\mathbb{H}$ -càdlàg (right-continuous with left limits) process taking values in the space of probability measures on $S = \{1, \dots, Q\}$ such that

$$(3.1) \quad \pi_t(f) := \mathbb{E}[f(Z_t) | \mathcal{H}_t],$$

for all measurable functions f on S . Since Z takes only a finite number of values, the filter is completely described by the vector

$$(3.2) \quad \pi_t(f_i) = \mathbb{P}(Z_t = i | \mathcal{H}_t), \quad i \in S,$$

where $f_i(z) := \mathbb{1}_{\{z=i\}}$, $i \in S$. With a slight abuse of notation, in the following we will denote by $\pi(i)$ the process $\pi(f_i)$, so that for all measurable functions f we have from (3.1), $\pi_t(f) = \sum_{i=1}^Q f(i) \pi_t(i)$.

Setting $\beta(Z_t) := r - g(Z_t)$ and, accordingly, $\beta_i := r - g(i)$, $i \in S$, notice that β is clearly a bounded function. Then, define the two processes I and I^1 such that for any $t \geq 0$

$$(3.3) \quad I_t := W_t - \int_0^t \sigma^{-1}(\pi_s(\beta) - \beta(Z_s)) ds, \quad I_t^1 := B_t - \int_0^t (\pi_s(\alpha(\eta_s, \cdot)) - \alpha(\eta_s, Z_s)) ds,$$

where

$$(3.4) \quad \alpha(q, i) := \sigma_2(q)^{-1} \left\{ b_1(q, i) - \sigma^{-1} \beta(i) \sigma_1(q) \right\}, \quad (q, i) \in \mathcal{I} \times S.$$

Henceforth, we will work under the following Novikov's condition.

Assumption 3.1.

$$(3.5) \quad \mathbb{E} \left[e^{\frac{1}{2} \int_0^t \alpha^2(\eta_s, Z_s) ds} \right] < \infty, \quad \text{for any } t \geq 0.$$

Under Assumption 3.1, by classical results from filtering theory (see, e.g., [36]), the *innovation processes* I and I^1 are Brownian motions with respect to the filtration $\mathbb{H}$. Moreover, given the assumed independence of B and W , they turn out to be independent.

The integer-valued random measure associated to the jumps of η is defined as

$$(3.6) \quad m(dt, dq) := \sum_{s: \Delta\eta_s \neq 0} \delta_{(s, \Delta\eta_s)}(ds, dq),$$

where $\delta_{(a_1, a_2)}$ denotes the Dirac measure at point $(a_1, a_2) \in \mathbb{R}_+ \times \mathbb{R}$. Notice that the $\mathbb{H}$ -adapted random measure m is such that

$$(3.7) \quad \int_0^t c(\eta_{s-}, Z_{s-}) \mathbb{1}_{\{c(\eta_{s-}, Z_{s-}) \neq 0\}} dN_s = \int_0^t \int_{\mathbb{R}} q m(ds, dq), \quad t > 0.$$

To proceed further we need the following useful definitions.

Definition 3.2. (*$\mathbb{G}$ -Predictable Process indexed by $\mathbb{R}$*). Given any filtration $\mathbb{G}$, let $\mathcal{P}(\mathbb{G})$ denote the predictable σ -field on $(0, \infty) \times \Omega$ and $\mathcal{B}(\mathbb{R})$ the Borel σ -algebra on $\mathbb{R}$. Any mapping $H : (0, \infty) \times \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ which is $\mathcal{P}(\mathbb{G}) \times \mathcal{B}(\mathbb{R})$ -measurable is called $\mathbb{G}$ -predictable process indexed by $\mathbb{R}$.

Letting

$$(3.8) \quad \mathcal{F}_t^m := \sigma\{m((0, s] \times A) : 0 \leq s \leq t, A \in \mathcal{B}(\mathbb{R})\},$$

we denote by $\mathbb{F}^m := (\mathcal{F}_t^m)_{t \geq 0}$ the filtration generated by the random measure $m(dt, dq)$.

Definition 3.3. (*Dual Predictable Projection of m*). Given any filtration $\mathbb{G}$, such that $\mathbb{F}^m \subseteq \mathbb{G}$, the $\mathbb{G}$ -dual predictable projection of m , denoted by $m^{p, \mathbb{G}}(dt, dq)$, is defined as the unique positive $\mathbb{G}$ -predictable random measure such that for any nonnegative, $\mathbb{G}$ -predictable process Φ indexed by $\mathbb{R}$

$$(3.9) \quad \mathbb{E} \left[\int_0^\infty \int_{\mathbb{R}} \Phi(s, q) m(ds, dq) \right] = \mathbb{E} \left[\int_0^\infty \int_{\mathbb{R}} \Phi(s, q) m^{p, \mathbb{G}}(ds, dq) \right].$$

To prove that a positive $\mathbb{G}$ -predictable random measure provides the $\mathbb{G}$ -dual predictable projection of m it suffices to prove that equation (3.9) holds true for any process of the form $\Phi(t, q) = C_t \mathbb{1}_A(q)$ with C a nonnegative $\mathbb{G}$ -predictable process and $A \in \mathcal{B}(\mathbb{R})$. For further details we refer to [4] and [28].

We now aim at deriving an equation for the evolution of the filter (the *filtering equation*). To this end we use the so-called *innovation approach* (see [4], [11], and [36], among others), which, in our setting, requires the introduction of the $\mathbb{H}$ -compensated jump measure of η

$$(3.10) \quad m^\pi(dt, dq) := m(dt, dq) - m^{p, \mathbb{H}}(dt, dq),$$

where $m^{p, \mathbb{H}}(dt, dq)$ is the $\mathbb{H}$ -dual predictable projection of m (cf. Definition 3.3 above). The triplet (I, I^1, m^π) also represents the building block of the construction of $\mathbb{H}$ -martingales, as it is shown in Proposition 3.5 below. We start determining the form of $m^{p, \mathbb{H}}$.

Proposition 3.4. *The $\mathbb{H}$ -dual predictable projection of m is given by*

$$(3.11) \quad m^{p, \mathbb{H}}(dt, dq) = \sum_{i=1}^Q \pi_{t-}(i) \lambda^N(i) \mathbb{1}_{\{c(\eta_{t-}, i) \neq 0\}} \delta_{c(\eta_{t-}, i)}(dq) dt,$$

where δ_a denotes the Dirac measure at point $a \in \mathbb{R}$.

Proof. Step 1. First, we prove that the $\mathbb{F}$ -dual predictable projection of m is given by

$$(3.12) \quad m^{p, \mathbb{F}}(dt, dq) := \lambda^N(Z_{t-}) \mathbb{1}_{\{c(\eta_{t-}, Z_{t-}) \neq 0\}} \delta_{c(\eta_{t-}, Z_{t-})}(dq) dt.$$

Let $A \in \mathcal{B}(\mathbb{R})$ and introduce

$$(3.13) \quad \mathcal{N}_t(A) := m((0, t] \times A) = \sum_{s \leq t} \mathbb{1}_{\{\Delta\eta_s \in A \setminus \{0\}\}}, \quad t \geq 0.$$

$\mathcal{N}(A)$ is the point process counting the number of jumps of η up to time t with jumps' size in the set A . Since by (2.2) one has that $\Delta\eta_s = c(\eta_{s-}, Z_{s-}) \mathbb{1}_{\{c(\eta_{s-}, Z_{s-}) \neq 0\}} \Delta N_s$, $\forall s \geq 0$, and N is a point

process with $\mathbb{F}$ -predictable intensity given by $\{\lambda^N(Z_{t-})\}_{t \geq 0}$, we obtain that for each C nonnegative $\mathbb{F}$ -predictable process

$$\mathbb{E} \left[\int_0^t C_s d\mathcal{N}_s(A) \right] = \mathbb{E} \left[\int_0^t C_s \mathbb{1}_{\{c(\eta_{s-}, Z_{s-}) \in A \setminus \{0\}\}} dN_s \right] = \mathbb{E} \left[\int_0^t C_s \mathbb{1}_{\{c(\eta_{s-}, Z_{s-}) \in A \setminus \{0\}\}} \lambda^N(Z_{s-}) ds \right].$$

That is, for any $A \in \mathcal{B}(\mathbb{R})$, we have that $\{\lambda^N(Z_{t-}) \mathbb{1}_{\{c(\eta_{t-}, Z_{t-}) \in A \setminus \{0\}\}}\}_{t \geq 0}$ provides the $\mathbb{F}$ -predictable intensity of the counting process $\mathcal{N}(A)$. Recalling (3.13) and Definition 3.3, this implies that $m^{p, \mathbb{F}}(dt, dq)$ given in (3.12) coincides with the $\mathbb{F}$ -dual predictable projection of m , since equation (3.9) holds with the choice $\mathbb{G} = \mathbb{F}$ and $\Phi(t, q) = C_t \mathbb{1}_A(q)$, with C an arbitrary nonnegative $\mathbb{F}$ -predictable process and $A \in \mathcal{B}(\mathbb{R})$.

Step 2. As in Proposition 2.3 in [10] we can now derive the $\mathbb{H}$ -dual predictable projection of $m^{p, \mathbb{F}}$, denoted by $m^{p, \mathbb{H}}(dt, dq)$, by simply projecting $m^{p, \mathbb{F}}$ with respect to the observation flow $\mathbb{H}$. Precisely, we have that the $\mathbb{H}$ -predictable intensity of the point process $\mathcal{N}(A)$, $\forall A \in \mathcal{B}(\mathbb{R})$, is given by

$$\pi_{t-}(\lambda^N(\cdot) \mathbb{1}_{\{c(\eta_{t-}, \cdot) \in A \setminus \{0\}\}}) = \sum_{i=1}^Q \pi_{t-}(i) \lambda^N(i) \mathbb{1}_{\{c(\eta_{t-}, i) \in A \setminus \{0\}\}}, \quad \forall A \in \mathcal{B}(\mathbb{R}).$$

This implies that $m^{p, \mathbb{H}}(dt, dq)$ is given by (3.11), since (3.9) holds with the choice $\mathbb{G} = \mathbb{H}$, $\Phi(t, q) = C_t \mathbb{1}_A(q)$, with $A \in \mathcal{B}(\mathbb{R})$ and C an arbitrary nonnegative $\mathbb{H}$ -predictable process. $\square$

An essential tool to prove that the original problem under partial information is equivalent to the separated one is the characterization of the filter as the unique solution to the filtering equation (see [7] and [20]). In order to derive the filtering equation solved by π we first give a representation theorem for $\mathbb{H}$ -martingales. The proof of the following technical result is given in Appendix A.

Proposition 3.5. *Under Assumptions 2.1 and 3.1, every $\mathbb{H}$ -local martingale M admits the decomposition*

$$M_t = M_0 + \int_0^t \varphi_s dI_s + \int_0^t \psi_s dI_s^1 + \int_0^t \int_{\mathbb{R}} w(s, q) m^\pi(ds, dq),$$

where φ and ψ are $\mathbb{H}$ -adapted processes, and w is an $\mathbb{H}$ -predictable process indexed by $\mathbb{R}$ such that a.s.

$$\int_0^t \varphi_s^2 ds < \infty, \quad \int_0^t \psi_s^2 ds < \infty, \quad \int_0^t \int_{\mathbb{R}} |w(s, q)| m^{p, \mathbb{H}}(dt, dq) < \infty, \quad t \geq 0.$$

We are now in the position to prove the following fundamental result, whose proof is postponed to Appendix A.

Theorem 3.6. *Recall (3.10), let $y \in \mathcal{Y}$ be the initial distribution of Z , and let Assumptions 2.1 and 3.1 hold. Then the filter $\underline{\pi}_t := (\pi_t(i); i \in S)_{t \geq 0}$ solves the Kushner-Stratonovich system*

$$\begin{aligned} \pi_t(i) &= y_i + \int_0^t \sum_{j=1}^Q \lambda_{ji} \pi_s(j) ds + \int_0^t \pi_s(i) \sigma^{-1} \left\{ \beta_i - \sum_{j=1}^Q \beta_j \pi_s(j) \right\} dI_s \\ &+ \int_0^t \pi_s(i) \left\{ \alpha(\eta_s, i) - \sum_{j=1}^Q \alpha(\eta_s, j) \pi_s(j) \right\} dI_s^1 + \int_0^t \int_{\mathbb{R}} (w_i^\pi(s, q) - \pi_{s-}(i)) m^\pi(ds, dq), \end{aligned} \quad (3.14)$$

for any $i \in S$. Here, $\beta_i = r - g(i)$ and

$$w_i^\pi(s, q) := \frac{d\lambda^N(i) \pi_{s-}(i) \mathbb{1}_{\{c(\eta_{s-}, i) \neq 0\}} \delta_{c(\eta_{s-}, i)}(dq)}{d \left[\sum_{j=1}^Q \pi_{s-}(j) \lambda^N(j) \mathbb{1}_{\{c(\eta_{s-}, j) \neq 0\}} \delta_{c(\eta_{s-}, j)}(dq) \right]} \quad (3.15)$$

denotes the Radon-Nikodym derivative of the measure $\lambda^N(i) \pi_{s-}(i) \mathbb{1}_{\{c(\eta_{s-}, i) \neq 0\}} \delta_{c(\eta_{s-}, i)}(dq)$ with respect to $\sum_{j=1}^Q \pi_{s-}(j) \lambda^N(j) \mathbb{1}_{\{c(\eta_{s-}, j) \neq 0\}} \delta_{c(\eta_{s-}, j)}(dq)$.

Let us introduce the sequence of jump times and jump sizes of the process η , denoted by $\{T_n, \zeta_n\}_{n \geq 1}$, and recursively defined as $T_1 := \inf\{t > 0 : \int_0^t c(\eta_{s-}, Z_{s-}) dN_s \neq 0\}$,

$$T_{n+1} := \inf\{t > T_n : \int_{T_n}^t c(\eta_{s-}, Z_{s-}) dN_s \neq 0\}, \quad \zeta_n := \eta_{T_n} - \eta_{T_n-} = c(\eta_{T_n-}, Z_{T_n-}), \quad n \geq 1.$$

In the definitions above we use the standard convention that $\inf \emptyset = +\infty$.

Then the integer-valued measure associated to the jumps of η (cf. (3.6)) can also be written as

$$(3.16) \quad m(dt, dq) = \sum_{n \geq 1} \delta_{(T_n, \zeta_n)}(ds, dq) \mathbb{1}_{\{T_n < +\infty\}}.$$

The filtering system of equations (3.14) has a natural recursive structure in terms of the sequence $\{T_n\}_{n \geq 1}$, as it is shown in the next proposition.

Proposition 3.7. *Between two consecutive jump times, $t \in [T_n, T_{n+1})$, the filtering system of equations (3.14) reads as*

$$(3.17) \quad \begin{aligned} \pi_t(i) &= \pi_{T_n}(i) + \int_{T_n}^t \left\{ \sum_{j=1}^Q \lambda_{ji} \pi_s(j) - \pi_s(i) \left[\lambda^N(i) \mathbb{1}_{\{c(\eta_{s-}, i) \neq 0\}} - \sum_{j=1}^Q \lambda^N(j) \pi_s(j) \mathbb{1}_{\{c(\eta_{s-}, j) \neq 0\}} \right] \right\} ds \\ &\quad + \int_{T_n}^t \sigma^{-1} \pi_s(i) \left\{ \beta_i - \sum_{j=1}^Q \beta_j \pi_s(j) \right\} dI_s + \int_{T_n}^t \pi_s(i) \left\{ \alpha(\eta_s, i) - \sum_{j=1}^Q \alpha(\eta_s, j) \pi_s(j) \right\} dI_s^1, \end{aligned}$$

for any $i \in S$. At a jump time of η , say T_n , $\pi_t = (\pi_t(i); i \in S)_{t \geq 0}$ jumps as well, and its value is given by

$$(3.18) \quad \pi_{T_n}(i) = \frac{\lambda^N(i) \pi_{T_n^-}(i) \mathbb{1}_{\{\zeta_n = c(\eta_{T_n^-}, i)\}}}{\sum_{j=1}^Q \lambda^N(j) \pi_{T_n^-}(j) \mathbb{1}_{\{\zeta_n = c(\eta_{T_n^-}, j)\}}}, \quad i \in S.$$

Proof. First, recalling that $m^\pi(dt, dq) = m(dt, dq) - m^{p, \mathbb{H}}(dt, dq)$, and that

$$m^{p, \mathbb{H}}(dt, dq) = \sum_{j=1}^Q \pi_{t-}(j) \lambda^N(j) \mathbb{1}_{\{c(\eta_{t-}, j) \neq 0\}} \delta_{c(\eta_{t-}, j)}(dq) dt,$$

we obtain that

$$\int_0^t \int_{\mathbb{R}} (w_i^\pi(s, q) - \pi_{s-}(i)) m^{p, \mathbb{H}}(ds, dq) = \int_0^t \pi_s(i) \left[\lambda^N(i) \mathbb{1}_{\{c(\eta_{s-}, i) \neq 0\}} - \sum_{j=1}^Q \lambda^N(j) \pi_s(j) \mathbb{1}_{\{c(\eta_{s-}, j) \neq 0\}} \right] ds,$$

which, from (3.14), implies that for any $t \in [T_n, T_{n+1})$, $\pi_t(i)$ solves equation (3.17).

Finally, equation (3.18) follows by (3.15) and

$$\pi_{T_n}(i) = w_i^\pi(T_n, \zeta_n) = \frac{\lambda^N(i) \pi_{T_n^-}(i) \mathbb{1}_{\{c(\eta_{T_n^-}, i) \neq 0\}} \delta_{c(\eta_{T_n^-}, i)}(\zeta_n)}{\sum_{j=1}^Q \pi_{T_n^-}(j) \lambda^N(j) \mathbb{1}_{\{c(\eta_{T_n^-}, j) \neq 0\}} \delta_{c(\eta_{T_n^-}, j)}(\zeta_n)}.$$

□

We want to stress that equation (3.18) shows that the vector π_{T_n} is completely determined by the observed data η and by the knowledge of π_t for $t \in [T_{n-1}, T_n)$, since $\pi_{T_n^-}(i) := \lim_{t \uparrow T_n} \pi_t(i)$, $i \in S$.

Remark 3.8. *A few comments on the filtering equation are worth being done.*

- (1) In the case $c(q, i) \equiv c \neq 0$, for any $i \in S$ and $q \in \mathcal{I}$, the sequences of jump times of η and N coincide, and the filtering system of equations (3.14) reduces to the simpler

$$\begin{aligned} \pi_t(i) &= y_i + \int_0^t \sum_{j=1}^Q \lambda_{ji} \pi_s(j) ds + \int_0^t \pi_s(i) \sigma^{-1} \left\{ \beta_i - \sum_{j=1}^Q \beta_j \pi_s(j) \right\} dI_s \\ &\quad + \int_0^t \pi_s(i) \left\{ \alpha(\eta_s, i) - \sum_{j=1}^Q \alpha(\eta_s, j) \pi_s(j) \right\} dI_s^1 \\ &\quad + \int_0^t \left[\frac{\lambda^N(i) \pi_{s-}(i)}{\sum_{j=1}^Q \pi_{s-}(j) \lambda^N(j)} - \pi_{s-}(i) \right] \left(dN_s - \sum_{j=1}^Q \pi_{s-}(j) \lambda^N(j) ds \right), \quad i \in S. \end{aligned}$$

(2) In the case $\alpha(q, i) = \alpha(i)$ and $c(q, i) \equiv 0$, for any $i \in S$ and $q \in \mathcal{I}$, the filtering system of equations (3.14) does not depend anymore explicitly on the process η . In particular, one has

$$(3.19) \quad \begin{aligned} \pi_t(i) = & y_i + \int_0^t \sum_{j=1}^Q \lambda_{ji} \pi_s(j) ds + \int_0^t \pi_s(i) \sigma^{-1} \left\{ \beta_i - \sum_{j=1}^Q \beta_j \pi_s(j) \right\} dI_s \\ & + \int_0^t \pi_s(i) \left\{ \alpha_i - \sum_{j=1}^Q \alpha_j \pi_s(j) \right\} dI_s^1, \quad i \in S, \end{aligned}$$

where we have set $\alpha_i := \sigma_2^{-1} \{b_1(i) - \sigma^{-1} \beta_i \sigma_1\}$. With reference to (2.2) and (3.4), this setting corresponds, e.g., to the purely diffusive arithmetic case $c(q, i) = 0$, $b_1(q, i) = b_1(i)$ and $\sigma_1(q) = \sigma_1 > 0$, $\sigma_2(q) = \sigma_2 > 0$, for any $i \in S$ and $q \in \mathcal{I}$, or to the purely diffusive geometric case $c(q, i) = 0$, $b_1(q, i) = b_1(i)q$ and $\sigma_1(q) = \sigma_1 q$, $\sigma_2(q) = \sigma_2 q$, for any $i \in S$ and $q \in \mathcal{I}$. In Section 4 we will provide the explicit solution to the optimal debt reduction problem within this setting.

3.2. The Separated Problem. Thanks to the introduction of the filter, equations (2.1), (2.2), and (2.4) can now be rewritten in terms of observable processes. In particular, we have that

$$(3.20) \quad dX_t^0 = \pi_t(\beta) X_t^0 dt + \sigma X_t^0 dI_t, \quad X_0^0 = x > 0,$$

$$(3.21) \quad d\eta_t = \pi_t(b_1(\eta_t, \cdot)) dt + \sigma_1(\eta_t) dI_t + \sigma_2(\eta_t) dI_t^1 + \int_{\mathbb{R}} \zeta m(dt, d\zeta), \quad \eta_0 = q \in \mathcal{I},$$

and

$$(3.22) \quad dX_t^\nu = \pi_t(\beta) X_t^\nu dt + \sigma X_t^\nu dI_t - d\nu_t, \quad X_{0-}^\nu = x > 0.$$

Notice that, for any $\nu \in \mathcal{M}(x, \underline{y}, q)$, the process X^ν turns out to be $\mathbb{H}$ -adapted, and depends on the vector $\underline{\pi}_t = (\pi_t(i); i \in S)_{t \geq 0}$, such that $\underline{\pi}_0 = \underline{y} \in \mathcal{Y}$.

Definition 3.9. (Strong Uniqueness). We say that a process $(\tilde{\pi}_t, \tilde{\eta}_t)_{t \geq 0}$ with values in $\mathcal{Y} \times \mathcal{I}$ is a strong solution to equations (3.14) and (3.21) if it satisfies pathwise those equations. We say that strong uniqueness for the system of equations (3.14) and (3.21) holds if, for any $(\tilde{\pi}_t, \tilde{\eta}_t)_{t \geq 0}$ strong solution to system (3.14) and (3.21), one has $\tilde{\pi}_t = \underline{\pi}_t$, $\tilde{\eta}_t = \eta_t$, a.s. for all $t \geq 0$.

Proposition 3.10. Let Assumptions 2.1 and 3.1 hold, and suppose that $\alpha(\cdot, i)$ is locally-Lipschitz for any $i \in S$, and there exists $M > 0$ such that $|\alpha(q, i)| \leq M(1 + |q|)$, for any $q \in \mathcal{I}$ and any $i \in S$. Then system (3.14) and (3.21) admits a unique strong solution.

Notice that, under Assumption 2.1, the requirement on α of Proposition 3.10 is verified, e.g., whenever $\sigma_2(q) \geq \kappa$, for some $\kappa > 0$ and for any $q \in \mathcal{I}$, or if b_1/σ_2 and σ_1/σ_2 are locally-Lipschitz on $q \in \mathcal{I}$ and have sublinear growth. The proof of Proposition 3.10 is postponed to Appendix A. As a byproduct, it also ensures strong uniqueness of the solution to (3.22). In the following, when there will be the need to stress the dependence with respect to the initial value $x > 0$, we shall denote the solution to (3.20) and (3.22) by $X^{x,0}$ and $X^{x,\nu}$, respectively.

Since

$$\mathbb{E} [\pi_t(h(X_t^{x,\nu}, \cdot))] = \mathbb{E} [\mathbb{E} [h(X_t^{x,\nu}, Z_t) | \mathcal{H}_t]],$$

an application of Fubini-Tonelli's theorem allows to rewrite also the cost functional of (2.6) in terms of observable quantities as

$$(3.23) \quad \mathcal{J}_{x,\underline{y},q}(\nu) = \mathbb{E}_{(x,\underline{y},q)} \left[\int_0^\infty e^{-\rho t} \pi_t(h(X_t^\nu, \cdot)) dt + \int_0^\infty e^{-\rho t} d\nu_t \right].$$

Here $\mathbb{E}_{(x,\underline{y},q)}$ denotes the expectation conditioned on $X_{0-}^\nu = x > 0$, $\underline{\pi}_0 = \underline{y} \in \mathcal{Y}$, and $\eta_0 = q \in \mathcal{I}$. Notice that the latter expression does not depend anymore on the unobservable process Z , and this allows us to introduce a control problem with complete information, the *separated problem*, in which the new state variable is given by the triplet $(X^\nu, \underline{\pi}, \eta)$. For this problem we rewrite the set $\mathcal{M}(x, \underline{y}, q)$

in terms of the observable processes given by (3.14), (3.21) and (3.22), and we denote by $\mathcal{A}(x, \underline{y}, q)$ such a representation of the set $\mathcal{M}(x, \underline{y}, q)$; that is,

$$\mathcal{A}(x, \underline{y}, q) := \left\{ \nu : \Omega \times \mathbb{R}_+ \rightarrow \mathbb{R}_+ : (\nu_t(\omega) := \nu(\omega, t))_{t \geq 0} \text{ is non decreasing, right-continuous,} \right. \\ \left. \mathbb{H}\text{-adapted, such that } X_t^{x, \nu} \geq 0 \ \forall t \geq 0, \ X_{0-}^{x, \nu} = x, \ \pi_{0-} = \underline{y}, \ \eta_0 = q \text{ a.s.} \right\},$$

for every $x \in (0, \infty)$ initial value of $X^{x, \nu}$ defined in (3.22), for any $\underline{y} \in \mathcal{Y}$ initial values of the process $\pi_t = (\pi_t(i); i \in S)_{t \geq 0}$ solution to equation (3.14), and for any $q \in \mathcal{I}$ initial value of η . In the following, we set $\nu_{0-} = 0$ a.s. for any $\nu \in \mathcal{A}(x, \underline{y}, q)$.

Given $\nu \in \mathcal{A}(x, \underline{y}, q)$, the triplet $\{(X_t^{x, \nu}, \pi_t, \eta_t)\}_{t \geq 0}$ solves (3.22), (3.14) and (3.21) and the jump measure associated to η has $\mathbb{H}$ -predictable dual projection given by equation (3.11). Hence, the process $\{(X_t^{x, \nu}, \pi_t, \eta_t)\}_{t \geq 0}$ is an $\mathbb{H}$ -Markov process and we therefore define the Markovian separated problem as

$$(\mathbf{P2}) \begin{cases} V(x, \underline{y}, q) := \inf_{\nu \in \mathcal{A}(x, \underline{y}, q)} \mathbb{E}_{(x, \underline{y}, q)} \left[\int_0^\infty e^{-\rho t} \pi_t(h(X_t^\nu, \cdot)) dt + \int_0^\infty e^{-\rho t} d\nu_t \right] & \text{with} \\ dX_t^{x, \nu} = \pi_t(\beta) X_t^{x, \nu} dt + \sigma X_t^{x, \nu} dI_t - d\nu_t, \quad X_{0-}^{x, \nu} = x > 0, \\ (\pi, \eta) \text{ solution to equations (3.14) and (3.21).} \end{cases}$$

This is now a *singular stochastic problem under complete information*, since all the processes involved are $\mathbb{H}$ -adapted.

The next proposition immediately follows from the previous construction of the separated problem, and from the strong uniqueness of the solutions to (3.14), (3.21), and (3.22).

Proposition 3.11. *Assume strong uniqueness for the system of equations (3.14) and (3.21), and let $(x, \underline{y}, q) \in (0, \infty) \times \mathcal{Y} \times \mathcal{I}$ be the initial values of the process (X, Z, η) in the problem under partial observation (P1). Then*

$$(3.24) \quad V_{po}(x, \underline{y}, q) = V(x, \underline{y}, q).$$

Moreover, $\nu^* \in \mathcal{A}(x, \underline{y}, q)$ is an optimal control for the separated problem (P2) if and only if $\nu^* \in \mathcal{M}(x, \underline{y}, q)$ is an optimal control for the original problem under partial observation (P1).

Remark 3.12. Notice that in the setting of Remark 3.8-(2), the pair $(X^{x, \nu}, \pi)$ solving equations (3.22) and (3.14), respectively, is an $\mathbb{H}$ -Markov process, for any $\nu \in \mathcal{A}(x, \underline{y}, q)$, $(x, \underline{y}, q) \in (0, \infty) \times \mathcal{Y} \times \mathcal{I}$. As a consequence, since the cost functional and the set of admissible controls do not depend explicitly on the process η , the value function of the separated problem (P2) does not depend anymore on the variable q . We will consider this setting as a case study in Section 4.

3.3. A Probabilistic Verification Theorem via Reduction to Optimal Stopping. In this section we relate the separated problem to a Markovian optimal stopping problem, and we show that the solution to the latter is directly related to the optimal control of the former. The following analysis is fully probabilistic and it is based on a change of variable formula for Lebesgue-Stieltjes integrals that has been already employed in singular control problems (see, e.g., [2] and [22]). The result of this section will then be employed in Section 4 where, in a case study, we determine the expression of the optimal debt reduction policy by solving an auxiliary optimal stopping problem.

With regard to Problem (P2), notice that $\pi_t(h(X_t^{x, \nu}, \cdot)) = \sum_{i=1}^Q \pi_t(i) h(X_t^{x, \nu}, i)$ a.s. for any $t \geq 0$. Then, for any $(x, \pi) \in (0, \infty) \times \mathcal{Y}$, set

$$(3.25) \quad \hat{h}(x, \pi) := \sum_{i=1}^Q \pi(i) h(x, i),$$

and, given $z \in (0, \infty)$, we introduce the optimal stopping problem

$$(3.26) \quad \tilde{U}_t(z) := \operatorname{ess\,inf}_{\tau \geq t} \mathbb{E} \left[\int_t^\tau e^{-\rho(s-t)} X_s^{1,0} \hat{h}_x(X_s^{z,0}, \pi_s) ds + e^{-\rho(\tau-t)} X_\tau^{1,0} \mid \mathcal{H}_t \right], \quad t \geq 0,$$

where the optimization is taken over all the $\mathbb{H}$ -stopping times $\tau \geq t$.

Under Assumption 2.4, the expectation in (3.26) is finite for any $\mathbb{H}$ -stopping time $\tau \geq t$, for any $t \geq 0$. To take care of the event $\{\tau = \infty\}$, in (3.26) we make use of the convention

$$(3.27) \quad e^{-\rho\tau} X_\tau^{1,0} := \liminf_{t \uparrow \infty} e^{-\rho t} X_t^{1,0} \quad \text{on } \{\tau = \infty\}.$$

Denote by $U_t(z)$ a càdlàg modification of $\tilde{U}_t(z)$, and observe that $0 \leq U_t(z) \leq X_t^{1,0}$, for any $t \geq 0$, a.s. Also, define the stopping time

$$(3.28) \quad \tau_t^*(z) := \inf\{s \geq t : U_s(z) \geq X_s^{1,0}\}, \quad z \in (0, \infty),$$

with the convention that $\tau_t^*(z) = \infty$ if the set on the right-hand side is empty. Then by Theorem D.12 in Appendix D of [33], $\tau_t^*(z)$ is an optimal stopping time for problem (3.26). In particular, $\tau^*(z) := \tau_0^*(z)$ is optimal for the problem

$$(3.29) \quad U_0(z) := \inf_{\tau \geq 0} \mathbb{E} \left[\int_0^\tau e^{-\rho t} X_t^{1,0} \hat{h}_x(X_t^{z,0}, \underline{\pi}_t) dt + e^{-\rho\tau} X_\tau^{1,0} \right]$$

Notice that since $h_x(\cdot, \underline{\pi})$ is a.s. increasing, then $z \mapsto \tau^*(z)$ is a.s. decreasing. Such monotonicity of $\tau^*(\cdot)$ will be important in the following as we will need to consider its generalized inverse. Moreover, since the triplet $(X_t^{z,0}, \underline{\pi}_t, \eta_t)$ is an homogenous $\mathbb{H}$ -Markov process, there exists a measurable function $U : (0, \infty) \times \mathcal{Y} \times \mathcal{I} \rightarrow \mathbb{R}$ such that $U_t(z) = U(X_t^{z,0}, \underline{\pi}_t, \eta_t)$ for any $t \geq 0$, a.s. Hence, $U_0(z) = U(z, \underline{y}, q)$, and for any $(x, \underline{y}, q) \in (0, \infty) \times \mathcal{Y} \times \mathcal{I}$, define

$$(3.30) \quad \tilde{V}(x, \underline{y}, q) := \int_0^x U(z, \underline{y}, q) dz.$$

Moreover, introduce the nondecreasing, right-continuous process

$$(3.31) \quad \bar{\nu}_t^* := \sup\{\alpha \in [0, x] : \tau^*(x - \alpha) \leq t\}, \quad t \geq 0, \quad \bar{\nu}_{0-}^* = 0,$$

and then also the process

$$(3.32) \quad \nu_t^* := \int_0^t X_s^{1,0} d\bar{\nu}_s^*, \quad t > 0, \quad \nu_{0-}^* = 0.$$

Notice that $\bar{\nu}^*$ is the right-continuous inverse of $\tau^*(\cdot)$.

Theorem 3.13. *Let $\tilde{V}$ be as in (3.30) and V as in the definition of Problem (P2). Then one has $\tilde{V} = V$, and ν^* is the (unique) optimal control for Problem (P2).*

Proof. Step 1. Let $x > 0$, $\underline{y} \in \mathcal{Y}$, and $q \in \mathcal{I}$ be given and fixed. For $\nu \in \mathcal{A}(x, \underline{y}, q)$, we introduce the process $\bar{\nu}$ such that $\bar{\nu}_t := \int_0^t \frac{d\nu_s}{X_s^{1,0}}$, $t \geq 0$, and define its inverse (see, e.g., Chapter 0, Section 4 of [42]) by

$$(3.33) \quad \tau^{\bar{\nu}}(z) := \inf\{t \geq 0 \mid x - \bar{\nu}_t < z\}, \quad 0 < z \leq x.$$

Notice that the process $\tau^{\bar{\nu}}(z) := \{\tau^{\bar{\nu}}(z), z \leq x\}$ has decreasing, left-continuous sample paths, and hence it admits right-limits

$$(3.34) \quad \tau_+^{\bar{\nu}}(z) := \inf\{t \geq 0 \mid x - \bar{\nu}_t \leq z\}, \quad z \leq x.$$

Moreover, the set of points $z \in \mathbb{R}$ at which $\tau^{\bar{\nu}}(z)(\omega) \neq \tau_+^{\bar{\nu}}(z)(\omega)$ is a.s. countable for a.e. $\omega \in \Omega$.

The random time $\tau^{\bar{\nu}}(z)$ is actually an $(\mathcal{H}_t)$ -stopping time because it is the entry time of an open set of the right-continuous process $\bar{\nu}$, and $(\mathcal{H}_t)_{t \geq 0}$ is right-continuous. Moreover, since $\tau_+^{\bar{\nu}}(z)$ is the first entry time of the right-continuous process $\bar{\nu}$ into a closed set, it is an $(\mathcal{H}_t)$ -stopping time as well for any $z \leq x$.

Proceeding then as in Step 1 of the proof of Theorem 3.1 in [22], by employing the change of variable formula in Chapter 0, Proposition 4.9 of [42], one finds that

$$\tilde{V}(x, \underline{y}, q) = \int_0^x U(z, \underline{y}, q) dz \leq \mathcal{J}_{x, \underline{y}, q}(\nu).$$

Hence, since ν was arbitrary, we find

$$(3.35) \quad \tilde{V}(x, \underline{y}, q) \leq V(x, \underline{y}, q), \quad (x, \underline{y}, q) \in (0, \infty) \times \mathcal{Y} \times \mathcal{I}.$$

Step 2. To complete the proof we have to show the reverse inequality. Let $x \in (0, \infty)$, $\underline{y} \in \mathcal{Y}$, and $q \in \mathcal{I}$, initial values of $X^{x,\nu}$, $\underline{\pi}$ and η . We first notice that $\nu^* \in \mathcal{A}(x, \underline{y}, q)$. Indeed, ν^* is nondecreasing, right-continuous and such that $X_t^{x,\nu^*} = X_t^{1,0}(x - \bar{\nu}_t^*) \geq 0$ a.s. for all $t \geq 0$, since one has by definition $\bar{\nu}_t^* \leq x$ a.s. Moreover, for any $0 < z \leq x$, we can write (cf. (3.31) and (3.34))

$$\tau_+^{\bar{\nu}^*}(z) \leq t \iff \bar{\nu}_t^* \geq x - z \iff \tau^*(z) \leq t.$$

Then, recalling that $\tau_+^{\bar{\nu}^*}(z) = \tau^{\bar{\nu}^*}(z)$ $\mathbb{P}$ -a.s. and for a.e. $z \leq x$, we pick $\nu = \nu^*$ (equivalently, $\bar{\nu} = \bar{\nu}^*$), and following Step 2 in the proof of Theorem 3.1 of [22], we obtain $\tilde{V}(x, \underline{y}, q) = \mathcal{J}_{x,\underline{y},q}(\nu^*)$. That is, $\tilde{V} = V$ by (3.35) and admissibility of ν^* . Therefore ν^* is optimal. In fact, ν^* is the unique optimal control in the class of controls belonging to $\mathcal{A}(x, \underline{y}, q)$ and such that $\mathcal{J}_{x,\underline{y},q}(\nu) < \infty$ by strict convexity of $\mathcal{J}_{x,\underline{y},q}(\cdot)$. $\square$

Remark 3.14. For any given $(x, \underline{y}, q) \in (0, \infty) \times \mathcal{Y} \times \mathcal{I}$, define the Markovian optimal stopping problem

$$v(x, \underline{y}, q) := \inf_{\tau \geq 0} \mathbb{E}_{(x,\underline{y},q)} \left[\int_0^\tau e^{-\rho t} X_t^{x,0} \hat{h}_x(X_t^{x,0}, \underline{\pi}_t) dt + e^{-\rho \tau} X_\tau^{x,0} \right],$$

where $\mathbb{E}_{(x,\underline{y},q)}$ denotes the expectation under the probability measure $\mathbb{P}_{(x,\underline{y},q)}$ such that $\mathbb{P}(\cdot) := \mathbb{P}(\cdot | X_0^{x,0} = x, \underline{\pi}_0 = \underline{y}, \eta_0 = q)$. Then, it is readily verified that $v(x, \underline{y}, q) = xU(x, \underline{y}, q)$. Moreover, it holds that the stopping time

$$\tau^*(x) := \inf\{t \geq 0 : v(X_t^{x,0}, \underline{\pi}_t, \eta_t) \geq X_t^{x,0}\}, \quad \mathbb{P}_{(x,\underline{y},q)} - \text{a.s.}$$

is optimal for $v(x, \underline{y}, q)$.

4. THE SOLUTION IN A CASE STUDY WITH $Q = 2$ ECONOMIC REGIMES

In this section, we build on the general filtering analysis developed in the previous sections and on the result of Theorem 3.13, and we provide the form of the optimal debt reduction policy in a case study that is defined through the following standing assumption.

Assumption 4.1.

- (1) Z takes values in $S = \{1, 2\}$, and, with reference to (2.4), we take $g_2 := g(2) < g(1) =: g_1$;
- (2) for any $q \in \mathcal{I}$ and any $i \in \{1, 2\}$ one has $c(q, i) = 0$ and, for α as in (3.4), we take $\alpha(q, i) = \alpha(i)$;
- (3) $h(x, i) = h(x)$ for all $(x, i) \in (0, \infty) \times \{1, 2\}$, with $h : \mathbb{R} \rightarrow \mathbb{R}$ such that:
 - (i) $x \mapsto h(x)$ is strictly convex, twice-continuously differentiable, and nondecreasing on $\mathbb{R}_+$ with $h(0) = 0$ and $\lim_{x \uparrow \infty} h(x) = \infty$;
 - (ii) there exist $\gamma > 1$, $0 < K_o < K$ and $K_1, K_2 > 0$ such that

$$K_o|x^+|^\gamma - K \leq h(x) \leq K(1 + |x|^\gamma), \quad |h'(x)| \leq K_1(1 + |x|^{\gamma-1})$$

and

$$|h''(x)| \leq K_2(1 + |x|^{(\gamma-2)^+}).$$

Notice that under Assumption 4.1-(2) the macroeconomic indicator η has a suitable diffusive dynamics whose coefficients b_1, σ_1, σ_2 are such that the function α is independent of q . As discussed in Remark 3.8-(2), this is the case of a geometric or arithmetic diffusive dynamics for η . In this setting the Kushner-Stratonovich system (3.14) reduces to

$$(4.1) \quad d\pi_t(1) = [\lambda_2 - (\lambda_1 + \lambda_2)\pi_t(1)]dt + \pi_t(1)(1 - \pi_t(1)) \left[\frac{\beta_1 - \beta_2}{\sigma} dI_t + (\alpha_1 - \alpha_2)dI_t^1 \right],$$

and $\pi_t(2) = 1 - \pi_t(1)$. Here, $\lambda_1 := \lambda_{12} > 0$ and $\lambda_2 := \lambda_{21} > 0$.

Denoting by $\pi_t := \pi_t(1)$, $t \geq 0$, problem **(P2)** then reads as

$$(\mathbf{P3}) \quad \begin{cases} V(x, y) = \inf_{\nu \in \mathcal{A}(x, y)} \mathbb{E}_{(x, y)} \left[\int_0^\infty e^{-\rho t} h(X_t^\nu) dt + \int_0^\infty e^{-\rho t} d\nu_t \right] & \text{with} \\ dX_t^{x, y, \nu} = [\beta_2 + \pi_t^y (g_2 - g_1)] X_t^{x, y, \nu} dt + \sigma X_t^{x, y, \nu} dI_t - d\nu_t, & X_{0-}^{x, y, \nu} = x > 0, \\ d\pi_t^y = [\lambda_2 - (\lambda_1 + \lambda_2)\pi_t^y] dt + \pi_t^y (1 - \pi_t^y) \left[\frac{(g_2 - g_1)}{\sigma} dI_t + (\alpha_1 - \alpha_2) dI_t^1 \right], & \pi_0 = y \in (0, 1), \end{cases}$$

where $g_i = r - \beta_i$, denotes the rate of economic growth in the state i , $i = 1, 2$.

It is worth noticing that there is no need to involve the process η in the Markovian formulation of problem **(P3)**. This is due to the fact that the couple (X^ν, π) , solving the two stochastic differential equations above is a strong Markov process, and the cost functional and the set of admissible controls (denoted by $\mathcal{A}(x, y)$ above) do not depend explicitly on η . For this reason the value function of Problem **(P3)** does not depend on the initial value q of the process η . However, memory of the macroeconomic indicator process η appears in the filter π through the constant term $\alpha_1 - \alpha_2$ in its dynamics.

Finally, we recall that, thanks to Proposition 3.11, by solving Problem **(P3)** we are also solving the original problem **(P1)**. Indeed, we have that

$$V_{po}(x, y) = V(x, y), \quad \text{for any given and fixed } (x, y) \in (0, \infty) \times (0, 1),$$

and a control is optimal for the separated problem **(P3)** if and only if it is such for the original problem under partial observation.

In the following analysis, we need (for technical reasons due to the infinite time-horizon of our problem) to take a discount factor sufficiently large. Namely, defining

$$\begin{aligned} \rho_o := & (\beta_2 + \frac{1}{2}\sigma^2) \vee \left[\gamma\beta_2 + \frac{1}{2}\sigma^2\gamma(\gamma - 1) \right] \vee (2\beta_2 + \sigma^2) \vee [24\theta^2 - (\lambda_1 + \lambda_2)] \\ & \vee (4\beta_2 + 6\sigma^2) \vee [4\beta_2(2 \vee \gamma) + 2\sigma^2(2 \vee \gamma)(4(2 \vee \gamma) - 1)], \end{aligned}$$

with $\theta^2 := \frac{1}{2} \left[\frac{(g_1 - g_2)^2}{\sigma^2} + (\alpha_1 - \alpha_2)^2 \right]$, we assume the following.

Assumption 4.2. *One has $\rho > \rho_o^+$.*

Due to the growth condition on h , Assumption 4.2 in particular ensures that $\rho > \gamma\beta_2 + \frac{1}{2}\sigma^2\gamma(\gamma - 1)$ so that the (trivial) admissible control $\nu \equiv 0$ has a finite total expected cost.

4.1. The Related Optimal Stopping Problem. Motivated by the results of the previous sections (see in particular Theorem 3.13), we now aim at solving Problem **(P3)** through the study of an auxiliary optimal stopping problem. Informally, the solution to such an optimal stopping problem gives the optimal time at which the government should reduce the debt ratio by one additional unit. The optimal stopping problem involves a two-dimensional diffusive process, and in the following we provide an almost exclusively probabilistic analysis.

4.1.1. Formulation and Preliminary Results. Recall that $(I_t, I_t^1)_{t \geq 0}$ is a two-dimensional, standard $\mathbb{H}$ -Brownian motion, and introduce the two-dimensional diffusion process $(\hat{X}, \pi) := (\hat{X}_t, \pi_t)_{t \geq 0}$ solving the stochastic differential equations (SDEs)

$$(4.2) \quad \begin{cases} d\hat{X}_t = \hat{X}_t [\beta_2 + (g_2 - g_1)\pi_t] dt + \sigma \hat{X}_t dI_t, \\ d\pi_t = [\lambda_2 - (\lambda_1 + \lambda_2)\pi_t] dt + \pi_t(1 - \pi_t) \left[\frac{(g_2 - g_1)}{\sigma} dI_t + (\alpha_1 - \alpha_2) dI_t^1 \right], \end{cases}$$

with initial conditions $\hat{X}_0 = x$, $\pi_0 = y$ for any $(x, y) := (0, \infty) \times (0, 1)$. In the following, we set $\mathcal{O} := (0, \infty) \times (0, 1)$. Recall that $\beta_2 = r - g_2$.

Since the process π is bounded, classical results on SDEs ensure that system (4.2) admits a unique strong solution, that, when needed, we shall denote by $(\hat{X}^{x, y}, \pi^y)$ in order to stress its dependence on the initial datum $(x, y) \in \mathcal{O}$. In particular, one easily obtains

$$(4.3) \quad \hat{X}_t^{x, y} = x e^{(\beta_2 - \frac{1}{2}\sigma^2)t + \sigma I_t + (g_2 - g_1) \int_0^t \pi_s^y ds}, \quad t \geq 0,$$

Moreover, it can be shown that the Feller's test of explosion (see, e.g., Chapter 5.5 in [32]) gives that $1 = \mathbb{P}(\pi_t^y \in (0, 1), \forall t \geq 0)$ for all $y \in (0, 1)$. In fact, the boundary points 0 and 1 are entrance-not-exit (cf. [3], p. 15), hence unattainable for the process π .

With regard to Remark 3.14, here we study the fully two-dimensional Markovian optimal stopping problem with value function

$$(4.4) \quad v(x, y) := \inf_{\tau \geq 0} \mathbb{E}_{(x, y)} \left[\int_0^\tau e^{-\rho t} \widehat{X}_t h'(\widehat{X}_t) dt + e^{-\rho \tau} \widehat{X}_\tau \right] =: \inf_{\tau \geq 0} \widehat{\mathcal{J}}_{(x, y)}(\tau), \quad (x, y) \in \mathcal{O}.$$

In (4.4) the optimization is taken over all the $\mathbb{H}$ -stopping times, and the symbol $\mathbb{E}_{(x, y)}$ denotes the expectation under the probability measure $\mathbb{P}_{(x, y)}$ on $(\Omega, \mathcal{F})$, defined as $\mathbb{P}_{(x, y)}(\cdot) := \mathbb{P}(\cdot | \widehat{X}_0 = x, \pi_0 = y)$, for any $(x, y) \in \mathcal{O}$.

Due to the fact that π is positive, $g_2 - g_1 < 0$, and $\rho > \beta_2$ by Assumption 4.2, one has from (4.3) that

$$(4.5) \quad \liminf_{t \uparrow \infty} e^{-\rho t} \widehat{X}_t = 0 \quad \mathbb{P}_{(x, y)} - a.s.,$$

which implies the convention (cf. (3.27)) $e^{-\rho \tau} \widehat{X}_\tau = 0$ on $\{\tau = \infty\}$.

Clearly, one has $v \geq 0$ since $\widehat{X}$ is positive and h is increasing on $\mathbb{R}_+$. Also, $v \leq x$ on $\mathcal{O}$, and we can therefore define the *continuation region* and the *stopping region* as

$$(4.6) \quad \mathcal{C} := \{(x, y) \in \mathcal{O} : v(x, y) < x\}, \quad \mathcal{S} := \{(x, y) \in \mathcal{O} : v(x, y) = x\}.$$

Notice that integrating by parts the term $e^{-\rho \tau} \widehat{X}_\tau$, taking expectations, and exploiting that for any $\mathbb{H}$ -stopping time τ one has $\mathbb{E}[\int_0^\tau e^{-\rho s} \widehat{X}_s dI_s] = 0$ (because $\rho > \beta_2 + \frac{1}{2}\sigma^2$ by Assumption 4.2), we can equivalently rewrite (4.4) as

$$(4.7) \quad v(x, y) := x + \inf_{\tau \geq 0} \mathbb{E}_{(x, y)} \left[\int_0^\tau e^{-\rho t} \widehat{X}_t \left(h'(\widehat{X}_t) - (\rho - \beta_2 - (g_2 - g_1)\pi_t) \right) dt \right],$$

for any $(x, y) \in \mathcal{O}$. From (4.7) it is readily seen that

$$(4.8) \quad \{(x, y) \in \mathcal{O} : h'(x) - (\rho - \beta_2 - (g_2 - g_1)y) < 0\} \subseteq \mathcal{C},$$

which implies

$$(4.9) \quad \mathcal{S} \subseteq \{(x, y) \in \mathcal{O} : h'(x) - (\rho - \beta_2 - (g_2 - g_1)y) \geq 0\},$$

Moreover, since ρ satisfies Assumption 4.2, and $0 \leq \pi_t \leq 1$ for any $(x, y) \in \mathcal{O}$, one has that

$$(4.10) \quad \mathbb{E}_{(x, y)} \left[\int_0^\infty e^{-\rho t} \widehat{X}_t \left(h'(\widehat{X}_t) + \rho + |\beta_2| + |g_2 - g_1| dt \right) \right] < \infty,$$

and the family of random variables

$$\left\{ \int_0^\tau e^{-\rho t} \widehat{X}_t \left(h'(\widehat{X}_t) - (\rho - \beta_2 - (g_2 - g_1)\pi_t) \right) dt, \tau \text{ } \mathbb{H}\text{-stopping time} \right\}$$

is therefore $\mathbb{H}$ -uniformly integrable under $\mathbb{P}_{(x, y)}$.

Preliminary properties of v are given in the next proposition.

Proposition 4.3. *The following hold:*

- (i) $x \mapsto v(x, y)$ is increasing for any $y \in (0, 1)$;
- (ii) $y \mapsto v(x, y)$ is decreasing for any $x \in (0, \infty)$;
- (iii) $(x, y) \mapsto v(x, y)$ is continuous in $\mathcal{O}$.

Proof. We prove each claim separately.

(i). Recall (4.4). By the strict convexity and the monotonicity of h and (4.3), it follows that $x \mapsto \widehat{\mathcal{J}}_{(x, y)}(\tau)$ is increasing for any $\mathbb{H}$ -stopping time τ , and for any $y \in (0, 1)$. Hence the claim is proved.

(ii). This is due to the fact that $y \mapsto \widehat{\mathcal{J}}_{(x, y)}(\tau)$ is decreasing for any stopping time τ and any $x \in (0, \infty)$. Indeed, the mapping $y \mapsto \widehat{X}_t^{x, y}$ is a.s. decreasing for any $t \geq 0$ (because $y \mapsto \pi_t^y$ is a.s. increasing by the comparison theorem of Yamada and Watanabe - see, e.g., Proposition 2.18 in Chapter 5.2 of [32] - and $g_2 - g_1 < 0$), and $x \mapsto xh'(x)$ is increasing.

(iii). Since $(x, y) \mapsto (\hat{X}_t^{x,y}, \pi_t^y)$ is a.s. continuous for any $t \geq 0$, it is not hard to verify that $(x, y) \mapsto \hat{\mathcal{J}}_{(x,y)}(\tau)$ is continuous for any given $\tau \geq 0$. Hence, v is upper semicontinuous. We now show that it is also lower semicontinuous.

Let $(x, y) \in \mathcal{O}$ and let $(x_n, y_n)_n \subseteq \mathcal{O}$ be any sequence converging to (x, y) . Without loss of generality, we may take $(x_n, y_n) \in (x - \delta, x + \delta) \times (y - \delta, y + \delta)$, for a suitable $\delta > 0$. Letting $\tau_\varepsilon^n := \tau_\varepsilon^n(x_n, y_n)$ be an ε -optimal for $v(x_n, y_n)$, but suboptimal for $v(x, y)$, we can then write

$$(4.11) \quad \begin{aligned} v(x, y) - v(x_n, y_n) &\leq \mathbb{E} \left[\int_0^{\tau_\varepsilon^n} e^{-\rho t} \left(\hat{X}_t^{x,y} h'(\hat{X}_t^{x,y}) - \hat{X}_t^{x_n, y_n} h'(\hat{X}_t^{x_n, y_n}) \right) dt \right] \\ &\quad + \mathbb{E} \left[e^{-\rho \tau_\varepsilon^n} \left(\hat{X}_{\tau_\varepsilon^n}^{x,y} - \hat{X}_{\tau_\varepsilon^n}^{x_n, y_n} \right) \right] + \varepsilon. \end{aligned}$$

Notice now that a.s.

$$\int_0^{\tau_\varepsilon^n} e^{-\rho t} \left| \hat{X}_t^{x,y} h'(\hat{X}_t^{x,y}) - \hat{X}_t^{x_n, y_n} h'(\hat{X}_t^{x_n, y_n}) \right| dt \leq \int_0^\infty e^{-\rho t} \left(\hat{X}_t^{x,y} h'(\hat{X}_t^{x,y}) + \hat{X}_t^{x+\delta, y-\delta} h'(\hat{X}_t^{x+\delta, y-\delta}) \right) dt,$$

where we have used that $x \mapsto \hat{X}^{x,y}$ is increasing, $y \mapsto \hat{X}^{x,y}$ is decreasing, and $x \mapsto xh'(x)$ is positive and increasing. The random variable on the right-hand side of the latter equation is independent of n and integrable due to (4.10).

Also, by an integration by parts, and performing standard estimates, we can write that a.s.

$$e^{-\rho \tau_\varepsilon^n} \left(\hat{X}_{\tau_\varepsilon^n}^{x,y} - \hat{X}_{\tau_\varepsilon^n}^{x_n, y_n} \right) \leq |x - x_n| + \int_0^\infty e^{-\rho s} (\rho + |\beta_2| + |g_2 - g_1|) (\hat{X}_s^{x,y} + \hat{X}_s^{x+\delta, y-\delta}) ds,$$

and the last integral above is independent of n and it has finite expectation due to (4.10).

Then, taking limits as $n \uparrow \infty$, invoking the dominated convergence theorem thanks to the previous estimates, and using that $(x, y) \mapsto (\hat{X}_t^{x,y}, \pi_t^y)$ is a.s. continuous for any $t \geq 0$ we find (after rearranging terms) that

$$\liminf_{n \uparrow \infty} v(x_n, y_n) \geq v(x, y) - \varepsilon.$$

We thus conclude that v is lower semicontinuous at (x, y) by arbitrariness of ε . Since $(x, y) \in \mathcal{O}$ was arbitrary as well, then v is lower semicontinuous on $\mathcal{O}$. $\square$

Due to Proposition 4.3-(iii) one has that the stopping region is closed, whereas the continuation region is open. Moreover, thanks to (4.10) and the $\mathbb{P}_{(x,y)}$ -a.s. continuity of $t \mapsto \int_0^t e^{-\rho s} \hat{X}_s (h'(\hat{X}_s) - (\rho - \beta_2 - (g_2 - g_1)\pi_s) ds)$, we can apply Theorem D.12 in Appendix D of [33] to obtain that the first entry time of $(\hat{X}, \pi)$ into $\mathcal{S}$ is optimal for (4.4); that is,

$$(4.12) \quad \tau^*(x, y) := \inf \{ t \geq 0 : (\hat{X}_t, \pi_t) \in \mathcal{S} \}, \quad \mathbb{P}_{(x,y)} - a.s., \quad (x, y) \in \mathcal{O},$$

attains the infimum in (4.4) (here we adopt the usual convention $\inf \emptyset = \infty$).

Also, by employing standard means based on the strong Markov property of $(\hat{X}, \pi)$ (see, e.g., [39], Ch. I, Sec. 2, Thm. 2.4), one can show that, $\mathbb{P}_{(x,y)}$ -a.s., the process $S := (S_t)_{t \geq 0}$, with

$$S_t := \left(e^{-\rho t} v(\hat{X}_t, \pi_t) + \int_0^t e^{-\rho s} \hat{X}_s h'(\hat{X}_s) ds \right)_{t \geq 0}, \quad \text{is an } \mathbb{H}\text{-submartingale,}$$

and that the stopped process $(S_{t \wedge \tau^*})_{t \geq 0}$ is an $\mathbb{H}$ -martingale. The latter two conditions are usually referred to as the *subharmonic characterization* of the value function v .

We now rule out the possibility of an empty stopping region.

Lemma 4.4. *The stopping region of (4.6) is not empty.*

Proof. We argue by contradiction and we suppose that $\mathcal{S} = \emptyset$. Hence, for any $(x, y) \in \mathcal{O}$ we can write

$$(4.13) \quad x > v(x, y) = \mathbb{E}_{(x,y)} \left[\int_0^\infty e^{-\rho t} \hat{X}_t h'(\hat{X}_t) dt \right] \geq K_o x^\gamma \mathbb{E}_{(1,y)} \left[\int_0^\infty e^{-\rho t} (\hat{X}_t)^\gamma dt \right] - \frac{K}{\rho},$$

where the inequality $xh'(x) \geq h(x)$, due to convexity of h , and the growth condition assumed on h (cf. Assumption 4.1) have been used. Now, by taking x sufficiently large, we reach a contradiction since $\gamma > 1$ by assumption. Hence $\mathcal{S} \neq \emptyset$. $\square$

Proposition 4.5. *For any $y \in (0, 1)$ let*

$$(4.14) \quad d(y) := \inf\{x > 0 : v(x, y) \geq x\},$$

where the convention $\inf \emptyset = +\infty$ has been used. Then

(i)

$$(4.15) \quad \mathcal{C} = \{(x, y) \in \mathcal{O} : x < d(y)\} \quad \text{and} \quad \mathcal{S} = \{(x, y) \in \mathcal{O} : x \geq d(y)\};$$

(ii) $y \mapsto d(y)$ is increasing and left-continuous;

(iii) there exist $0 < x_\star < x^\star < \infty$ such that for any $y \in [0, 1]$

$$(h')^{-1}(\rho - \beta_2) \vee x_\star \leq d(y) \leq x^\star.$$

Proof. (i). To show that (4.15) holds true it suffices to show that if $(x_1, y) \in \mathcal{S}$, then $(x_2, y) \in \mathcal{S}$ for any $x_2 \geq x_1$. Let $\tau^\varepsilon := \tau^\varepsilon(x_2, y)$ be an ε -optimal stopping time for $v(x_2, y)$. Then, exploiting the fact that $\hat{X}_t^{x_2, y} = \frac{x_2}{x_1} \hat{X}_t^{x_1, y} \geq \hat{X}_t^{x_1, y}$ a.s. and the monotonicity of h' , we can write from (4.7)

$$(4.16) \quad \begin{aligned} 0 &\geq v(x_2, y) - x_2 \geq \mathbb{E} \left[\int_0^{\tau^\varepsilon} e^{-\rho t} \hat{X}_t^{x_2, y} \left(h'(\hat{X}_t^{x_2, y}) - (\rho - \beta_2 - (g_2 - g_1)\pi_t^y) \right) dt \right] - \varepsilon \\ &\geq \frac{x_2}{x_1} \mathbb{E} \left[\int_0^{\tau^\varepsilon} e^{-\rho t} \hat{X}_t^{x_1, y} \left(h'(\hat{X}_t^{x_1, y}) - (\rho - \beta_2 - (g_2 - g_1)\pi_t^y) \right) dt \right] - \varepsilon \\ &\geq \frac{x_2}{x_1} (v(x_1, y) - x_1) - \varepsilon = -\varepsilon. \end{aligned}$$

Therefore, by arbitrariness of ε , we conclude that $(x_2, y) \in \mathcal{S}$ as well, and therefore that d as in (4.14) splits $\mathcal{C}$ and $\mathcal{S}$ as in (4.15).

(ii). Let $(x, y_1) \in \mathcal{C}$. Since $y \mapsto v(x, y)$ is decreasing by Proposition 4.3-(ii), it thus follows that $(x, y_2) \in \mathcal{C}$ for any $y_2 \geq y_1$. This in turn implies that $y \mapsto d(y)$ is increasing. The monotonicity of $y \mapsto d(y)$, together with the fact that $\mathcal{S}$ is closed, then give the claimed left-continuity by standard arguments.

(iii). Let $\Theta_t^x := x \exp\{(\beta_2 - \frac{1}{2}\sigma^2 + (g_2 - g_1))t + \sigma I_t\}$, and introduce the one-dimensional optimal stopping problem

$$(4.17) \quad v^\star(x) := \inf_{\tau \geq 0} \mathbb{E} \left[\int_0^\tau e^{-\rho t} \Theta_t^x h'(\Theta_t^x) dt + e^{-\rho \tau} \Theta_\tau^x \right], \quad x > 0.$$

Because $g_2 - g_1 < 0$, h' is increasing, and $\pi_t^y \leq 1$ a.s. for all $t \geq 0$ and $y \in (0, 1)$, it is not hard to see that $v(x, y) \geq v^\star(x)$ for any $(x, y) \in \mathcal{O}$.

By arguments similar to those employed to prove (i) above one can show that there exists $x^\star$ such that $\{x \in (0, \infty) : v^\star(x) \geq x\} = \{x \in (0, \infty) : x \geq x^\star\}$. In fact, by arguing as in the proof of Lemma 4.4, one has that the latter set is not empty. Then the following inclusions hold

$$\{x \in (0, \infty) : x \geq x^\star\} \subseteq \{(x, y) \in \mathcal{O} : v(x, y) \geq x\} = \{(x, y) \in \mathcal{O} : x \geq d(y)\},$$

which in turn show that $d(y) \leq x^\star$ for all $y \in (0, 1)$. Hence, also $d(y) \leq x^\star$ for all $y \in [0, 1]$, by setting $d(0+) := \lim_{y \downarrow 0} d(y)$ by monotonicity, and $d(1) := \lim_{y \uparrow 1} d(y)$ by left-continuity.

As for the lower bound of d , notice that (4.9) implies

$$(4.18) \quad d(y) \geq (h')^{-1}(\rho - \beta_2 - (g_2 - g_1)y) =: \zeta(y), \quad y \in (0, 1),$$

where $(h')^{-1}(\cdot)$ is the inverse of the strictly increasing function $h' : [0, \infty) \mapsto (0, \infty)$ (notice that $\rho - \beta_2 - (g_2 - g_1)y \geq 0$ since $\rho > \beta_2$, $g_2 - g_1 < 0$, and $y > 0$). Since $(h')^{-1}$ is strictly increasing, and $-(g_2 - g_1)y \geq 0$, we can conclude from (4.18) that $d(y) \geq (h')^{-1}(\rho - \beta_2)$ for every $y \in [0, 1]$.

Moreover, setting $\Psi_t^x := x \exp\{(\beta_2 - \frac{1}{2}\sigma^2)t + \sigma I_t\}$ and introducing the one-dimensional optimal stopping problem

$$(4.19) \quad v_\star(x) := \inf_{\tau \geq 0} \mathbb{E} \left[\int_0^\tau e^{-\rho t} \Psi_t^x h'(\Psi_t^x) dt + e^{-\rho \tau} \Psi_\tau^x \right], \quad x > 0,$$

one has that $v(x, y) \leq v_\star(x)$ for any $(x, y) \in \mathcal{O}$. Following arguments as those employed above, the last inequality implies that $d(y) \geq x_\star$ for all $y \in [0, 1]$, where $x_\star := \inf\{x > 0 : v_\star(x) \geq x\} \in (0, \infty)$. $\square$

4.1.2. *Smooth-Fit Property and Continuity of the Free Boundary.* We now aim at proving further regularity of v and of the free boundary d .

The second-order linear elliptic differential operator

$$(4.20) \quad \begin{aligned} \mathbb{L} := & (\beta_2 + (g_2 - g_1)y)x \frac{\partial}{\partial x} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2}{\partial x^2} + (\lambda_2 - (\lambda_1 + \lambda_2)y) \frac{\partial}{\partial y} \\ & + \frac{1}{2} \left((\alpha_1 - \alpha_2)^2 + \frac{(g_2 - g_1)^2}{\sigma^2} \right) y^2 (1 - y)^2 \frac{\partial^2}{\partial y^2}, \end{aligned}$$

acting on any function $f \in C^2(\mathcal{O})$, is the infinitesimal generator of the process $(\hat{X}, \pi)$. The non-degeneracy of the process $(\hat{X}, \pi)$, the smoothness of the coefficients in (4.20), together with the subharmonic characterization of v , allow to prove by standard arguments (see, e.g., [39], Ch. 3, Sec. 7.1) and classical regularity results for elliptic partial differential equations (see, e.g., [26]) the following result.

Lemma 4.6. *The value function v of (4.4) belongs to C^2 separately strictly inside $\mathcal{C}$ and $\mathcal{S}$ (i.e. away from the boundary $\partial\mathcal{C}$ of $\mathcal{C}$). Moreover, inside $\mathcal{C}$ it uniquely solves*

$$(4.21) \quad (\mathbb{L} - \rho)v(x, y) = -xh'(x),$$

with $\mathbb{L}$ as in (4.20).

We continue our analysis by proving that the value function of (4.4) belongs to $C^1((0, \infty) \times (0, 1))$. This will be obtained through probabilistic methods that rely on the regularity (in the sense of diffusions) of the stopping set $\mathcal{S}$ for the process $(\hat{X}, \pi)$ (see also [30] for similar ideas and arguments in a different context, and [16] for general results). Recall that the boundary points are regular for $\mathcal{S}$ relative to $(\hat{X}, \pi)$ if (cf. Definition 2.9 p. 249 in [32])

$$(4.22) \quad \hat{\tau}(x_o, y_o) := \inf\{t > 0 : (\hat{X}_t^{x_o, y_o}, \pi_t^{y_o}) \in \mathcal{S}\} = 0 \quad \text{a.s.} \quad \forall (x_o, y_o) \in \partial\mathcal{C}.$$

The time $\hat{\tau}(x_o, y_o)$ is the first hitting time of $(\hat{X}^{x_o, y_o}, \pi^{y_o})$ to $\mathcal{S}$.

Notice that for every bounded Borel function $f : \mathbb{R}^2 \mapsto \mathbb{R}$ one has $\mathbb{E}_{(x, y)}[f(\hat{X}_t, \pi_t)] = \mathbb{E}_{(u, y)}[f(e^{U_t}, \pi_t)]$, where $u := \ln(x)$ and $U_t := \ln(\hat{X}_t)$ is such that $dU_t = (\beta_2 + (g_2 - g_1)\pi_t - \frac{1}{2}\sigma^2)dt + \sigma dI_t$. Due the nondegeneracy of the process (U, π) , and the smoothness and boundedness of its coefficients, we have that (U, π) has a continuous transition density $\hat{p}(\cdot, \cdot, \cdot; u, y)$, $(u, y) \in \mathbb{R} \times (0, 1)$, such that for any $t \geq 0$ and $(u', y') \in \mathbb{R} \times (0, 1)$ (see, e.g., [1])

$$(4.23) \quad \begin{aligned} & \frac{M}{t} \exp \left\{ -\lambda \frac{((u - u')^2 + (y - y')^2)}{t} \right\} \geq \hat{p}(t, u', y'; u, y) \\ & \geq \frac{m}{t} \exp \left\{ -\Lambda \frac{((u - u')^2 + (y - y')^2)}{t} \right\}, \end{aligned}$$

for some constants $M > m > 0$ and $\Lambda > \lambda > 0$. It thus follows that $(u, y) \mapsto \mathbb{E}_{(u, y)}[f(e^{U_t}, \pi_t)]$ is continuous, so that (U, π) is a strong Feller process. Hence, $(\hat{X}, \pi)$ is strong Feller as well, and we can therefore conclude that (4.22) holds true if and only if (see [17], pp. 32-40)

$$(4.24) \quad \tau^*(x_n, y_n) \rightarrow 0 \quad \text{a.s.} \quad \text{whenever} \quad \mathcal{C} \supseteq (x_n, y_n)_n \rightarrow (x_o, y_o) \in \partial\mathcal{C},$$

where τ^* is as in (4.12).

The next proposition shows the validity of (4.22).

Proposition 4.7. *The boundary points in $\partial\mathcal{C}$ are regular for $\mathcal{S}$ relative to $(\hat{X}, \pi)$; that is, (4.22) holds.*

Proof. Let $(x_o, y_o) \in \partial\mathcal{C}$, and set $u_o := \ln(x_o)$. With U as defined above, we set $\hat{\sigma}(u_o, y_o) := \hat{\tau}(e^{u_o}, y_o)$, $(u_o, y_o) \in \mathbb{R} \times (0, 1)$, and we equivalently rewrite (4.22) in terms of the process (U, π) as

$$\hat{\sigma}(u_o, y_o) := \inf\{t > 0 : U_t^{u_o, y_o} \geq \ln(d(\pi^{y_o}))\} = 0 \quad \text{a.s.} \quad \forall (u_o, y_o) \text{ such that } u_o = \ln(d(y_o)).$$

Given that $y \mapsto \ln(d(y))$ is increasing (since $y \mapsto d(y)$ is such), then the region $\hat{\mathcal{S}} := \{(u, y) \in \mathbb{R} \times (0, 1) : u \geq \ln(d(y))\}$ enjoys the so-called *cone property* (see [32], p. 250). In particular,

we can always construct a cone C_o with vertex in (u_o, y_o) and aperture $0 \leq \phi \leq \pi/2$ such that $C_o \cap (\mathbb{R} \times (0, 1)) \subseteq \hat{\mathcal{S}}$, and for any $t_o \geq 0$ we can write that

$$(4.25) \quad \mathbb{P}(\hat{\sigma}(u_o, y_o) \leq t_o) \geq \mathbb{P}((U_{t_o}^{u_o, y_o}, \pi_{t_o}^{y_o}) \in C_o).$$

Then using (4.23) one has

$$(4.26) \quad \begin{aligned} \mathbb{P}((U_{t_o}^{u_o, y_o}, \pi_{t_o}^{y_o}) \in C_o) &= \int_{C_o} \hat{p}(t_o, u_o, y_o; u, y) du dy \geq \int_{C_o} \frac{m}{t_o} e^{-\Lambda \frac{((u-u_o)^2 + (y-y_o)^2)}{t_o}} du dy \\ &= m \int_{C_o} e^{-\Lambda((u')^2 + (y')^2)} du' dy' =: \ell > 0, \end{aligned}$$

where we have used that the change of variable $u' := (u - u_o)/\sqrt{t_o}$ and $y' := (y - y_o)/\sqrt{t_o}$ maps the cone C_o into itself. The number ℓ above depends on u_o, y_o , but it is independent of t_o . From (4.25) and (4.26) we thus have that $\mathbb{P}(\hat{\sigma}(u_o, y_o) \leq t_o) \geq \ell$, and letting $t_o \downarrow 0$ we obtain $\mathbb{P}(\hat{\sigma}(u_o, y_o) = 0) \geq \ell > 0$. However, $\{\hat{\sigma}(u_o, y_o) = 0\} \in \mathcal{H}_0$, and by the Blumenthal's 0-1 Law we obtain $\mathbb{P}(\hat{\sigma}(u_o, y_o) = 0) = 1$, which completes the proof. $\square$

Theorem 4.8. *One has that $v \in C^1(\mathcal{C})$.*

Proof. The value function belongs to C^2 strictly inside the continuation region due to Lemma 4.6, and it is C^∞ strictly inside the stopping region where $v = x$. It thus only remains to prove that v is continuously differentiable across $\partial\mathcal{C}$. In the following, we will prove that: (i) the function $\bar{w} := \frac{1}{x}(v - x)$ has continuous derivative with respect to x across $\partial\mathcal{C}$ (and this clearly implies the continuity of v_x across $\partial\mathcal{C}$); (ii) that the function v_y is continuous across $\partial\mathcal{C}$.

(i) *Continuity of v_x across $\partial\mathcal{C}$.* For the subsequent arguments it is useful to notice that the function $\bar{w} = \frac{1}{x}(v - x)$ admits the representation (recall (4.7))

$$(4.27) \quad \bar{w}(x, y) = \inf_{\tau \geq 0} \mathbb{E} \left[\int_0^\tau e^{-\rho t} \hat{X}_t^{1, y} \left(h'(\hat{X}_t^{x, y}) - (\rho - \beta_2 - (g_2 - g_1)\pi_s^y) \right) ds \right],$$

and to bear in mind that the optimal stopping time τ^* for v as in (4.12) is also optimal for $\bar{w}$ since $v \geq x$ if and only if $\bar{w} \geq 0$. We now prove that $\bar{w}_x$ is continuous across $\partial\mathcal{C}$, thus implying continuity of v_x across $\partial\mathcal{C}$.

Take $(x, y) \in \mathcal{C}$, and let $\varepsilon > 0$ be such that $x - \varepsilon > 0$. Since $x \mapsto \bar{w}(x, y)$ is increasing (due to the monotonicity of h') it is clear that $(x - \varepsilon, y) \in \mathcal{C}$ as well. Denote by $\tau_\varepsilon^*(x, y) := \tau^*(x - \varepsilon, y)$ the optimal stopping time for $\bar{w}(x - \varepsilon, y)$, and notice that $\tau_\varepsilon^*(x, y)$ is suboptimal for $\bar{w}(x, y)$ and $\tau_\varepsilon^*(x, y) \rightarrow \tau^*(x, y)$ a.s. To simplify exposition in the following we write $\tau_\varepsilon^* := \tau_\varepsilon^*(x, y)$ and $\tau^* := \tau^*(x, y)$. We can then write from (4.27)

$$\begin{aligned} 0 \leq \frac{\bar{w}(x, y) - \bar{w}(x - \varepsilon, y)}{\varepsilon} &\leq \frac{1}{\varepsilon} \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} \hat{X}_t^{1, y} \left(h'(\hat{X}_t^{x, y}) - h'(\hat{X}_t^{x - \varepsilon, y}) \right) dt \right] \\ &= \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} (\hat{X}_t^{1, y})^2 h''(\hat{X}_t^{\xi_\varepsilon, y}) dt \right], \end{aligned}$$

for some $\xi_\varepsilon \in (x - \varepsilon, x)$, and where in the last step we have used the mean value theorem, and the fact that $\hat{X}_t^{x, y} - \hat{X}_t^{x - \varepsilon, y} = \varepsilon \hat{X}_t^{1, y}$. Letting $\varepsilon \downarrow 0$, invoking the dominated convergence theorem (thanks to the fact that $\rho > (\gamma\beta_2 + \frac{1}{2}\sigma^2\gamma(\gamma - 1)) \vee (2\beta_2 + \sigma^2)$ by Assumption 4.2), and using that $\bar{w} \in C^1(\mathcal{C})$ (since $v \in C^1(\mathcal{C})$), we then find from the latter that

$$(4.28) \quad 0 \leq \bar{w}_x(x, y) \leq \mathbb{E} \left[\int_0^{\tau^*} e^{-\rho t} (\hat{X}_t^{1, y})^2 h''(\hat{X}_t^{x, y}) dt \right].$$

Let now (x_o, y_o) be any arbitrary point belonging to $\partial\mathcal{C}$. Taking limits in (4.28) as $(x, y) \rightarrow (x_o, y_o)$, by the dominated convergence theorem and thanks to Proposition 4.7 we obtain that

$$0 \leq \liminf_{(x, y) \rightarrow (x_o, y_o) \in \partial\mathcal{C}} \bar{w}_x(x, y) \leq \limsup_{(x, y) \rightarrow (x_o, y_o) \in \partial\mathcal{C}} \bar{w}_x(x, y) \leq 0,$$

thus proving that $\bar{w}_x$ is continuous across $\partial\mathcal{C}$. This immediately implies the continuity of v_x across $\partial\mathcal{C}$, upon recalling that $v = x(\bar{w} + 1)$.

(ii) *Continuity of v_y across $\partial\mathcal{C}$.* Take again $(x, y) \in \mathcal{C}$, and let $\varepsilon > 0$ be such that $y + \varepsilon < 1$. Since $y \mapsto v(x, y)$ is decreasing (cf. Proposition 4.3-(ii)), it is clear that $(x, y + \varepsilon) \in \mathcal{C}$ as well. Denote by $\tau_\varepsilon^*(x, y) := \tau^*(x, y + \varepsilon)$ the optimal stopping time for $v(x, y + \varepsilon)$ and notice that $\tau_\varepsilon^*(x, y)$ is suboptimal for $v(x, y)$ and $\tau^*(x, y + \varepsilon) \rightarrow \tau^*(x, y)$ a.s. as $\varepsilon \downarrow 0$. In order to simplify the notation, in the following we write τ_ε^* instead of $\tau_\varepsilon^*(x, y)$.

From Proposition 4.3-(ii) and (4.7) we can then write

$$\begin{aligned} 0 &\geq \frac{v(x, y + \varepsilon) - v(x, y)}{\varepsilon} \geq \frac{1}{\varepsilon} \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} \hat{X}_t^{x, y + \varepsilon} \left[h'(\hat{X}_t^{x, y + \varepsilon}) - (\rho - \beta_2 - \pi_t^{y + \varepsilon}(g_2 - g_1)) \right] dt \right] \\ &\quad - \frac{1}{\varepsilon} \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} \hat{X}_t^{x, y} \left[h'(\hat{X}_t^{x, y}) - (\rho - \beta_2 - \pi_t^y(g_2 - g_1)) \right] dt \right] \\ &= \frac{1}{\varepsilon} \left(\mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} \left[\hat{X}_t^{x, y + \varepsilon} h'(\hat{X}_t^{x, y + \varepsilon}) - \hat{X}_t^{x, y} h'(\hat{X}_t^{x, y}) \right] dt - \int_0^{\tau_\varepsilon^*} e^{-\rho t} (\rho - \beta_2) (\hat{X}_t^{x, y + \varepsilon} - \hat{X}_t^{x, y}) dt \right] \right) \\ &\quad + \frac{1}{\varepsilon} \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} (g_2 - g_1) (\hat{X}_t^{x, y + \varepsilon} \pi_t^{y + \varepsilon} - \hat{X}_t^{x, y} \pi_t^y) dt \right]. \end{aligned}$$

Now, add and subtract both $\mathbb{E}[\int_0^{\tau_\varepsilon^*} e^{-\rho t} \hat{X}_t^{x, y + \varepsilon} h'(\hat{X}_t^{x, y}) dt]$ and $(g_2 - g_1)\mathbb{E}[\int_0^{\tau_\varepsilon^*} e^{-\rho t} \hat{X}_t^{x, y + \varepsilon} \pi_t^y dt]$ in the right-hand side of the latter, and recall that $(g_2 - g_1) < 0$, that $\hat{X}_t^{x, y} \geq 0$ a.s. for every $t \geq 0$, as well as that $(\pi_t^{y + \varepsilon} - \pi_t^y) \geq 0$ a.s. for every $t \geq 0$. Then, after rearranging terms and employing the integral mean value theorem (for some $L_t^\varepsilon \in (\hat{X}_t^{x, y + \varepsilon}, \hat{X}_t^{x, y})$ a.s.), we obtain from the equation above that

$$\begin{aligned} 0 &\geq \frac{v(x, y + \varepsilon) - v(x, y)}{\varepsilon} \geq \frac{1}{\varepsilon} \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} \hat{X}_t^{x, y + \varepsilon} \left[h'(\hat{X}_t^{x, y + \varepsilon}) - h'(\hat{X}_t^{x, y}) \right] dt \right] \\ &\quad + \frac{1}{\varepsilon} \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} (\hat{X}_t^{x, y + \varepsilon} - \hat{X}_t^{x, y}) \left[h'(\hat{X}_t^{x, y}) - (\rho - \beta_2 - \pi_t^y(g_2 - g_1)) \right] dt \right] \\ (4.29) \quad &\quad - \frac{1}{\varepsilon} |g_2 - g_1| \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} \hat{X}_t^{x, y + \varepsilon} (\pi_t^{y + \varepsilon} - \pi_t^y) dt \right] \\ &\geq \frac{1}{\varepsilon} \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} (\hat{X}_t^{x, y + \varepsilon} - \hat{X}_t^{x, y}) (\hat{X}_t^{x, y + \varepsilon} h''(L_t^\varepsilon) + h'(\hat{X}_t^{x, y})) dt \right] \\ &\quad - \frac{1}{\varepsilon} |g_2 - g_1| \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} \hat{X}_t^{x, y + \varepsilon} (\pi_t^{y + \varepsilon} - \pi_t^y) dt \right]. \end{aligned}$$

In the last inequality we have used that $\rho - \beta_2 - \pi_t^y(g_2 - g_1) \geq 0$, since $\rho > \beta_2$ by Assumption 4.2, that $g_2 - g_1 < 0$, and that $\hat{X}_t^{x, y + \varepsilon} \leq \hat{X}_t^{x, y}$.

Define now $\Delta\pi_t^y := \frac{1}{\varepsilon}(\pi_t^{y + \varepsilon} - \pi_t^y)$, $t \geq 0$, and notice that, by using the second equation in (4.2), we can write

$$d\Delta\pi_t^y = -(\lambda_1 + \lambda_2)\Delta\pi_t^y dt + \Delta\pi_t^y \left(1 - \pi_t^{y + \varepsilon} - \pi_t^y \right) \left[\frac{(g_2 - g_1)}{\sigma} dI_t + (\alpha_1 - \alpha_2) dI_t^1 \right], \quad t > 0,$$

with $\Delta\pi_0^y = 1$. With the help of Itô's formula, it can be easily shown that

$$(4.30) \quad \Delta\pi_t^y = \exp \left\{ -(\lambda_1 + \lambda_2)t - \theta^2 \int_0^t (1 - \pi_s^{y + \varepsilon} - \pi_s^y)^2 ds + \int_0^t (1 - \pi_s^{y + \varepsilon} - \pi_s^y) \left[\frac{(g_2 - g_1)}{\sigma} dI_s + (\alpha_1 - \alpha_2) dI_s^1 \right] \right\},$$

with $\theta^2 := \frac{1}{2} \left[\frac{(g_2 - g_1)^2}{\sigma^2} + (\alpha_1 - \alpha_2)^2 \right]$, solves the previous stochastic differential equation.

Also, by (4.3) and simple algebra,

$$(4.31) \quad \frac{1}{\varepsilon} (\hat{X}_t^{x, y + \varepsilon} - \hat{X}_t^{x, y}) = \hat{X}_t^{x, y} \left(\frac{e^{\varepsilon(g_2 - g_1) \int_0^t \Delta\pi_s^y ds} - 1}{\varepsilon} \right).$$

Employing the definition of $\Delta\pi_t^y$ and (4.31) in (4.29), and using that $\hat{X}_t^{x,y+\varepsilon} \leq \hat{X}_t^{x,y}$, one finds

$$(4.32) \quad 0 \geq \frac{v(x, y + \varepsilon) - v(x, y)}{\varepsilon} \geq \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} \hat{X}_t^{x,y} \left(\frac{e^{\varepsilon(g_2 - g_1) \int_0^t \Delta\pi_s^y ds} - 1}{\varepsilon} \right) \cdot \left(\hat{X}_t^{x,y} h''(L_t^\varepsilon) + h'(\hat{X}_t^{x,y}) \right) dt \right] - |g_2 - g_1| \mathbb{E} \left[\int_0^{\tau_\varepsilon^*} e^{-\rho t} \hat{X}_t^{x,y} \Delta\pi_t^y dt \right].$$

We now aim at taking limits as $\varepsilon \downarrow 0$ in (4.32). To this end, notice that $\Delta\pi_t^y \rightarrow Z_t^y$ a.s. for all $t \geq 0$, as $\varepsilon \downarrow 0$, where, by Theorem 39 in Chapter V.7 of [41], $(Z_t^y)_{t \geq 0}$ is the unique strong solution to

$$(4.33) \quad dZ_t^y = -(\lambda_1 + \lambda_2) Z_t^y dt + Z_t^y (1 - 2\pi_s^y) \left[\frac{(g_2 - g_1)}{\sigma} dI_t + (\alpha_1 - \alpha_2) dI_t^1 \right], \quad t > 0,$$

with $Z_0^y = 1$. Then, if we were allowed to invoke the dominated convergence theorem when taking limits as $\varepsilon \downarrow 0$ in (4.32), we would obtain that

$$(4.34) \quad 0 \geq v_y(x, y) \geq (g_2 - g_1) \mathbb{E} \left[\int_0^{\tau^*} e^{-\rho t} \hat{X}_t^{x,y} \left(\int_0^t Z_s^y ds \right) \left(\hat{X}_t^{x,y} h''(\hat{X}_t^{x,y}) + h'(\hat{X}_t^{x,y}) \right) dt \right] - |g_2 - g_1| \mathbb{E} \left[\int_0^{\tau^*} e^{-\rho t} \hat{X}_t^{x,y} Z_t^y dt \right],$$

upon recalling that $v \in C^2(\mathcal{C})$. Therefore, letting (x_o, y_o) be any arbitrary point belonging to $\partial\mathcal{C}$, by taking limits in (4.34) as $(x, y) \rightarrow (x_o, y_o)$, by the dominated convergence theorem and thanks to Proposition 4.7 we obtain that

$$0 \geq \limsup_{(x,y) \rightarrow (x_o, y_o) \in \partial\mathcal{C}} v_y(x, y) \geq \liminf_{(x,y) \rightarrow (x_o, y_o) \in \partial\mathcal{C}} v_y(x, y) \geq 0,$$

thus proving that v_y is continuous across $\partial\mathcal{C}$.

In order to complete the proof it thus only remains to show that the dominated convergence theorem can indeed be applied when taking limits as $\varepsilon \downarrow 0$ in (4.32). This is what we are going to show in the two following technical steps.

Step 1. To prove that the dominated convergence theorem can be invoked when taking $\varepsilon \downarrow 0$ in the first expectation on the right-hand side of (4.32), we set

$$\Lambda_\varepsilon := \int_0^{\tau_\varepsilon^*} e^{-\rho t} \hat{X}_t^{x,y} \left(\frac{e^{\varepsilon(g_2 - g_1) \int_0^t \Delta\pi_s^y ds} - 1}{\varepsilon} \right) \left(\hat{X}_t^{x,y} h''(L_t^\varepsilon) + h'(\hat{X}_t^{x,y}) \right) dt,$$

and we show that the family of random variables $\{\Lambda_\varepsilon, \varepsilon \in (0, 1 - y)\}$ is bounded in $L^2(\Omega, \mathcal{F}, \mathbb{P})$, hence uniformly integrable.

Notice that by Assumption 4.1-(ii) and the fact that $L_t^\varepsilon \leq \hat{X}_t^{x,y}$ a.s., one has a.s. for any $t \geq 0$

$$\hat{X}_t^{x,y} \left[\hat{X}_t^{x,y} h''(L_t^\varepsilon) + h'(\hat{X}_t^{x,y}) \right] \leq \hat{K} \left(1 + (\hat{X}_t^{x,y})^{\gamma \vee 2} \right),$$

for some constant $\hat{K} > 0$ (independent of ε), so that by Jensen's inequality

$$|\Lambda_\varepsilon|^2 \leq \frac{2\hat{K}^2}{\rho^2} \int_0^\infty \rho e^{-\rho t} \left(\frac{1 - e^{\varepsilon(g_2 - g_1) \int_0^t \Delta\pi_s^y ds}}{\varepsilon} \right)^2 \left(1 + (\hat{X}_t^{x,y})^{2\gamma \vee 4} \right) dt.$$

Then, taking expectations and applying Hölder's inequality

$$(4.35) \quad \mathbb{E} \left[|\Lambda_\varepsilon|^2 \right]^{\frac{1}{2}} \leq K' \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(\frac{1 - e^{\varepsilon(g_2 - g_1) \int_0^t \Delta\pi_s^y ds}}{\varepsilon} \right)^4 dt \right]^{\frac{1}{4}} \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(1 + (\hat{X}_t^{x,y})^{4\gamma \vee 8} \right) dt \right]^{\frac{1}{4}},$$

for some other constant $K' > 0$, independent of ε , that in the following will be varying from line to line.

The standard inequality $1 - e^{-x} \leq x$, with $x = \varepsilon(g_1 - g_2) \int_0^t \Delta\pi_s^y ds \geq 0$, allows us to continue from (4.35) and write

$$(4.36) \quad \mathbb{E} \left[|\Lambda_\varepsilon|^2 \right]^{\frac{1}{2}} \leq K' \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(\int_0^t \Delta\pi_s^y ds \right)^4 dt \right]^{\frac{1}{4}} \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(1 + (\hat{X}_t^{x,y})^{4(\gamma \vee 2)} \right) dt \right]^{\frac{1}{4}}.$$

We now treat the two expectations in (4.36) separately. First of all, notice that by Jensen's inequality

$$(4.37) \quad \left(\int_0^t \Delta \pi_s^y ds \right)^4 = \left(\frac{1}{t} \int_0^t t \Delta \pi_s^y ds \right)^4 \leq t^3 \int_0^t (\Delta \pi_s^y)^4 ds.$$

Second of all, thanks to the nonnegativity of $(\Delta \pi^y)^4$, we can invoke Fubini-Tonelli's theorem and using also (4.37), obtain

$$(4.38) \quad \begin{aligned} \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(\int_0^t \Delta \pi_s^y ds \right)^4 dt \right] &\leq \mathbb{E} \left[\int_0^\infty e^{-\rho t} t^3 \left(\int_0^t (\Delta \pi_s^y)^4 ds \right) dt \right] \\ &= \frac{1}{\rho^4} \int_0^\infty e^{-\rho s} (\rho^3 s^3 + 3\rho^2 s^2 + 6\rho s + 6) \mathbb{E} \left[(\Delta \pi_s^y)^4 \right] ds. \end{aligned}$$

We now aim at evaluating the expectation in the last integral above.

To accomplish that, notice that by applying Itô's formula to the process $\xi_t^y := (\Delta \pi_t^y)^4$, and using (4.30), we have for any $t > 0$

$$d\xi_t^y = \xi_t^y \left(-(\lambda_1 + \lambda_2) + 12\theta^2(1 - \pi_t^{y+\varepsilon} - \pi_t^y)^2 \right) dt + 4\xi_t^y (1 - \pi_t^{y+\varepsilon} - \pi_t^y) \left[\frac{(g_2 - g_1)}{\sigma} dI_s + (\alpha_1 - \alpha_2) dI_s^1 \right],$$

with $\xi_0^y = 1$ and $\theta^2 = \frac{1}{2} \left[\frac{(g_2 - g_1)^2}{\sigma^2} + (\alpha_1 - \alpha_2)^2 \right]$. Because $(1 - \pi_t^{y+\varepsilon} - \pi_t^y)^2 \leq 2$ a.s. for all $t \geq 0$, and

$$\xi_t^y = e^{-(\lambda_1 + \lambda_2)t + 12\theta^2 \int_0^t (1 - \pi_s^{y+\varepsilon} - \pi_s^y)^2 ds} M_t^y,$$

where $(M_t^y)_{t \geq 0}$ is an exponential martingale, it is easy to see that

$$(4.39) \quad \mathbb{E} \left[(\Delta \pi_t^y)^4 \right] \leq e^{-(\lambda_1 + \lambda_2)t + 24\theta^2 t}, \quad t \geq 0.$$

Using the latter estimate in (4.38), together with Assumption 4.2, we deduce that

$$(4.40) \quad \sup_{\varepsilon \in (0, 1-y)} \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(\int_0^t \Delta \pi_s^y ds \right)^4 dt \right] < \infty.$$

As for the second expectation in (4.36), Assumption 4.2 and standard estimates employing (4.3) (together with the fact that $(g_2 - g_1) \int_0^t \pi_s^y ds < 0$) guarantee that it is finite. Moreover, it is independent of ε . Combining this with (4.40) we thus find from (4.36) that $\sup_{\varepsilon \in (0, 1-y)} \mathbb{E} \left[|\Lambda_\varepsilon|^2 \right]^{\frac{1}{2}} < \infty$, thus implying that the family of random variables $\{\Lambda_\varepsilon, \varepsilon \in (0, 1-y)\}$ is bounded in $L^2(\Omega, \mathcal{F}, \mathbb{P})$, hence uniformly integrable.

Step 2. We consider the second expectation on the right-hand side of (4.32), and setting

$$\Xi_\varepsilon := \int_0^{\tau_\varepsilon^*} e^{-\rho t} \widehat{X}_t^{x,y} \Delta \pi_t^y dt,$$

we aim at proving that the family of random variables $\{\Xi_\varepsilon, \varepsilon \in (0, 1-y)\}$ is bounded in $L^2(\Omega, \mathcal{F}, \mathbb{P})$, hence uniformly integrable.

By Jensen's inequality first, and Hölder's inequality then, one finds that

$$(4.41) \quad \mathbb{E} \left[|\Xi_\varepsilon|^2 \right]^{\frac{1}{2}} \leq \widehat{K} \mathbb{E} \left[\int_0^\infty e^{-\rho t} (\widehat{X}_t^{x,y})^4 dt \right]^{\frac{1}{4}} \mathbb{E} \left[\int_0^\infty e^{-\rho t} (\Delta \pi_t^y)^4 dt \right]^{\frac{1}{4}},$$

for some $\widehat{K} > 0$, independent of ε .

The first expectation on the right-hand side of (4.41) is finite thanks to Assumption 4.2 and standard estimates employing (4.3) (together with the fact that $(g_2 - g_1) \int_0^t \pi_s^y ds < 0$). Moreover, it is independent of ε .

As for the second one, by interchanging expectation and time integral by Fubini-Tonelli's theorem, and using (4.39), we obtain

$$\mathbb{E} \left[\int_0^\infty e^{-\rho t} (\Delta \pi_t^y)^4 dt \right]^{\frac{1}{4}} \leq \frac{1}{(\rho + \lambda_1 + \lambda_2 - 24\theta^2)^{\frac{1}{4}}},$$

due to Assumption 4.2. We therefore conclude that (cf. (4.41)) $\sup_{\varepsilon \in (0, 1-y)} \mathbb{E} \left[|\Xi_\varepsilon|^2 \right]^{\frac{1}{2}} < \infty$, thus completing the proof. $\square$

The previous theorem in particular implies the so-called *smooth-fit* property, a well known optimality principle in optimal stopping theory. Moreover, by standard arguments based on the strong Markov property of $(\widehat{X}, \pi)$ (see Chapter III in [39]) it follows from the results collected so far that the couple (v, d) solves the free-boundary problem

$$(4.42) \quad \begin{cases} (\mathbb{L} - \rho)v(x, y) = -xh'(x) & \text{on } \mathcal{C}, \\ v(x, y) = x & \text{on } \mathcal{S}, \\ v_x(x, y) = 1 & \text{at } x = d(y), y \in (0, 1), \\ v_y(x, y) = 0 & \text{at } x = d(y), y \in (0, 1), \end{cases}$$

with $v \in C^2(\mathcal{C})$.

An important consequence of Theorem 4.8 is the following.

Proposition 4.9. *One has that $y \mapsto d(y)$ is continuous on $[0, 1]$.*

Proof. Define the probability measure $\widehat{\mathbb{P}}$ on $(\Omega, \mathcal{F})$ such that $\frac{d\widehat{\mathbb{P}}}{d\mathbb{P}} \Big|_{\mathcal{F}_t} = e^{-\frac{1}{2}\sigma^2 t + \sigma I_t}$, $t \geq 0$. Such a measure is equivalent to $\mathbb{P}$ on $\mathcal{F}_t$, and defining $\widehat{I}_t := I_t - \sigma t$, by Girsanov's theorem the latter is a standard $\mathbb{H}$ -Brownian motion under $\widehat{\mathbb{P}}$.

By a change of measure (see, e.g., Section 12 in Chapter IV of [39]) it is then not difficult to see that v as in (4.7) is such that $v(x, y) := x - \widehat{V}(x, y)$, where, for any $(x, y) \in \mathcal{O}$, we have set

$$(4.43) \quad \widehat{V}(x, y) := \sup_{\tau \geq 0} \widehat{\mathbb{E}}_{(x, y)} \left[\int_0^\tau e^{-(\rho - \beta_2)t + (g_2 - g_1) \int_0^t \pi_s ds} \widehat{H}(\widehat{X}_t, \pi_t) dt \right],$$

with $\widehat{H}(x, y) := (\rho - \beta_2 - (g_2 - g_1)y - h'(x))$. In (4.43) above $\widehat{\mathbb{E}}_{(x, y)}$ denotes the expectation conditioned on the fact that $(\widehat{X}_0, \pi_0) = (x, y)$ $\widehat{\mathbb{P}}$ -a.s. Since $\{(x, y) \in \mathcal{O} : v(x, y) \geq x\} = \{(x, y) \in \mathcal{O} : \widehat{V}(x, y) \leq 0\}$, $d(\cdot)$ is the optimal stopping boundary for the problem with value $\widehat{V}$ as well.

In order to prove the continuity of $d(\cdot)$, we now aim at applying Theorem 10 in [40] for problem (4.43). Notice that $\widehat{V}_x \leq 0$ on $\mathcal{O}$ since $x \mapsto h(x)$ is strictly convex. Moreover, recalling $\theta^2 = \frac{1}{2}[(\alpha_1 - \alpha_2)^2 + \frac{(g_2 - g_1)^2}{\sigma^2}]$, we have $\partial_x \left(\frac{\widehat{H}}{\theta^2 y^2 (1-y)^2} \right) < 0$ on $\mathcal{O}$ thanks, again, to the strict convexity of h . Also, $\widehat{V}_y$ is continuous across the boundary, due to the C^1 -property shown in Theorem 4.8 for $v = x - \widehat{V}$; hence, the horizontal smooth-fit property holds.

We can therefore apply Theorem 10 of [40] (upon noticing that in [40] x is the horizontal axis and y is the vertical one, while, in our paper, x is the vertical axis and y is the horizontal one), and conclude that d cannot have discontinuities of the first kind at any point $y \in [0, 1]$. Finally, d is also continuous at $y = 1$ since it is left-continuous by Proposition 4.5-(ii). $\square$

4.2. The Optimal Control for Problem (P3). In this section, we provide the form of the optimal debt reduction policy. It is given in terms of the free boundary studied in the previous section.

For d as in (4.14), introduce under $\mathbb{P}_{(x, y)}$ the nondecreasing process

$$(4.44) \quad \overline{\nu}_t^* = \left[x - \inf_{0 \leq s \leq t} \left(d(\pi_s) e^{-(\beta_2 - \frac{1}{2}\sigma^2)s - \sigma I_s - (g_2 - g_1) \int_0^s \pi_u du} \right) \right] \vee 0, \quad t \geq 0,$$

with $\overline{\nu}_{0-}^* = 0$, and then the process

$$(4.45) \quad \nu_t^* := \int_0^t e^{-(\beta_2 - \frac{1}{2}\sigma^2)s - \sigma I_s - (g_2 - g_1) \int_0^s \pi_u du} d\overline{\nu}_s^*, \quad t \geq 0, \quad \nu_{0-}^* = 0.$$

Notice that since $\overline{\nu}_t^* \leq x$ a.s. for all $t \geq 0$, and $t \mapsto \overline{\nu}_t^*$ is nondecreasing, it does follow from (4.45) that ν^* is admissible. Moreover, $t \mapsto \overline{\nu}_t^*$ is continuous (with the exception of a possible initial jump at initial time), due to the continuity of $y \mapsto d(y)$ and to that of $t \mapsto I_t$, $t \mapsto \pi_t$, and $t \mapsto \int_0^t \pi_s ds$.

Theorem 4.10. *Let $\tilde{V}(x, y) := \int_0^x \frac{1}{z} v(z, y) dz$, $(x, y) \in [0, \infty) \times [0, 1]$. Then one has $\tilde{V} = V$ on $[0, \infty) \times [0, 1]$, and ν^* as in (4.45) is optimal for Problem (P3).*

Proof. Recall $U = U_0$ as in (3.26), and notice that in our Markovian setting one has $\frac{1}{z} v(z, y) = U(z)$. By the proof of Theorem 3.13 it suffices to show that the right-continuous inverse of the stopping time $\tau^*(z, y) = \inf\{t \geq 0 \mid \hat{X}_t^{z, y} \geq d(\pi_t^y)\}$ (which is optimal for $v(z, y)$, cf. (4.12)) coincides (up to a null set) with $\bar{\nu}^*$.

Then, recall (3.34) from the proof of Theorem 3.13, fix $(x, y) \in (0, \infty) \times (0, 1)$, take $t \geq 0$ arbitrary, and notice that by (4.12) we have $P_{(z, y)}$ -a.s. the equivalences

$$\begin{aligned} \tau^*(z, y) \leq t &\iff \hat{X}_\theta \geq d(\pi_\theta) \text{ for some } \theta \in [0, t] \iff \\ z &\geq e^{-(\beta_2 - \frac{1}{2}\sigma^2)\theta - \sigma I_\theta - (g_2 - g_1) \int_0^\theta \pi_u du} d(\pi_\theta) \text{ for some } \theta \in [0, t] \iff \\ \left[x - \inf_{0 \leq s \leq t} \left(d(\pi_s) e^{-(\beta_2 - \frac{1}{2}\sigma^2)s - (g_2 - g_1) \int_0^s \pi_u du - \sigma I_s} \right) \right] \vee 0 &\geq x - z \iff \bar{\nu}_t^* \geq x - z \iff \\ \tau_+^{\bar{\nu}^*}(z) &\leq t. \end{aligned}$$

Hence, $\tau_+^{\bar{\nu}^*}(z) = \tau^*(z, y)$ a.s., and $\bar{\nu}^*$ is the right-continuous inverse of $\tau^*(\cdot, y)$. Since $\bar{\nu}^*$ is admissible, by arguing as in *Step 2* of the proof of Theorem 3.13 the claim follows. $\square$

Notice that the equation of X^{x, y, ν^*} in the formulation of Problem (P3), and (4.45), yield

$$X_t^{x, y, \nu^*} = e^{(\beta_2 - \frac{1}{2}\sigma^2)t + (g_2 - g_1) \int_0^t \pi_s^y ds + \sigma I_t} [x - \bar{\nu}_t^*],$$

which, with regard to (4.44), shows that

$$(4.46) \quad 0 \leq X_t^{x, y, \nu^*} \leq d(\pi_t^y), \quad t \geq 0, \mathbb{P} - a.s.$$

Moreover, it is easy to see that we can express $\bar{\nu}^*$ of (4.44) as

$$(4.47) \quad \bar{\nu}_t^* = \sup_{0 \leq u \leq t} \left(\frac{X_s^{x, y, 0} - d(\pi_s^y)}{X_s^{1, y, 0}} \right) \vee 0, \quad \bar{\nu}_{0-}^* = 0.$$

The previous equations allow us to make some remarks about the optimal debt management policy of our problem.

- (i) If at initial time the level of the debt ratio x is above $d(y)$, then an immediate lump sum reduction of amplitude $(x - d(y))$ is optimal.
- (ii) At any $t \geq 0$, it is optimal to keep the debt ratio level below the belief-dependent ceiling d .
- (iii) If the level of the debt ratio at time t is strictly below $d(\pi_t)$, there is no need for interventions. The government should intervene to reduce its debt only at those (random) times t at which the debt ratio attempts to rise above $d(\pi_t)$. These interventions are then minimal, in the sense that $(X^{x, y, \nu^*}, \pi^y, \nu^*)$ solves a Skorokhod reflection problem at the free boundary d .
- (iv) Recall that the debt ceiling d is an increasing function of the government's belief that the economy is enjoying a phase of fast growth. Then, with regard to the previous description of the optimal debt reduction rule, we have that the more the government believes that the economy is in a good shape, the more the fiscal space is, and the less strict the optimal debt reduction policy should be.

4.3. Regularity of the Value Function of Problem (P3) and Related HJB Equation.

Combining the results collected so far, we are now able to prove that the value function V of control Problem (P3) is a twice-continuously differentiable function. As a byproduct, V is a classical solution to the corresponding Hamilton-Jacobi-Bellman (HJB) equation.

By Theorem 4.10 we know that $V(x, y) = \int_0^x \frac{1}{z} v(z, y) dz$, for all $(x, y) \in \bar{\mathcal{O}} := [0, \infty) \times [0, 1]$. Hence, thanks to Theorem 4.8 and to the dominated convergence theorem, we immediately obtain the following result.

Lemma 4.11. *One has that $V \in C^1(\mathcal{O}) \cap C(\bar{\mathcal{O}})$. Moreover, $V_{xx} \in C(\mathcal{O})$, as well as $V_{xy} \in C(\mathcal{O})$.*

To take care of the second derivative V_{yy} , we determine the second weak derivative of V (recall that V_y is continuous by Theorem 4.8), and we then show that it is a continuous function. This is accomplished in the next proposition.

Proposition 4.12. *Let $\theta^2 := \frac{1}{2}[(\alpha_1 - \alpha_2)^2 + \frac{(g_2 - g_1)^2}{\sigma^2}]$. We have $V_{yy} \in C(\mathcal{O})$ with*

$$(4.48) \quad \begin{aligned} V_{yy}(x, y) = & -\frac{1}{\theta^2 y^2 (1-y)^2} \left[(\beta_2 + (g_2 - g_1)y - \frac{1}{2}\sigma^2) (v(x \wedge d(y), y) - v(0+, y)) \right. \\ & + h(x \wedge d(y)) + \frac{1}{2}\sigma^2 (x \wedge d(y)) v_x(x \wedge d(y), y) \left. \right] + \frac{(\lambda_2 - (\lambda_1 + \lambda_2)y)}{\theta^2 y^2 (1-y)^2} \left(\int_0^{x \wedge d(y)} \frac{1}{z} v_y(z, y) dz \right) \\ & - \frac{\rho}{\theta^2 y^2 (1-y)^2} \left(\int_0^{x \wedge d(y)} \frac{1}{z} v(z, y) dz \right). \end{aligned}$$

Proof. Notice that $V_y(x, y) = \int_0^x \frac{1}{z} v_y(z, y) dz$, and therefore $V_y(x, \cdot)$ is a continuous function for all $x > 0$ by Theorem 4.8 (notice indeed that by the bounds in (4.34) and the multiplicative dependence of $\widehat{X}^{z,y}$ with respect to z one has that $\frac{1}{z} v_y(z, y)$ is integrable at zero). Hence, its weak derivative with respect to y is a function $g \in L^1_{loc}(\mathcal{O})$ such that for any test function $\varphi \in C_c^\infty((0, 1))$ one has

$$(4.49) \quad \int_0^1 V_y(x, y) \varphi'(y) dy = - \int_0^1 g(x, y) \varphi(y) dy.$$

We now aim at evaluating g and at showing that it coincides with the right-hand side of (4.48).

Denote by $m(x)$, $x > 0$, the generalized right-continuous inverse of $d(y)$, $y \in [0, 1]$; that is, $m(x) := \inf\{y \in [0, 1] : d(y) \geq x\}$. Then, noticing that $v_y = 0$ on $\{(x, y) \in \mathcal{O} : x > d(y)\}$ and using Fubini's theorem, we can write

$$(4.50) \quad \begin{aligned} \int_0^1 V_y(x, y) \varphi'(y) dy &= \int_0^1 \left(\int_0^{x \wedge d(y)} \frac{1}{z} v_y(z, y) dz \right) \varphi'(y) dy \\ &= \int_0^x \frac{1}{z} \left(\int_{m(z)}^1 v_y(z, y) \varphi'(y) dy \right) dz \\ &= \int_0^x \frac{1}{z} \left[v_y(z, 1) \varphi(1) - v_y(z, m(z)) \varphi(m(z)) - \int_{m(z)}^1 v_{yy}(z, y) \varphi(y) dy \right] dz \\ &= - \int_0^x \frac{1}{z} \left(\int_{m(z)}^1 v_{yy}(z, y) \varphi(y) dy \right) dz, \end{aligned}$$

where we have used that $v_y(z, m(z)) = 0$ for all $z \in (0, x)$, $x > 0$, as well as $\varphi(1) = 0$.

By Lemma 4.6 (cf. also (4.20)), for any $y > m(z)$, for any $z \in (0, x)$ with $x > 0$, we have that

$$(4.51) \quad \begin{aligned} v_{yy}(z, y) = & \frac{1}{\theta^2 y^2 (1-y)^2} \left[\rho v(z, y) - (\lambda_2 - (\lambda_1 + \lambda_2)y) v_y(z, y) - z h'(z) \right. \\ & \left. - \frac{1}{2}\sigma^2 z^2 v_{xx}(z, y) - (\beta_2 + (g_2 - g_1)y) z v_x(z, y) \right]. \end{aligned}$$

Inserting the latter expression in the last integral term on the right-hand side of (4.50), using again Fubini's theorem and then integrating the derivatives with respect to x , we find

$$(4.52) \quad \begin{aligned} \int_0^1 V_y(x, y) \varphi'(y) dy &= - \int_0^x \frac{1}{z} \left(\int_{m(z)}^1 v_{yy}(z, y) \varphi(y) dy \right) dz \\ &= \int_0^1 \frac{(\lambda_2 - (\lambda_1 + \lambda_2)y)}{\theta^2 y^2 (1-y)^2} \left(\int_0^{x \wedge d(y)} \frac{1}{z} v_y(z, y) dz \right) \varphi(y) dy \\ &\quad - \int_0^1 \frac{\rho}{\theta^2 y^2 (1-y)^2} \left(\int_0^{x \wedge d(y)} \frac{1}{z} v(z, y) dz \right) \varphi(y) dy \\ &\quad + \int_0^1 \left[h(x \wedge d(y)) + (\beta_2 + (g_2 - g_1)y) (v(x \wedge d(y), y) - v(0+, y)) \right. \\ &\quad \left. + \frac{1}{2}\sigma^2 (x \wedge d(y)) v_x(x \wedge d(y), y) - \frac{1}{2}\sigma^2 (v(x \wedge d(y), y) - v(0+, y)) \right] \frac{\varphi(y)}{\theta^2 y^2 (1-y)^2} dy, \end{aligned}$$

where we have also used that $h(0) = 0$. Finally, setting

$$\begin{aligned} g(x, y) := & -\frac{1}{\theta^2 y^2 (1-y)^2} \left[h(x \wedge d(y)) + (\beta_2 + (g_2 - g_1)y - \frac{1}{2}\sigma^2)(v(x \wedge d(y), y) - v(0+, y)) \right. \\ & \left. + \frac{1}{2}\sigma^2(x \wedge d(y))v_x(x \wedge d(y), y) \right] + \frac{(\lambda_2 - (\lambda_1 + \lambda_2)y)}{\theta^2 y^2 (1-y)^2} \left(\int_0^{x \wedge d(y)} \frac{1}{z} v_y(z, y) dz \right) \\ & - \frac{\rho}{\theta^2 y^2 (1-y)^2} \left(\int_0^{x \wedge d(y)} \frac{1}{z} v(z, y) dz \right), \end{aligned}$$

we see that (4.52) reads $\int_0^1 V_y(x, y) \varphi'(y) dy = - \int_0^1 g(x, y) \varphi(y) dy$, so that g identifies with the second weak derivative of V with respect to y . Notice that g is continuous by the continuity of d , v , v_x , h , and the fact that $\int_0^{x \wedge d(y)} \frac{1}{z} v(z, y) dz$ and $\int_0^{x \wedge d(y)} \frac{1}{z} v_y(z, y) dz$ are finite due to (4.3), (4.4), and (4.34). The proof is therefore completed. $\square$

Thanks to Lemma 4.11 and Proposition 4.12 we have that $V \in C^2(\mathcal{O}) \cap C(\overline{\mathcal{O}})$. As a byproduct of this, by the Dynamic Programming Principle and standard means based on an application of Dynkin's formula, we obtain the next result.

Proposition 4.13. *Recall the second-order differential operator $\mathbb{L}$ defined in (4.20). The value function V of Problem (P3) is a classical solution to the HJB equation*

$$\min \{ (\mathbb{L} - \rho)V(x, y) + h(x), 1 - V(x, y) \} = 0, \quad (x, y) \in \mathcal{O},$$

with boundary condition $V(0, y) = 0$ for any $y \in [0, 1]$.

ACKNOWLEDGMENTS

The research of Claudia Ceci is partially supported by ‘‘Gruppo Nazionale per l’Analisi Matematica, la Probabilit  e le loro Applicazioni’’ (GNAMPA) of ‘‘Istituto Nazionale di Alta Matematica’’ (INdAM). Financial support by the German Research Foundation (DFG) through the Collaborative Research Centre 1283 is gratefully acknowledged by Giorgio Ferrari. We wish to thank Luciano Campi, Paola Mannucci, Fabio Paronetto, Paavo Salminen, and Wolfgang Runggaldier for useful discussions.

APPENDIX A. FILTERING RESULTS

Proof of Proposition 3.5. Since the innovation processes (I, I^1) (see (3.3)) and the random measure $m(dt, dq)$ (see (3.6) and (3.8)) are $\mathbb{H}$ -adapted, then $\mathbb{F}^I \vee \mathbb{F}^{I^1} \vee \mathbb{F}^m \subseteq \mathbb{H}$. In general, the latter inclusion could be strict. Let us now consider the exponential $\mathbb{F}$ -martingale solving

$$(A.1) \quad dL_t = -L_t \left\{ \frac{\beta(Z_t)}{\sigma} dW_t + \alpha(\eta_t, Z_t) dB_t \right\}, \quad t \geq 0,$$

and define a probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{F})$, equivalent to $\mathbb{P}$ on $\mathcal{F}_t$, and such that

$$\frac{d\mathbb{Q}}{d\mathbb{P}} \Big|_{\mathcal{F}_t} = L_t, \quad t \geq 0.$$

Notice that Assumption (3.1) ensures that L is indeed an $\mathbb{F}$ -martingale. By Girsanov's theorem, the processes

$$(A.2) \quad \widetilde{W}_t := W_t + \int_0^t \frac{\beta(Z_s)}{\sigma} ds, \quad \widetilde{B}_t := B_t + \int_0^t \alpha(\eta_s, Z_s) ds, \quad t \geq 0,$$

are $(\mathbb{Q}, \mathbb{F})$ -independent Brownian motions. We now prove that $\mathbb{F}^{\widetilde{W}} \vee \mathbb{F}^{\widetilde{B}} \vee \mathbb{F}^m = \mathbb{H}$. On the one hand, the inclusion $\mathbb{F}^{\widetilde{W}} \vee \mathbb{F}^{\widetilde{B}} \vee \mathbb{F}^m \subseteq \mathbb{H}$ follows from the fact that $\widetilde{W}$ and $\widetilde{B}$ turn out to be $\mathbb{H}$ -adapted since they can be written as

$$(A.3) \quad \widetilde{W}_t = I_t + \int_0^t \frac{\pi_s(\beta)}{\sigma} ds, \quad \widetilde{B}_t := I_t^1 + \int_0^t \pi_s(\alpha(\eta_s, \cdot)) ds, \quad t \geq 0.$$

To prove the converse, let us observe that, under the probability measure $\mathbb{Q}$, the process X^0 and η solve the following stochastic differential equations

$$dX_t^0 = X_t^0 \sigma d\widetilde{W}_t, \quad X_0^0 = x > 0,$$

$$(A.4) \quad d\eta_t = \sigma_1(\eta_t)d\widetilde{W}_t + \sigma_2(\eta_t)d\widetilde{B}_t + \int_{\mathbb{R}} qm(dt, dq), \quad \eta_0 = q \in \mathcal{I},$$

respectively. Clearly X^0 is $\mathbb{F}^{\widetilde{W}}$ -adapted. Recalling (3.16), the solution to equation (A.4) can be constructed iteratively. More precisely, $\forall t \in [0, T_1)$, the process η solves

$$d\eta_t = \sigma_1(\eta_t)d\widetilde{W}_t + \sigma_2(\eta_t)d\widetilde{B}_t, \quad \eta_0 = q \in \mathcal{I},$$

and for any time between two consecutive jump times, i.e. $t \in [T_n, T_{n+1})$, $n \geq 1$, one has

$$d\eta_t = \sigma_1(\eta_t)d\widetilde{W}_t + \sigma_2(\eta_t)d\widetilde{B}_t, \quad \eta_{T_n} = \eta_{T_n-} + \zeta_n.$$

By Assumption 2.1, this sequence stochastic differential equations has a unique strong solution on any interval $[T_n, T_{n+1})$, and this in turn gives the unique strong solution η to (A.4). Moreover, η turns out to be $F^{\widetilde{W}} \vee \mathbb{F}^{\widetilde{B}} \vee \mathbb{F}^m$ -adapted.

Then, by applying Corollary III.4.3.1 in [29], we have that every $(\mathbb{Q}, \mathbb{H})$ -local martingale $\widetilde{M}$ admits the decomposition

$$\widetilde{M}_t = \widetilde{M}_0 + \int_0^t \widetilde{\varphi}_s d\widetilde{W}_s + \int_0^t \widetilde{\psi}_s d\widetilde{B}_s + \int_0^t \int_{\mathbb{R}} \widetilde{w}(s, q) m^\pi(dt, dq),$$

where $\widetilde{\varphi}$ and $\widetilde{\psi}$ are $\mathbb{H}$ -adapted processes, and $\widetilde{w}$ is an $\mathbb{H}$ -predictable process indexed by $\mathbb{R}$, such that for all $t \geq 0$

$$\int_0^t \widetilde{\varphi}_s^2 ds < \infty, \quad \int_0^t \widetilde{\psi}_s^2 ds < \infty, \quad \int_0^t \int_{\mathbb{R}} |\widetilde{w}(s, q)| m^{p, \mathbb{H}}(dt, dq) < \infty \quad \mathbb{Q} - a.s.$$

Let now M be a $(\mathbb{P}, \mathbb{H})$ -local martingale, then $\widetilde{M} := M\widetilde{L}^{-1}$ is a $(\mathbb{Q}, \mathbb{H})$ -local martingale, where

$$\widetilde{L}_t := \mathbb{E}[L_t | \mathcal{H}_t] = \frac{d\mathbb{Q}}{d\mathbb{P}} \Big|_{\mathcal{H}_t}, \quad t \geq 0.$$

Taking into account (A.3), we have that $\widetilde{L}$ solves

$$d\widetilde{L}_t = -\widetilde{L}_t \left\{ \frac{\pi_t(\beta)}{\sigma} dI_t + \pi_t(\alpha(\eta_t, \cdot)) dI_t^1 \right\}, \quad \widetilde{L}_0 = 1,$$

and by applying the product formula to $M = \widetilde{M}\widetilde{L}$, we easily obtain that

$$\begin{aligned} dM_t &= \widetilde{M}_t d\widetilde{L}_t + \widetilde{L}_t d\widetilde{M}_t + d\langle \widetilde{M}^c, \widetilde{L}^c \rangle_t \\ &= (\widetilde{L}_t \widetilde{\varphi}_t - M_t \frac{\pi_t(\beta)}{\sigma}) dI_t + (\widetilde{L}_t \widetilde{\psi}_t - M_t \pi_t(\alpha(\eta_t, \cdot))) dI_t^1 + \int_{\mathbb{R}} \widetilde{w}(t, q) \widetilde{L}_t m^\pi(dt, dq). \end{aligned}$$

To conclude, we thus only need to set

$$\varphi_t := \widetilde{L}_t \widetilde{\varphi}_t - M_t \frac{\pi_t(\beta)}{\sigma}, \quad \psi_t := \widetilde{L}_t \widetilde{\psi}_t - M_t \pi_t(\alpha(\eta_t, \cdot)), \quad w(t, q) := \widetilde{w}(t, q) \widetilde{L}_t.$$

□

Proof of Theorem 3.6. In order to derive the filtering equation solved by $\pi_t = (\pi_t(i); i \in S)_{t \geq 0}$, we apply the innovation approach (see, for instance, Chapter IV in [4]). In this proof we shall use the two well-known facts:

- (i) for every $\mathbb{F}$ -martingale m , the projection over $\mathbb{H}$ is an $\mathbb{H}$ -martingale; that is, $\hat{m}_t := \mathbb{E}[m_t | \mathcal{H}_t]$, $t \geq 0$, is an $\mathbb{H}$ -martingale;
- (ii) for any $\mathbb{F}$ -progressively measurable and integrable process Ψ we have that

$$\mathbb{E} \left[\int_0^t \Psi_s ds \Big| \mathcal{H}_t \right] - \int_0^t \mathbb{E}[\Psi_s | \mathcal{H}_s] ds$$

is an $\mathbb{H}$ -martingale.

The first step of the innovation method consists in writing the process $\mathbb{1}_{\{Z_t=i\}}$, $i \in S$, as a semimartingale. Denoting by L^Z the Markov generator of the state process Z , we have that

$$L^Z f_i(j) = \sum_{k \in S} \lambda_{ki} f_k(j), \quad i, j \in S,$$

where $f_k(j) = \mathbb{1}_{\{j=k\}}$. Hence, for any $i \in S$, we can write

$$\mathbb{1}_{\{Z_t=i\}} = f_i(Z_t) = f_i(Z_0) + \int_0^t L^Z f_i(Z_s) ds + m_t(i),$$

where $(m_t(i))_{t \geq 0}$ is an $\mathbb{F}$ -martingale. By taking the conditional expectation with respect to $\mathcal{H}_t$, and using (i) and (ii) above, we obtain that

$$(A.5) \quad \pi_t(i) = y_i + \int_0^t \sum_{k \in S} \lambda_{ki} \pi_s(k) ds + M_t(i),$$

where $M(i)$ is an $\mathbb{H}$ -martingale null at zero. Proposition 3.5 ensures the existence of processes $\psi(i)$ and $\varphi(i)$ that are $\mathbb{H}$ -predictable, and w_i which is $\mathbb{H}$ -predictable and indexed by $\mathbb{R}$, such that

$$(A.6) \quad M_t(i) = \int_0^t \psi_s(i) dI_s + \int_0^t \varphi_s(i) dI_s^1 + \int_0^t \int_{\mathbb{R}} w_i(s, q) m^\pi(ds, dq),$$

with $\mathbb{E}[\int_0^t \varphi_s^2(i) ds] < \infty$, $\mathbb{E}[\int_0^t \psi_s^2(i) ds] < \infty$ and $\mathbb{E}[\int_0^t \int_{\mathbb{R}} |w_i(s, q)| m^{p, \mathbb{H}}(dt, dq)] < \infty$, $t \geq 0$. To obtain equation (3.14) it only remains to prove that

$$\begin{aligned} \psi_s(i) &= \pi_s(i) \sigma^{-1} \left\{ \beta_i - \sum_{j=1}^Q \beta_j \pi_s(j) \right\}, \quad \varphi_s(i) = \pi_s(i) \left\{ \alpha(\eta_s, i) - \sum_{j=1}^Q \alpha(\eta_s, j) \pi_s(j) \right\} \\ w_i(s, q) &= w_i^\pi(s, q) - \pi_{s-}(i) \end{aligned}$$

with w_i^π given in (3.15).

Following the same lines of the proof of Theorem 3.1 in [11] we can derive the structure of the processes $\psi(i)$, $\varphi(i)$ by imposing the following equalities

$$\forall i \in S, \quad \mathbb{E}[f_i(Z_t) \widetilde{W}_t | \mathcal{H}_t] = \pi_t(i) \widetilde{W}_t, \quad \mathbb{E}[f_i(Z_t) \widetilde{B}_t | \mathcal{H}_t] = \pi_t(i) \widetilde{B}_t,$$

where $\widetilde{W}$ and $\widetilde{B}$ are the $\mathbb{H}$ -Brownian motions defined in (A.2). To derive the expression of w_i , we consider a bounded process Γ of the form $\Gamma_t = \int_0^t \int_{\mathbb{R}} \gamma(s, q) m(ds, dq)$, with γ $\mathbb{H}$ -predictable process indexed by $\mathbb{R}$. Since Γ is $\mathbb{H}$ -adapted, the equality

$$(A.7) \quad \forall i \in S, \quad \mathbb{E}[f_i(Z_t) \Gamma_t | \mathcal{H}_t] = \pi_t(i) \Gamma_t$$

holds. By applying the product rule (taking into account no common jumps between Z and N) we obtain

$$d(f_i(Z_t) \Gamma_t) = f_i(Z_{t-}) d\Gamma_t + \Gamma_{t-} df_i(Z_t) = \Gamma_{t-} L^Z f_i(Z_t) dt + \int_{\mathbb{R}} f_i(Z_{t-}) \gamma(t, q) m^{p, \mathbb{F}}(dt, dq) + \mathcal{M}_t^{\mathbb{F}},$$

where $m^{p, \mathbb{F}}(dt, dq)$ is the $\mathbb{F}$ -dual predictable projection of $m(dt, dq)$ given in (3.12), and $\mathcal{M}^{\mathbb{F}}$ is an $\mathbb{F}$ -martingale. By projection onto $\mathcal{H}_t$, and denoting by $\mathcal{M}^{\mathbb{H}}$ an $\mathbb{H}$ -martingale, we have that

$$(A.8) \quad d\mathbb{E}[f_i(Z_t) \Gamma_t | \mathcal{H}_t] = \Gamma_t \pi_t(L^Z f_i) dt + \lambda^N(i) \pi_{t-}(i) \gamma(t, c(\eta_{t-}, i)) \mathbb{1}_{\{c(\eta_{t-}, i) \neq 0\}} dt + \mathcal{M}_t^{\mathbb{H}}.$$

On the other hand, the product rule and (A.5) and (A.6) yield

$$\begin{aligned} d(\pi_t(i) \Gamma_t) &= \pi_{t-}(i) d\Gamma_t + \Gamma_{t-} d\pi_t(i) + d\langle \pi(i), \Gamma \rangle_t \\ &= \Gamma_t \pi_t(L^Z f_i) dt + \Gamma_{t-} dM_t(i) + \int_{\mathbb{R}} \gamma(t, q) w_i(t, q) m(dt, dq). \end{aligned}$$

Recalling that $m^{p, \mathbb{H}}(dt, dq)$ is the $\mathbb{H}$ -dual predictable projection of $m(dt, dq)$ given in (3.11), we find

$$(A.9) \quad d(\pi_t(i) \Gamma_t) = \Gamma_t \pi_t(L^Z f_i) dt + \int_{\mathbb{R}} \gamma(t, q) w_i(t, q) m^{p, \mathbb{H}}(dt, dq) + \mathcal{M}_t^{\mathbb{H}},$$

where, again, $\mathcal{M}^{\mathbb{H}}$ is an $\mathbb{H}$ -martingale.

Gathering equations (A.7), (A.8), and (A.9), we obtain that for a.e. $t \geq 0$

$$\begin{aligned} & \lambda^N(i) \pi_{t-}(i) \gamma(t, c(\eta_{t-}, i)) \mathbb{1}_{\{c(\eta_{t-}, i) \neq 0\}} \\ &= \sum_{j=1}^Q \pi_{t-}(j) \lambda^N(j) \gamma(t, c(\eta_{t-}, j)) \mathbb{1}_{\{c(\eta_{t-}, j) \neq 0\}} (\pi_{t-}(i) + w_i(t, c(\eta_{t-}, j))). \end{aligned}$$

Choose now $\gamma(t, q)$ of the form $\gamma(t, q) = C_t \mathbb{1}_A(q) \mathbb{1}_{t \leq T_n}$, with C any bounded, $\mathbb{H}$ -predictable, positive process and $A \in \mathcal{B}(\mathbb{R})$. Observe that Γ is bounded since $|\Gamma_t| \leq \int_0^{t \wedge T_n} C_s dN_s \leq Dn$, with D a positive constant. Then the following equality holds on $\{t \leq T_n\}$

$$\forall \nu \in \mathcal{B}(\mathbb{R}), \quad \int_A \nu_t(i, dq) = \int_A (\pi_{t-}(i) + w_i(t, q)) \nu_t(dq),$$

where we have set

$$\nu_t(i, dq) := \lambda^N(i) \pi_{t-}(i) \mathbb{1}_{\{c(\eta_{t-}, i) \neq 0\}} \delta_{c(\eta_{t-}, i)}(dq), \quad \nu_t(dq) := \sum_{i=1}^Q \nu_t(i, dq).$$

Thus, on $\{t \leq T_n\}$,

$$w_i^\pi(t, q) = w_i(t, q) - \pi_{t-}(i) = \frac{d\nu_t(i, dq)}{\nu_t(dq)}, \quad \forall i \in S.$$

Finally, since the counting process N is nonexplosive, $T_n \uparrow \infty$ a.s. for $n \uparrow \infty$, and this yields (3.15). $\square$

Proof of Proposition 3.10. By Proposition 3.7, equations (3.14) and (3.21) are equivalent to a system of recursive equations between consecutive jump times, i.e. for $t \in [T_n, T_{n+1})$, $n = 0, 1, \dots$

$$\begin{aligned} \pi_t(i) &= \pi_{T_n}(i) + \int_{T_n}^t b^\pi(\underline{\pi}_s, \eta_s, i) ds + \int_{T_n}^t \sigma_1^\pi(\underline{\pi}_s, i) dI_s + \int_{T_n}^t \sigma_2^\pi(\underline{\pi}_s, i) dI_s^1, \quad i \in S, \\ \eta_t &= \eta_{T_n} + \int_{T_n}^t b^\eta(\underline{\pi}_s, \eta_s) ds + \int_{T_n}^t \sigma_1(\eta_s) dI_s + \int_{T_n}^t \sigma_2(\eta_s) dI_s^1, \end{aligned}$$

where we have set

$$\begin{aligned} b^\pi(\underline{y}, q, i) &:= \sum_{j=1}^Q \lambda_{ji} y_j - y_i \left[\lambda^N(i) \mathbb{1}_{\{c(q, i) \neq 0\}} - \sum_{j=1}^Q \lambda^N(j) y_j \mathbb{1}_{\{c(q, j) \neq 0\}} \right], \quad b^\eta(\underline{y}, q) := \sum_{j=1}^Q y_j b_1(q, j), \\ \sigma_1^\pi(\underline{y}, i) &:= \sigma^{-1} y_i \left\{ \beta_i - \sum_{j=1}^Q \beta_j y_j \right\}, \quad \sigma_2^\pi(\underline{y}, i) := y_i \left\{ \alpha(q, i) - \sum_{j=1}^Q \alpha(q, j) y_j \right\}, \end{aligned}$$

with the update at time T_n given by

$$(A.10) \quad \pi_{T_n}(i) = \frac{\lambda^N(i) \pi_{T_n^-}(i) \mathbb{1}_{\{\zeta_n = c(\eta_{T_n^-}, i)\}}}{\sum_{j=1}^Q \lambda^N(j) \pi_{T_n^-}(j) \mathbb{1}_{\{\zeta_n = c(\eta_{T_n^-}, j)\}}}, \quad i \in S, \quad \eta_{T_n} = \eta_{T_n^-} + \zeta_n.$$

Recall that, by assumption, the function $\alpha(q, i)$ given in (3.4) is locally-Lipschitz with respect to q and satisfies a (global) sublinear growth condition with respect to $q \in \mathcal{I}$, uniformly in $i \in S$.

We now develop the proof of uniqueness by distinguishing among three different cases related to the jumps' amplitude c ; namely, $c \neq 0$, $c = 0$, and $c \in \mathbb{R}$.

In the case $c \neq 0$, we have that $b^\pi(\underline{y}, q, i) := \sum_{j=1}^Q \lambda_{ji} y_j - y_i \left[\lambda^N(i) - \sum_{j=1}^Q \lambda^N(j) y_j \right]$, and it is easy to verify that between two consecutive jump times the pair $(\underline{\pi}, \eta)$ solves a $(Q+1)$ -dimensional stochastic differential equation with coefficients satisfying locally-Lipschitz conditions and (global) sublinear growth conditions with respect to $(\underline{y}, q) \in \mathcal{Y} \times \mathbb{R}$, uniformly in $i \in S$. As a consequence, strong uniqueness holds between two consecutive jump times; i.e. for $t \in [T_{n-1}, T_n)$, $n = 1, \dots$. Moreover, since the update at jump time T_n (see (A.10)) depends on the process $(\underline{\pi}_t, \eta_t)$ for $t \in [T_{n-1}, T_n)$, we have strong uniqueness of the solution to system (3.14) and (3.21) for all $t \geq 0$.

In the case $c = 0$, equations (3.14) and (3.21) reduce to

$$d\pi_t(i) = b^\pi(\underline{\pi}_t, \eta_t, i) dt + \sigma_1^\pi(\underline{\pi}_t, i) dI_t + \sigma_2^\pi(\underline{\pi}_t, i) dI_t^1, \quad i \in S, \quad t \geq 0,$$

$$d\eta_t = b^\eta(\pi_t, \eta_t)dt + \sigma_1(\eta_t)dI_t + \sigma_2(\eta_t)dI_t^1, \quad t \geq 0,$$

where, in particular, $b^\pi(y, q, i) = \sum_{j=1}^Q \lambda_{ji} y_j$. It is easy to check that also in this case strong uniqueness follows by the locally-Lipschitz property of the coefficients and by their (global) sublinear growth condition.

In the case $c \in \mathbb{R}$ the jumps' amplitude can assume any possible real value. In particular, c can be such that η and N do not have only common jumps: N might jump at a time at which $c(\eta_{t-}, Z_{t-}) = 0$, so that η does not jump at that time. The treatment of this case is more delicate and should be performed separately. Indeed, the uniqueness cannot be proved by using the arguments employed in the previous two cases because of the presence of $\mathbb{1}_{\{c(q,i) \neq 0\}}$ in the coefficient b^π which prevents to prove Lipschitz-continuity of b^π with respect to q . However, one might prove uniqueness by relying on the filtered martingale problem associated to the infinitesimal generator of the triplet (Z, X^0, η) . We refer to the seminal paper [35] and to Theorem 3.3 and Appendix B in the more recent [11]. $\square$

REFERENCES

- [1] ARONSON, D.G. (1967). Bounds for the Fundamental Solution of a Parabolic Equation. *Bull. Amer. Math. Soc.* **73(6)** pp. 890–896.
- [2] BALDURSSON, F.M., KARATZAS, I. (1997). Irreversible Investment and Industry Equilibrium. *Finance Stoch.* **1** pp. 69–89.
- [3] BORODIN, A.N., SALMINEN, P. (2015). *Handbook of Brownian Motion - Facts and Formulae* (2nd edition). Birkhäuser.
- [4] BRÉMAUD, P. (1980). *Point Processes and Queues: Martingale Dynamics*. Springer-Verlag.
- [5] CADENILLAS, A., HUAMÁN-AGUILAR, R. (2016). Explicit Formula for the Optimal Government Debt Ceiling. *Ann. Oper. Res.* **247(2)** pp. 415–449.
- [6] CADENILLAS, A., HUAMÁN-AGUILAR, R. (2018). On the Failure to Reach the Optimal Government Debt Ceiling. Forthcoming on *Risks*.
- [7] CECI, C., GERARDI, A. (1998). Partially Observed Control of a Markov Jump Process with Counting Observations: Equivalence with the Separated Problem. *Stoch. Process. Appl.* **78**, pp. 245–260.
- [8] CECI, C., GERARDI, A. (2000). Filtering of a Markov Jump Process with Counting Observations. *Appl. Math. Optim.* **42(1)**, pp. 1–18.
- [9] CECI, C., GERARDI, A. (2006). A Model for High Frequency Data under Partial Information: A Filtering Approach. *Int. J. Theoretical Appl. Finance* **9(4)**, pp. 1–22.
- [10] CECI, C. (2012). Utility Maximization with Intermediate Consumption under Restricted Information for Jump Market Models. *Int. J. Theoretical Appl. Finance* **15(6)** pp. 1–34.
- [11] CECI, C., COLANERI, K. (2012). Nonlinear Filtering for Jump Diffusion Observations. *Adv. Appl. Probab.* **44(3)**, pp. 678–701.
- [12] CECI, C., COLANERI, K. (2014). The Zakai Equation of Nonlinear Filtering for Jump-Diffusion Observation: Existence and Uniqueness. *Appl. Math. Optim.* **69(1)**, pp. 47–82.
- [13] CHRISTENSEN, S., CROCCE, F., MORDECKI, E., SALMINEN, P. (2018). On Optimal Stopping of Multidimensional Diffusions. [ArXiv:1611.00959](#). Forthcoming on *Stoch. Process. Appl.*
- [14] DE ANGELIS, T. (2018). Optimal Dividends with Partial Information and Stopping of a Degenerate Reflecting Diffusion. [ArXiv:1805.12035](#).
- [15] DE ANGELIS, T., FEDERICO, S., FERRARI, G. (2017). Optimal Boundary Surface for Irreversible Investment with Stochastic Costs. *Math. Oper. Res.* **42(4)**, pp. 1135–1161.
- [16] DE ANGELIS, T., PESKIR, G. (2018). Global C^1 Regularity of the Value Function in Optimal Stopping Problems. [ArXiv:1812.04564](#).
- [17] DYNKIN, E.B. (1965). *Markov Processes*, Vol. II. Springer.
- [18] DÉCAMPS, J.P., VILLENEUVE, S. (2015). Integrating Profitability Prospects and Cash Management. TSE Working Paper, n. 15-570.
- [19] ELLIOTT, R.J., AGGOUN, L., MOORE J.B. (1995). *Hidden Markov Models: Estimation and Control*, Springer.
- [20] EL KAROUI, N., HUU NGUYEN, D., JEANBLANC-PICQUÉ, M. (1988). Existence of an Optimal Markovian Filter for Control under Partial Observations. *SIAM J. Control Optim.* **26(5)** pp. 1025–1061.
- [21] EUROPEAN CENTRAL BANK (2013). Growth Effects of High Government Debt. *Monthly Bulletin March 2013*, pp. 82–84.
- [22] FERRARI, G. (2018). On the Optimal Management of Public Debt: a Singular Stochastic Control Problem. *SIAM J. Control Optim.* **56(3)** pp. 2036–2073.
- [23] FERRARI, G., RODOSTHENOUS, N. (2018). Optimal Management of Debt-to-GDP Ratio with Regime-Switching Interest Rate. [ArXiv:1808.01499](#).
- [24] FREY, R., SCHIMDT, T. (2012)- Pricing and Hedging of Credit Derivatives via the Innovation Approach to Nonlinear Filtering. *Finance Stoch.* **16**, pp. 105–133.

- [25] GIHMAN, I.J., SKOROHOD, A.V. (1972). *Stochastic Differential Equations*. Springer-Verlag.
- [26] GILBARG, D., TRUDINGER, N.S. (2001). *Elliptic Partial Differential Equations of the Second Order*. Springer.
- [27] HAMILTON, J.D. (1989). A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica* **57**(2) pp. 357–384.
- [28] JACOD, J. (1979). *Calcul Stochastique et Problèmes de Martingales*. Springer-Verlag.
- [29] JACOD, J., SHIRYAEV, A.N. (2003). *Limit Theorems for Stochastic Processes*. Springer-Verlag, Berlin Heidelberg.
- [30] JOHNSON, P., PESKIR, G. (2017). Quickest Detection Problems for Bessel Processes. *Ann. Appl. Probab.* **27**(2) pp. 1003–1056.
- [31] KALLIANPUR, G. (1980). *Stochastic Filtering Theory*. Springer-Verlag.
- [32] KARATZAS, I., SHREVE, S.E. (1991). *Brownian Motion and Stochastic Calculus* (Second Edition). Graduate Texts in Mathematics 113, Springer-Verlag, New York.
- [33] KARATZAS, I., SHREVE, S.E. (1998). *Methods of Mathematical Finance*. Applications of Mathematics (New York), 39. Springer-Verlag, New York.
- [34] KLIEMANN, W.H., KOCH, G., MARCHETTI F. (1990). On the Unnormalized Solution of the Filtering Problem with Counting Process Observations. *IEEE Transactions on Information Theory* **36**, pp. 1415–1425.
- [35] KURTZ, T.G., Ocone, D. (1988). Unique Characterization of Conditional Distributions in Nonlinear Filtering. *Ann. Probab.* **16**(1), pp. 80–107.
- [36] LIPTSER, R., SHIRYAEV, A.N. (2001). *Statistics of Random Processes. I. General Theory*. Springer-Verlag, Berlin Heidelberg.
- [37] MENALDI, J.L., ROBIN, M. (1983). On Optimal Correction Problems with Partial Information. *Stoch. Anal. Appl.* **3**(1), pp. 63–92.
- [38] ØKSENDAL, B., SULEM, A. (2012). Singular Stochastic Control and Optimal Stopping with Partial-Information of Itô-Lévy Processes. *SIAM J. Control Optim.* **50**(4), pp. 2254–2287.
- [39] PESKIR, G., SHIRYAEV, A.N. (2006). *Optimal Stopping and Free-Boundary Problems*. Lectures in Mathematics ETH, Birkhauser.
- [40] PESKIR, G. (2018). *Continuity of the Optimal Stopping Boundary for Two-Dimensional Diffusions*. Research Report No. 4 (2015), Probab. Statist. Group Manchester. Forthcoming on *Ann. Appl. Probab.*
- [41] PROTTER, P. (2004). *Stochastic Integration and Differential Equations*. Springer-Verlag, Berlin Heidelberg.
- [42] REVUZ, D., YOR, M. (1999). *Continuous Martingales and Brownian Motion*. Springer-Verlag, Berlin.
- [43] STOCK J.H., WATSON, M.W. (2003). How Did Leading Indicator Forecasts Perform During the 2001 Recession? *Federal Reserve Bank of Richmond - Economic Quarterly* **89**(3) pp. 71–90.
- [44] WOO, J., KUMAR, M.S. (2015). Public Debt and Growth. *Economica* **82**(328) pp. 705–739.
- [45] ZAKAI, M. (1969). On the Optimal Filtering of Diffusion Processes. *Z. Wahrscheinlichkeitstheor. Verw. Geb.* **11**(3), pp. 230–243.

G. CALLEGARO: DEPARTMENT OF MATHEMATICS “TULLIO LEVI-CIVITA”, UNIVERSITY OF PADOVA, VIA TRIESTE, 35121 PADOVA, ITALY

E-mail address: gcallega@math.unipd.it

C. CECI: DEPARTMENT OF ECONOMICS, UNIVERSITY “G. D’ANNUNZIO” OF CHIETI-PESCARA, VIALE PINDARO 42, I-65127 PESCARA, ITALY

E-mail address: claudia.ceci@unich.it

G. FERRARI: CENTER FOR MATHEMATICAL ECONOMICS (IMW), BIELEFELD UNIVERSITY, UNIVERSITÄTSSTRASSE 25, 33615 BIELEFELD, GERMANY

E-mail address: giorgio.ferrari@uni-bielefeld.de